Available online at www.sciencedirect.com
ScienceDirect
Lingua 137 (2013) 193–213
Lingua
www.elsevier.com/locate/lingua

Is diachronic lenition a factor in the asymmetry in intelligibility between Danish and Swedish?

Sandrien van Ommen [a,*], Petra Hendriks [b], Dicky Gilbers [b], Vincent van Heuven [c], Charlotte Gooskens [b]

[a] Utrecht University, The Netherlands
[b] University of Groningen, The Netherlands
[c] Leiden University, The Netherlands

Received 28 October 2011; received in revised form 17 September 2013; accepted 19 September 2013
Available online

Abstract

In this study, we investigate whether diachronic lenition is a factor in the previously found asymmetry in intelligibility between Danish and Swedish. Due to the historical process of consonantal lenition in Danish, the aspiration distinction between intervocalic, originally long, stops has disappeared. In two experiments, we tested the hypothesis that the absence of this distinction in Danish but not in Swedish results in better comprehension of Swedish intervocalic stops by Danish listeners than of Danish intervocalic stops by Swedish listeners. Our production experiment confirmed, not surprisingly, that there is no contrast for intervocalic stops in Danish, whereas there is one in Swedish. However, our perception experiment revealed that Danish listeners were not better in their perception of words with Swedish intervocalic stops than Swedish listeners were in their perception of words with Danish intervocalic stops. Contrary to the expectations of exemplar dynamics, in cross-language perception listeners did not always use their native categories in their perception of sounds of the non-native language and sometimes adjusted their native segment category boundaries to the values of the non-native language.

Keywords: Phonology; Intelligibility; Diachronic lenition; Danish; Swedish

1. Introduction

Danish and Swedish, two of the mainland Scandinavian languages, are rather similar. As a result, often speakers of these languages can communicate with each other using their own language while largely understanding the other language. However, the fact that these languages are very similar and mutually intelligible does not mean speakers understand each other to an equal extent. Previous studies (Maurud, 1976; Bø, 1978; Delsing and Lundin Åkesson, 2005; Gooskens, 2007; Gooskens et al., 2010) have consistently found that mutual intelligibility between Swedish and Danish is asymmetrical. In all of these studies the intelligibility of spoken Danish for Swedish listeners turned out to be lower than the intelligibility of spoken Swedish for Danish listeners. A still largely unresolved issue is *why* this relation is asymmetrical. Several different factors are likely to be involved. In the current study, we will investigate one of these: the production and perception of prevocalic and intervocalic stops in these languages.

Mutual intelligibility has been related to both linguistic and extra-linguistic factors. Linguistic factors are, for example, phonetic, lexical and morpho-syntactic distances between languages. In previous studies on Scandinavian languages

* Corresponding author. Tel.: +31 302536717.
E-mail address: s.vanommen@uu.nl (S. van Ommen).

0024-3841/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.lingua.2013.09.009

and dialects (Gooskens, 2007; Gooskens et al., 2008), linguistic distances have been shown to correlate with mutual intelligibility. Lexical and phonetic distances show strong correlations with scores on spoken mutual intelligibility tasks. In Gooskens (2007) the lexical distance was computed by counting the number of cognates (words with a common etymology). The lexical distance is large when the percentage of non-cognates is high. The phonetic distance between closely related languages also makes use of cognates, but in this case the difference between two pronunciations of a cognate is computed using the Levenshtein distance (see a.o. Levenshtein, 1965; Heeringa, 2004; Nerbonne and Heeringa, 2010), which is a string distance measure based on the number of phonetic differences between two cognates.

Linguistic distances are useful for quantifying differences between languages, but an intrinsic feature of distance that makes it an insufficient measure for explaining asymmetrical mutual intelligibility is the fact that distance is symmetrical.[1] The linguistic distance between Danish and Swedish is exactly as large as the linguistic distance between Swedish and Danish. Since mutual intelligibility between Swedish and Danish has been found to be asymmetrical, linguistic distances will not be able to provide us with a full explanation of the phenomenon of mutual intelligibility. In previous literature, extra-linguistic factors such as contact and attitude have also been investigated (see, e.g. Maurud, 1976; Delsing and Lundin Åkesson, 2005; Gooskens, 2006; Gooskens et al., 2010). These factors are hypothesized to influence mutual intelligibility asymmetrically, for example, because Swedish is more prevalent in the media than Danish (asymmetrical contact) and Danish is considered to be a less appealing language than Swedish (asymmetrical attitudes) (Gooskens, 2006). Gooskens (2006) investigated contact and attitude with respect to their role in the asymmetrical mutual intelligibility relation. However, she found only weak correlations, and Schüppert and Gooskens (2011) found no relation between attitude and intelligibility among Swedish and Danish children and adults. More recent mutual intelligibility experiments in which differences in language contact and attitudes were eliminated have confirmed the asymmetry in the cross-language intelligibility scores between Swedes and Danes (Gooskens et al., 2010). Thus, the existence of asymmetries in intelligibility remains a puzzle.

Because phonetic differences between languages may have a different impact on listeners from one language than on listeners from another language, it is conceivable that phonetic differences may actually be part of the explanation for the observed asymmetrical intelligibility. In the current study we hypothesize that cross-linguistic sound perception is influenced asymmetrically. We test this hypothesis on intervocalic stops. The expectation is that a historical loss of contrast in Danish intervocalic long stops impedes comprehension of Danish words with these consonants for Swedish listeners.[2] At the same time, Danish listeners are expected to be able to comprehend words with this contrast in Swedish, if they are able to interpret the phonetic manifestation of the stop consonants in this position in the non-native language. We will test these predictions in a production study (section 3) and a perception study (section 4), both with Danish and Swedish participants.

2. Theoretical background

There is reason to believe that phonetic differences between two languages need not affect intelligibility for both languages to an equal extent. Several phonetic differences between Danish and Swedish may be impeding Swedish listeners more than Danish listeners. An important role seems to be played by phonetic reduction. Danish, for example, has been shown to have a higher articulation rate than Swedish, whether in terms of the number of uttered words, or of underlying syllables per second (Hilton et al., 2011). A higher articulation rate by itself could already cause an asymmetrical mutual intelligibility relation. However, the higher articulation rate of Danish could also be a side effect of phonetic reduction, which would impede non-native perception even more. Hilton et al. found a higher articulation rate in terms of the number of assumed underlying (*phonological*) syllables per second, but when they looked at the number of actually realized (*phonetic*) syllables per second, this was lower in Danish than in Swedish. This is an indication that reduction processes are more pervasive in colloquial Danish than in Swedish. The frequent occurrence of elisions or

[1] The proportion of shared cognates by two languages (lexical distance) or the Levenshtein distance between cognates (phonetic distance) is symmetrical by definition. Asymmetrical distance measures have also been proposed, such as conditional entropy (Moberg et al., 2006, see p. 5 for an explanation of the concept) and the Phonological Correspondence Index (PCI; Tang & van Heuven, 2009; Cheng, 1997). The latter measure is based on the number and complexity of rules needed to convert a phonemic transcription of a sample of words in language A to their cognate counterparts in language B, and of a second set of rules that converts the transcriptions from B to A. Deriving these rule sets, which are often asymmetrical between language pairs, requires the availability of large computer-readable lists of phonemic transcriptions of lists of cognates in the languages at issue. Such resources were not available for the comparison of Danish and Swedish.

[2] It should be clear that these words would be ambiguous for Danish listeners, too. Ambiguities are common in any language, and this should not pose a problem to native speakers. The problem is expected to arise when listeners of a language in which the ambiguity does not exist (Swedish, in this case) try to comprehend forms in a language in which the ambiguity does exist (in this case: Danish). An anonymous reviewer, quite rightly, stressed that the existence of the ambiguity in Danish may prompt Danish listeners to ignore the contrast in Swedish, because it is not operative in that position in their native language, a point that we, in fact, turn to in the discussion.

reductions of whole syllables results in a lower number of realized syllables in spoken Danish than would be expected on the basis of the underlying forms. Phonetic reduction, in particular strength and saliency of segmentation cues as a result of phonetic reduction, has also been shown to result in a delay in the acquisition of past-tense forms for Danish children as compared to Swedish, Norwegian and Icelandic children (Bleses et al., 2011). This illustrates the influence that a seemingly low-level linguistic factor as reduction has on various linguistic tasks.

Another important factor in asymmetrical cross-linguistic perception may be orthography. The phonetic realization of words in Danish is not as straightforwardly reflected in their orthography as in Swedish (Elbro, 2006). This aspect may facilitate the intelligibility of Swedish for Danes in cognates. Since written language is more conservative than spoken language, and spoken Danish has diverged further from its North Germanic roots than Swedish (Elbro, 2006), Danish orthography is closer to Swedish spoken language than Swedish orthography is to Danish spoken language. This would allow Danes to benefit from orthography when mapping Swedish phonology to Danish words (Schüppert et al., in preparation; Doetjes and Gooskens, 2009), but less so the other way around.

Both the articulation rate and the orthography of Danish suggest that the phonological changes Danish has undergone may affect the mutual intelligibility of Danish and Swedish asymmetrically. One aspect of these changes is consonantal lenition. The phonological structure of Danish native (prototypical) disyllables is as follows: the first syllable is stressed and the final syllable has a schwa or non-full vowel. Before a full vowel, Danish has two stop series: the aspirated/affricated and the unaspirated/unaffricated series, which we will refer to as *fortis* and *lenis* stops, respectively. It is important to realize that we use fortis and lenis as phonological rather than phonetic terms, to describe the two members of a voicing or aspiration contrast. In the literature, the terms fortis and lenis have been used in various ways. Most commonly, the contrast between fortis and lenis sounds is seen as one of articulatory strength (Ladefoged and Maddieson, 1996). This notion is intended to capture the complex of measurable acoustic properties that distinguish sounds with a voicing or aspiration distinction, as voice or aspiration alone generally does not capture all characteristics of the contrast. Indeed, in languages with a voicing distinction, the ‘voiced’ members are often realized without vocal fold vibration. For example, in Dutch, a language with a clear voicing contrast, the contrast remains in situations without vocal fold vibration, such as in whispered speech (Slis and Cohen, 1969). Rather than referring to a separate contrast, the force of articulation is tightly linked to voice and aspiration contrasts: ‘‘Only a relatively small handful of languages have been proposed as possibly having articulatory strength differences that are independent of voicing’’ (Ladefoged and Maddieson, 1996). Thus we seem to be able to safely proceed from a *phonological* fortis–lenis distinction to refer to the aspirated/affricated stop (fortis) versus the unaspirated/unaffricated stop and in Swedish also the prevoiced stop (lenis) in prevocalic position, and to the aspirated and voiceless stop (fortis) versus the unaspirated and voiced stop (lenis) in intervocalic position.

An advantage of the phonological fortis–lenis distinction is that we can refer to members of a contrast independent of their syllabic position, since the syllabic position changes the phonetic realization of these segments. Furthermore, this distinction allows us to make a cross-linguistic comparison of phonetically different, but phonologically comparable, segments. We will describe the phonetic characteristics in the following section, distinguishing between phonological labels and phonetic descriptions. So, to summarize, without suggesting that Danish fortis stops are phonetically prototypical fortis sounds,[3] we group the Danish aspirated and affricated stops with the Swedish aspirated stops under the term ‘fortis’. These ‘fortis’ sounds are prevocalic and intervocalic /p/, /t/, /k/. On the other hand, we group the Danish unaspirated and unaffricated stop and the Swedish prevoiced and unaspirated stop (Helgason and Ringen, 2008) under the term ‘lenis’. These ‘lenis’ sounds, then, are prevocalic and intervocalic /b/, /d/, /g/.

As stated above, this fortis–lenis distinction exists in Danish in the position before a full vowel. We will refer to this ‘strong position’ as the *prevocalic* position. The inventory of intervocalic non-syllabic segments (the position before the schwa or non-full vowel) is identical to that found in word-final position (a.o. Basbøll, 2005), also called ‘weak position’ (Rischel, 1970, cf. Jakobson et al., 1961) and in this position, Danish has only one stop series. In the remainder of this paper, we will refer to this medial weak position as the *intervocalic* position. The fact that there is only one stop series in intervocalic position is the result of diachronic lenition, and the effect of this result on interlingual intelligibility is the focus of our study.

Consonantal lenition can be described as a ‘weakening’ of articulatory strength, reducing acoustic distinctiveness and articulatory effort (cf. Steriade, 1997; Flemming, 1995; Kirchner, 1998; Boersma, 1998). The diachronic lenition process we focus on in this study concerns both loss of phonological *strength* (a consonant becomes more vowel-like) and loss of phonological *contrast* (a minimal pair of sounds merges). The result of this historical process of lenition may be part of the explanation for the asymmetrical mutual intelligibility relationship between Danish and Swedish. This is because Danish developed differently over time than Swedish, yielding stop consonants that have no length or voicing opposition. The only contrast left in Danish is the aspiration contrast which is operative in initial, *strong*, position only. This means that stops in

[3] In fact, they may be considered phonetically lenis when it comes to force of articulation, even though they are aspirated or affricated (Fischer-Jørgensen, 1954).

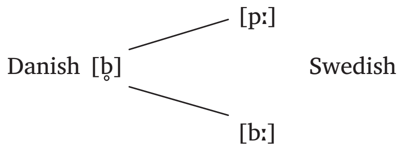

Fig. 1. Interlingual mapping of Danish (left) and Swedish(right) intervocalic stops.

medial or final, *weak*, position are non-contrastive. In the words of Fischer-Jørgensen (1954) there is 'free variation' between fortis and lenis stops in final position, and, before schwa, the normal manifestation is unaspirated and unaffricated. This is different in Swedish. Regarding stops in intervocalic position there is a length and voicing contrast in Swedish, but no such contrast in Danish. If a current speaker of Swedish or Danish produces a cognate such as 'lappar' (in Swedish)-'lapper' (in Danish) ('rags'), the resulting interlingual mapping of this word between Swedish and Danish is asymmetric, because the Danish word is a homophone with 'labber'. Fig. 1 illustrates this asymmetry.

As Fig. 1 shows, two Swedish long stops correspond to one Danish stop. As a result, the chance that Swedish listeners will map the Danish intervocalic stop to the correct Swedish intervocalic stop is p = .5 (for simplicity we do not take the frequency of words and sounds into account for this example). For Danish listeners, the chance is p = 1.0 that they will map the Swedish intervocalic stop to the correct Danish stop, as both the [pː] and the [bː] stop are mapped to [b̥] in Danish (see Moberg et al., 2006, for an elaboration on conditional entropy).

Even though there is no contrast in Danish *intervocalic* stops, Danish still has a contrast in *prevocalic* stops. This would possibly allow Danish listeners to perceive the Swedish intervocalic stops as contrastive, aided by knowledge of their native prevocalic stop contrast. However, phonetic differences between Danish and Swedish stops may interfere with this mapping. Phonetic differences are regarded both as differences in articulation (e.g. Best, 1995) and as differences in acoustics (e.g. Pierrehumbert, 2001, 2003). Because we aim to investigate the relation between the production of intervocalic stops in the native language and the cross-language perception of these sounds, we focus on acoustic rather than articulatory similarities and dissimilarities, as listeners have immediate access to acoustic information. The asymmetry we investigate could be caused by the inability of Swedish listeners to correctly classify Danish intervocalic stops due to acoustic dissimilarities. Pierrehumbert (2001, 2003) outlines the theory of exemplar dynamics for classifying speech sounds based on acoustic similarities and dissimilarities. According to this theory, a sound category is construed on the basis of the distinct examples of this category encountered by the language user, that is, on an 'exemplar cloud' in phonetic space. Each exemplar of a sound category has distinct but similar phonetic characteristics. The mapping of non-native phonemes to native phonemes will thus depend on the acoustic similarities and dissimilarities between the phonemes of the two languages. If an exemplar in one language is sufficiently similar to exemplars of a particular sound category in the other language, it will be classified as belonging to the same category. If not, it will be classified as belonging to a different category.

It should be added that the diachronic consonantal lenition of Danish intervocalic stops to approximants has also led to long vocalic sequences, which are likely to be hard to segment by Swedish listeners. This type of lenition does not exclusively occur in intervocalic position, as is illustrated by the words 'køb! mad, tog', which in Danish are pronounced as [kʰøʊ̯ˀ], [mað], [tˢoːˀ] and in Swedish as [cøːpʰ] [mɑːtʰ] [tʰuːkʰ]. Because it is hard to distinguish syllables in long vocalic sequences, this may result in a segmentation problem for Swedish listeners. However, this is a different phenomenon than the one we focus on in our study, which is a classification problem caused by differences in small phonetic detail.

Hypothesizing that listeners will use the exemplars from their native language to classify speech sounds from their native language as well as from a non-native language, our predictions for the production and comprehension of words with prevocalic and intervocalic stops in Danish and Swedish are the following.

(1) *Prevocalic* stops in Danish and Swedish are produced more or less similarly (although Danish prevocalic /t/ is affricated rather than purely aspirated), as a result of which cross-language classification is expected to be successful.
(2) The *intervocalic* stop contrast has disappeared in Danish, whereas it is present in Swedish. As a consequence, cross-language perception is asymmetrical: Swedish listeners are unsuccessful in the classification of words with Danish stops,[4] whereas Danish listeners differentiate Swedish words containing an intervocalic fortis–lenis contrast, possibly helped by their knowledge of the native prevocalic contrast.

This could then be part of the explanation for the observed asymmetry in intelligibility between Danish and Swedish. We test these hypotheses in two experiments: a production experiment investigating the production of Danish and Swedish

[4] See footnote 2.

Table 1
The eleven acoustic correlates used to identify fortis and lenis stops in Danish and Swedish.

In prevocalic position	In intervocalic position
1. Duration of release	5. Duration of release
2. Intensity of release	6. Intensity of release
3. Rise of following vowel	7. Rise of following vowel
4. Voice Onset Time (VOT)	8. Decay of preceding vowel
	9. Duration of preceding vowel
	10. Occlusion duration
	11. Intensity of voice bar during occlusion

words containing prevocalic or intervocalic stops by native speakers of these languages (Experiment 1), and a perception experiment investigating native and interlingual comprehension of the words that were produced in the production experiment (Experiment 2).

3. Production

3.1. Methods and design

Three Danish and three Swedish speakers were asked to read aloud a list of sentences in which 24 target items (words and non-words) were embedded. This list was interspersed with 38 sentences with monosyllabic and disyllabic fillers (words and non-words) with a different segmental structure than the target items. The fillers and target items were presented in random order, to prevent the speaker from using contrastive emphasis.[5] The items were judged to be natural-sounding by a native speaker of Danish. Words and non-words were presented in separate lists, which, again, contained fillers. Each speaker pronounced each word three times: the first time in a word list, and the second and third time in a sentence list. The sentences in which the items and fillers were embedded were of the following structure:

Danish	Jeg siger nu pakker igen. Pakker.
Swedish	Jag säger nu packar igen. Packar.
Translation	I now say ‘pakker’ again. Pakker.

The only version of the word used in the analysis was the word following the sentence. All speakers were male native speakers. They all spoke the standard language and originated from the capitals, i.e. Stockholm in Sweden or Copenhagen in Denmark. Of the 24 target items, 12 are Danish-Swedish cognates (historically related words), 6 are non-cognates (the six words in Danish are of different origin than the six words in Swedish) and 6 are non-words (see Appendix for all materials). Because the number of cognates with the desired segmental structure was very small, we added non-cognates and non-words to the analysis.

In the produced items, we then measured the eleven acoustic correlates in Table 1. These features are known to contribute to the fortis–lenis contrast. Three correlates are measured both in prevocalic and intervocalic position. These three correlates will be used to compare the expression of contrast in prevocalic and intervocalic position. If these correlates contribute significantly to the fortis–lenis contrast in prevocalic as well as in intervocalic position, and if each of these correlates contributes in approximately the same way to the contrast in Danish and Swedish, knowledge of the contrast in Danish prevocalic stops may aid Danish listeners in the perception of Swedish intervocalic stops.

Before we turn to the results of the measurements in section 3.2, we motivate why and how we measured the eleven acoustic correlates in Table 1.

Differences between fortis and lenis sounds are often regarded as a difference in voicing and ‘force of articulation’ (see, e.g. Malécot, 1966; Debrock, 1977), where fortis sounds are pronounced with more energy than lenis sounds. Fig. 2 is the spectrogram with the intensity contour of the word [lɑpːɑr] (containing an intervocalic fortis stop). The total amount of energy is the product of duration and intensity, that is, the area under the intensity contour. The two large peaks are the

[5] An anonymous reviewer pointed out that the carrier sentences may prompt a somewhat over-articulated production. We are aware of the fact that reading a list of carrier sentences is not the most natural production task. However, we believe the random order of the segmentally very different items has distracted the readers sufficiently from the contrasts we were interested in and a native speaker of Danish judged the items to be natural-sounding. Furthermore, if, despite these efforts, any over-articulated production occurred, we expect this to facilitate comprehension to an equal degree for Danish and Swedish listeners. A subsequently found lack of comprehension by a listener in the perception experiment would then be due to cross-linguistic differences rather than to an over-articulated production of the items.

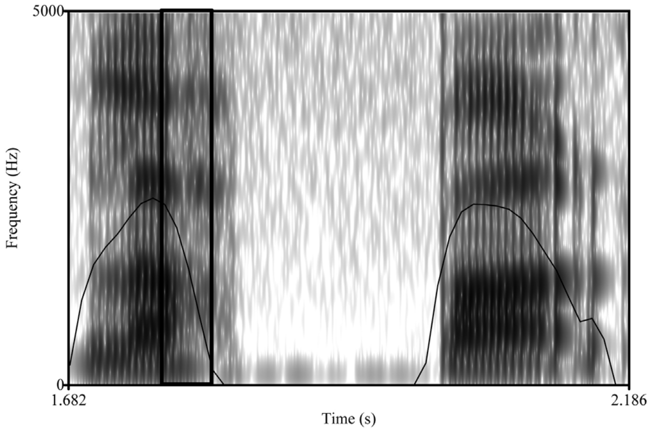

Fig. 2. Spectrogram with intensity contour of the Swedish word [lap:ar].

preceding and following vowel, the silent interval between the two vowels is the occlusion, and the small peak after the silent interval is the release. Fant (1973) identifies three stages of the release: the release transient, the fricative segment and the aspirative segment. In this study, two stages of the release will be measured: the release transient and the fricative segment. The aspirative segment (i.e. the voiceless initial part of the vowel, or, aspiration) is not considered as a part of the release in this study, because any frication in this stage is caused by a different articulator (glottal).

If force of articulation is important in the Danish and Swedish contrast, it is expected that the duration of the release of the fortis sound is longer than that of the lenis sound (correlates 1 and 5), or the intensity of the fortis sound is higher than that of the lenis sound (correlates 2 and 6), or, most likely, both. The two correlates are measured independently.

The difference in voicing is generally expressed in Voice Onset Time (VOT, correlate 4). This correlate is measured reliably in prevocalic position, but in intervocalic position, voicing is influenced by preceding sounds. This is why we decided to measure in prevocalic position only. Here, the VOT is measured from the start of the release to the point at which the vocal cords start to vibrate (if the vocal cords begin to vibrate before the release burst, VOT is negative).

The rise and decay in intensity of vowels (correlates 3, 7 and 8) are identified by Debrock (1977) as important correlates of the fortis–lenis contrast. In this study, recordings by Smalley (1961–1964) were used to investigate the intensity curves of vowels after and before fortis and lenis stops and fricatives. The rise of a vowel is that part of the curve starting from the end of the preceding consonant to the maximum intensity of the vowel (minus 10%, for reliability reasons). The decay of a vowel is the time from the maximum intensity (minus 10%) to the minimum (plus 10%). In Fig. 2, the black box around part of the first vowel shows the duration of the decay. The intervocalic fortis stop has a steeper intensity contour than its lenis counterpart does, resulting in a shorter decay. The same mechanism applies to the rise of the following vowel. When measuring the phonetic correlates rise and decay, we measured the duration of the rise and decay of the vowels along the horizontal axis.

The duration of the preceding vowel (correlate 9) has long been a generally accepted correlate of the intervocalic fortis–lenis contrast (Malécot, 1966; Maddieson and Gandour, 1977). The difference in duration of the preceding vowel has been considered to be due to some inherent property of the speech mechanism, since it has been found in many languages (see Cho and Ladefoged, 1999 for a discussion on phonetic universals). Lenis stops are prototypically voiced, fortis stops unvoiced. After a vowel, the more forcefully articulated stop will cut off the vowel earlier and faster and maintain its closure longer than the corresponding lenis stop. In some languages this apparent trade-off between vowel and closure duration may be considered inherent to the fortis–lenis difference, as in French. In other languages, speakers/listeners have caught on to the difference in vowel duration and exaggerate it, i.e. they lengthen the vowels before voiced stops disproportionately – to the effect that vowel duration becomes the auditory cue to the nature of the succeeding consonant, as in English.

Finally, the intensity of the voice bar (i.e. the band of low frequency voiced energy in the acoustic signal resulting from vibration of the vocal chords) during occlusion (correlate 11) is very likely to differ between fortis and lenis sounds. Fortis sounds will not be actively voiced (or will be actively devoiced) during this silent interval, causing the intensity to reach a lower level than it would in lenis sounds, which are actively voiced (or not actively devoiced) (Jansen, 2004). This correlate can be seen as an intervocalic alternative to VOT. To control for variation in our recordings and variation by the speaker, we related the intensity of release and the intensity of the voice bar during occlusion to the intensity of a word in the preceding carrier phrase, by subtracting the latter from the first. These two forms are pronounced by the same speaker on the same expiration. The selected word in the carrier phrase is the word *nu* 'now' in Danish (as in: *Jeg siger <u>nu</u> pakker igen. <u>Pakker</u>*) and in Swedish (as in: *Jag säger <u>nu</u> packar igen. <u>Packar</u>*).

3.2. Results and discussion

First, we determined whether the eleven correlates contributed significantly to the production of the contrast in prevocalic and intervocalic position. To this end, we analyzed the data using mixed-effects modeling (Baayen, 2008; Quené and van den Bergh, 2008) with restricted maximum likelihood estimations, in which each phonetic correlate is a dependent variable. Fixed factors are category (fortis and lenis), language (Danish and Swedish), and position (prevocalic and intervocalic) for those correlates that were measured in both positions. Item and speaker are crossed random factors, except in the models in which these factors proved to be redundant, in which case they were removed from the model. We employ a significance level of 5% and we report the actual degrees of freedom (rounded to the nearest integer) that were used in the statistical test. The significant results per phonetic correlate are discussed in the order in which they are given in Table 1. As a first analysis revealed that there were no significant differences between words and non-words, we combined the words and non-words in our further analyses. In Table 2, the means are given per combination of phonetic correlate, position, category and language of speaker that is compared in the mixed model. In intervocalic position, the originally long alveolar stop in Danish has lenited to a semivowel and thus is not relevant for the discussion of the production and perception of stops. Therefore, the Danish words with intervocalic [d] and [ð] have not been analyzed and will not be discussed in the results. The fact that alveolars are not included in the analyses in intervocalic position in Danish, but are retained in the other positions may introduce a slight skewing of the data, since stops in different places of articulation have intrinsically different durations. However, we believe that the advantage of retaining the data in the other positions outweighs the, most probably negligible, effect of this skewing of the data on the results.

Let us first look at the three correlates measured in *prevocalic* as well as *intervocalic* position. The analysis reveals that *duration of release* (correlates 1 and 5) is significantly different for position ($F(1,120) = 8.31$, $p = .005$) and category ($F(1,120) = 4.17$, $p < .05$, but not for language. There are, however, significant interactions for position and category ($F(1,120) = 6.26$, $p < .05$), position and language ($F(1,120) = 23.50$, $p = .000$), category and language ($F(1,120) = 24.69$, $p = .000$) and an interaction for category, position and language ($F(1,120) = 22.71$, $p = .000$). These interaction effects are largely caused by the deviant production of release duration in Danish prevocalic fortis stops compared to the other sounds. The Danish speakers in our production experiment pronounced the prevocalic /t/ with affrication, which explains the higher duration of release. This affrication is typical for Copenhagen Danish. For Swedish an unexpected pattern emerges, namely a longer release duration for lenis than fortis stops in prevocalic position. A closer look at the data reveals that this effect is due to two data points. We have only few data points in prevocalic position and we would need more data points to find out whether this is deviant behavior. The remaining data points for Swedish prevocalic stops suggest no difference in release duration between fortis and lenis stops. For *intensity of release* (correlates 2 and 6), a significant difference was found for position ($F(1,36) = 13.87$, $p = .001$). None of the other effects were significant. *Rise of the following vowel* (correlates 3 and 7), too, is significantly different for position only ($F(1,120) = 60.96$, $p = .000$).

The correlate measured in *prevocalic* position only, *VOT* (correlate 4), is significantly different for category ($F(1,8) = 31.31$, $p = .001$), and a trend toward a difference between languages ($F(1,6) = 5.90$, $p = .051$) but there is no significant interaction of category and language. An inspection of the VOT values in Danish and Swedish shows that this lack of interaction means that for both languages the effect has the same direction.

In *intervocalic* position, *decay of the preceding vowel* (correlate 8), is significantly different for language ($F(1,92) = 8.79$, $p < .005$) and there is a significant interaction of category*language ($F(1,92) = 4{,}53$, $p < .05$). Furthermore, the *duration of the preceding vowel* (correlate 9) is significantly different for language only ($F(1,6) = 32.42$, $p = .001$); the Swedish vowels are shorter than the Danish vowels, which could be due to the fact that the vowels are part of open syllables in Danish, but of closed syllables in Swedish. The *occlusion duration* (correlate 10) proves to be significantly different for category ($F(1,25) = 21.81$, $p = .000$), language ($F(1,10) = 381.58$, $p = .000$) and there is a significant interaction of category*language ($F(1,86) = 44{,}21$, $p = .000$). The last correlate in intervocalic position, the *intensity of the voice bar* (correlate 11), has significant effects for category ($F(1,22) = 163.33$, $p = .000$), language ($F(1,4) = 14.67$, $p < .05$) and an interaction of category*language ($F(1,54) = 166.82$, $p = .000$).

These analyses show that only three of the eleven measured phonetic correlates do not contribute to the fortis–lenis contrast in any position: intensity of release, rise of the following vowel and duration of the preceding vowel. For these correlates, there is no significant effect of category, nor any significant interaction involving category. The remaining correlates do not only significantly contribute to the fortis–lenis contrast (either in one of the languages, or in both) but also reveal an effect of language or an interaction of category and language. This means that for each of the correlates that contribute significantly to the fortis–lenis contrast, the pronunciation of this contrast differs significantly between the two languages. Whether this difference impedes perception of the contrast for listeners is a question which will be answered in Experiment 2.

Importantly, there is a significant effect of position for all three correlates measured both in prevocalic and intervocalic position (duration of release, intensity of release and rise of the following vowel). This means that listeners are not very likely to use these correlates in perception to classify intervocalic stops by using knowledge of prevocalic stops, but on the

Table 2
Significant effects at the 5% level, means and standard errors (SE) of the two and three-way interactions in a multilevel analysis of the production of prevocalic and intervocalic stops.

Phonetic correlate	Significant effects	Position	Category[a]	Language of speaker	EM means	*N*	SE
Duration of release (ms)	Position	Prevocalic	Lenis	Danish	22	9	4
	Voice			Swedish	28	9	4
	Position*voice		Fortis	Danish	53	9	4
	Position*language			Swedish	17	9	4
	Voice*language	Intervocalic	Lenis	Danish	21	18	3
	Position*voice*language			Swedish	27	27	2
			Fortis	Danish	20	18	3
				Swedish	26	27	2
Intensity of release (dB)[b]	Position	Prevocalic	Lenis	Danish	−19.9	9	3.7
				Swedish	−12.7	9	3.7
			Fortis	Danish	−17.5	9	3.7
				Swedish	−16.9	9	3.7
		Intervocalic	Lenis	Danish	−8.6	18	3.1
				Swedish	−12.2	27	2.9
			Fortis	Danish	−10.6	18	3.1
				Swedish	−11.0	27	2.9
Rise of following vowel (ms)	Position	Prevocalic	Lenis	Danish	44	9	5
				Swedish	28	9	5
			Fortis	Danish	35	9	5
				Swedish	36	9	5
		Intervocalic	Lenis	Danish	17	18	4
				Swedish	16	27	3
			Fortis	Danish	17	18	4
				Swedish	16	27	3
Voice Onset Time (ms)	Voice	Prevocalic	Lenis	Danish	37	9	9
	Language (trend: p = .051)			Swedish	15	9	9
			Fortis	Danish	76	9	9
				Swedish	49	9	9
Decay of preceding vowel (ms)	Language	Intervocalic	Lenis	Danish	54	18	5
	Voice*language			Swedish	58	27	4
			Fortis	Danish	51	18	5
				Swedish	73	27	4
Duration of preceding vowel (ms)	Language	Intervocalic	Lenis	Danish	128	18	8
				Swedish	89	27	7
			Fortis	Danish	133	18	8
				Swedish	80	27	7
Occlusion duration (ms)	Voice	Intervocalic	Lenis	Danish	44	18	7
	Language			Swedish	134	27	6
	Voice*language		Fortis	Danish	45	18	7
				Swedish	192	27	6
Intensity of voice bar during occlusion (dB)	Voice	Intervocalic	Lenis	Danish	−12.0	18	1.9
	Language			Swedish	−11.8	27	1.9
	Voice*language		Fortis	Danish	−12.6	18	1.9
				Swedish	−32.3	27	1.9

Durations are in milliseconds, intensities in decibels.

[a] With Danish sounds in intervocalic position, the terms ‘fortis’ and ‘lenis’ both refer to a lenis sound, due to lenition. The terms are used in this table to distinguish between orthographic words with <bb>, <dd> or <gg> and orthographic words with <pp>, <tt> or <kk>.

[b] The intensity values are negative because they represent the intensity of the target minus the intensity of an earlier word in the same sentence (the word *nu* ‘now’). This way the intensity is not dependent on the speaker or the level of the recordings. The reason why the difference is negative is that the earlier word *nu* consists of a nasal and a vowel, which have a higher intensity than a voiced or voiceless occlusion.

other hand, the fact that acoustic properties are of a different magnitude before stressed and unstressed vowels, does not normally hinder correct identification: Listeners are also speakers and know implicitly what to expect from native consonants under different conditions of stress. However, as stated before, there are also significant main and interaction effects of language. This makes it unlikely that Danish listeners use their knowledge of Danish prevocalic stops in their classification of Swedish intervocalic stops. In fact, the correlates that do not show effects of language are intensity of release and rise of the following vowel, but these correlates reveal no significant effect of category either, meaning that they do not contribute to expressing the contrast between fortis and lenis stops in either language.

Taking a closer look at the means and standard errors of Danish intervocalic stops, we can conclude that, as expected, there is no trace of a contrast left in Danish. All effects of category and language found in intervocalic position are due to the fortis–lenis contrast in Swedish and to the difference between Danish and Swedish. More specifically, the Swedish intervocalic fortis stop has a highly different mean from the other sounds in intensity of the voice bar (lower intensity) and duration of decay (longer decay).[6] This means that in these correlates, the Swedish lenis stop groups with the Danish intervocalic stops. The means of the duration of the occlusion of Swedish fortis and lenis stops are both higher than those of Danish stops. This, too, is as expected, because the Swedish stops are phonetically long, and the Danish are not (anymore).

On the basis of the observed differences in the production of Danish and Swedish prevocalic and intervocalic stops in Experiment 1, we can make predictions for cross-language classification. Our analysis of the produced forms showed that the *prevocalic* fortis–lenis contrast is produced phonetically differently in Danish and Swedish. Swedish fortis and lenis stops are both closer to Danish lenis stops than to Danish fortis stops. This means that from the perspective of a Danish listener, Swedish stops may not be sufficiently contrastive. In particular, they may have problems with the classification of the Swedish fortis stop. From the perspective of a Swedish listener, the Danish fortis stop is more similar to the Swedish fortis stop than to the Swedish lenis stop. Hence, Swedish listeners may succeed in correctly classifying the Danish fortis stop. On the other hand, the Danish lenis stop is more similar to the Swedish fortis stop, too, which means Swedish listeners may perform poorly in the classification of Danish lenis stops.

In *intervocalic* position, the contrast between Swedish stops is phonetically very salient, whereas there is no contrast in Danish. The Swedish lenis stop groups together with all Danish stops in most correlates significantly contributing to the contrast, although there is a large difference in occlusion duration in the two languages. This latter difference is not surprising, as the Swedish stops are long, whereas the Danish stops are short. From the perspective of a Swedish listener, Danish stops could all be categorized as Swedish lenis stops. From the perspective of a Danish listener, the difference between Swedish stops is acoustically salient but both categories are dissimilar to native Danish stops, which is in large part due to the long duration of the Swedish stops as compared to the Danish stops.

We hypothesize that listeners use the exemplar clouds of native sounds in perception. This would mean that the boundaries of the native categories coincide with the exemplars given for each category. In this case, the information about the exemplar clouds we have are the productions of the words that we elicited and the boundaries between categories are the points between these clouds. If a new word containing this sound in the same segmental structure would fall between the categories, we assume that the categorization of these sounds would be based on the proximity of the sound to one of the categories. So, if listeners base their classification of non-native sounds on the acoustic properties of the sounds in their native language, and on these properties only, we predict the following results for a comprehension experiment[7]:
Prevocalic stops:

(1) Comprehension of words with native prevocalic stops is excellent for Swedish and Danish.
(2) Swedish comprehension of Danish words with prevocalic fortis stops is moderately good, but Swedish comprehension of Danish words with prevocalic lenis stops is at chance level.
(3) Danish comprehension of Swedish words with prevocalic stops is poor, as all prevocalic stops will be classified as lenis.

Intervocalic stops:

(1) Swedish comprehension of native words with intervocalic stops is excellent.
(2) Danish comprehension of native words with intervocalic stops is at chance level.
(3) Swedish comprehension of Danish words with intervocalic stops is poor, as most or all Danish stops will be classified as lenis.

[6] This result is surprising, as rise and decay times are normally found to be shorter before and after fortis sounds than before and after lenis sounds (Debrock, 1977). As Helgason and Ringen (2008) found previously and as discussed by e.g. Kortlandt (2003), some Swedish speakers tend to preaspirate word-medial or word-final fortis stops: they utter a short [h]-like sound before occlusion (i.e. after the vowel). The speakers who participated in the current study tended to do so as well, be it to a variable extent. The decay of the vowel was measured from the highest point of intensity to the lowest point of intensity, which includes this pre-aspiration. As intervocalic lenis stops are not pre-aspirated, this latter decay is shorter. The same is true for the duration of the preceding vowel, which is measured up until the lowest point of intensity, thereby including occasional pre-aspiration and therefore lengthening this duration before fortis stops. This is most likely the reason that no difference was found in preceding vowel duration between intervocalic fortis and lenis stops in Swedish. Even though the articulation of pre-aspiration is variable across speakers, it may be a good option to measure the role of this correlate in production and perception of intervocalic stops in Swedish in future endeavors.

[7] Of course, alternative assumptions are possible. Alternative assumptions would be that the cross-over point is defined differently for the two languages, or that the proximity to the exemplar clouds is not linear, but based on the weight of a category. These are possibilities we have not investigated. We do not pretend to have recreated the exemplar clouds of our listeners, either, as this would be an impossible task. It would, however, be interesting to investigate larger databases than the current one to investigate questions of this latter kind.

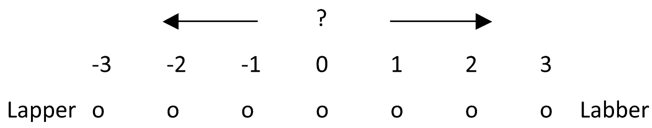

Fig. 3. Item in the auditory classification task (taken from the Danish form). Participants had to indicate on a 7-point confidence scale which word they thought they heard.

(4) Danish comprehension of Swedish words with intervocalic stops is either poor because the sounds are dissimilar from Danish sounds and Danish listeners probably are not aided by their knowledge of the prevocalic fortis–lenis contrast, or good because the Swedish intervocalic stops are phonetically very distinct. A third possibility is that Danish listeners do not pay attention to a contrast that has no role in that position in their native language.

In the next section, we discuss our second experiment, which investigates the perception by Danish and Swedish listeners of the Danish and Swedish stops that were produced in the production experiment.

4. Perception

4.1. Methods and design

Two schools, one in Denmark (Tønder) and one in Sweden (Kungälv), were asked to participate in the perception experiment. At each school two groups with young adults in the age between 16 and 20 years old participated, resulting in 40 Danish participants (*F*: 29; M: 11; mean age: 17.0) and 37 Swedish participants (*F*: 23; M: 14; mean age: 17.4). The schools were chosen for their comparable remoteness from the neighboring country: Tønder is approximately 270 km from Sweden and Kungälv is approximately 260 km from Denmark. In this way we hoped to exclude participants who had more than average experience with the neighboring language. This distance is, at the same time, the distance to the respective capitals.[8]

All participants spoke Danish and Swedish, respectively, as their native language. None of the participants had any known hearing problems. An auditory translation task showed that none of the participants was proficient in the neighboring language at the time of testing: the Danish participants were not proficient in Swedish (4 out of 40 participants correctly translated one out of five words, the other participants did not translate any of the words correctly) and the Swedish participants were not proficient in Danish either (none of the participants translated any of the words correctly). These results were consistent with the level of proficiency in the neighboring language and amount of contact with the language as reported by the participants themselves, in a post-test questionnaire. The participants were in a class-room setting, listening over loudspeakers. The session lasted 8 min and started with 3 training items.

The experiment consisted of an auditory classification task. The participants listened to 18 target items (the 12 cognates and the 6 non-words, see Appendix), pronounced by the six speakers in Experiment 1 (three Danish and three Swedish speakers), rendering a total of 72 words and 36 non-words. The audio file started with a three-word practice item in the native language, followed by six blocks of twelve words, and finally three blocks of twelve non-words. While listening to the audio file, the participants had a form in front of them to fill in. All instructions were given in the native language. An item on the form is as in Fig. 3. The participants were asked to tick the word they thought they heard, thereby also indicating the level of confidence in their choice. The word forms including fortis and lenis stops were presented as the two extremes on a 7-point confidence scale. When very sure, the participants were to choose the extreme. When completely unsure, participants were to choose the middle option.

Swedish and Danish pronunciations were randomized and presented to the participants without prior notice of the language in which the item was pronounced. All items on the form, including the items pronounced in the neighboring language, were given in the orthography of the native language of the participant, again making sure that the participants did not a priori know in which language the item was pronounced. The interstimulus interval was 5 s, allowing the participants enough time to make their choice. After each block the participants heard a beep, so they could keep track of the experiment. On half of the forms the fortis stops were presented on the left hand side and the lenis stops on the right hand side, on the other half of the forms this was done the other way around. The forms contained 20 items per page.

[8] It is conceivable that differences between the language varieties of the listeners and the language varieties of the speakers affect the listeners' performance in the task. This is especially true for the Danish listeners, because Copenhagen speakers have affricated prevocalic /t/, and Tønder speakers do not. This would, as a consequence, be reflected in a diminished comprehension of the prevocalic stops by the Danish participants. However, we did not see any such difficulties reflected in our results.

Table 3
Responses on the auditory classification task, broken down by the participant's native language and the position of the sound.

Participants	Scoring	Prevocalic	Intervocalic	Total
Danish ($N = 40$)	Fortis (−3)	46.7%	23.5%	35.1%
	Lenis (3)	38.0%	32.2%	35.1%
	Unsure (−2 to 2)	15.3%	44.3%	29.8%
	Mean score	−.26	.37	.13
Swedish ($N = 37$)	Fortis (−3)	42.9%	32.8%	37.9%
	Lenis (3)	29.5%	48.8%	39.2%
	Unsure (−2 to 2)	27.7%	18.4%	23.0%
	Mean score	−.46	.65	.23

4.2. Results

The perception experiment can provide us with information on the accuracy with which Danish and Swedish participants perceive native and non-native sounds, as well as on the perceived category of these sounds. Furthermore, the use of a confidence scale allows us to determine how confident participants are in their choices and perhaps how prototypical the items are as exemplars of the participants' native category. The task of the participants was to judge which of two words they heard by choosing between two orthographic forms. In principle, this choice is based on the perception of the stop, since this is where the two words differed minimally. We will therefore refer to the choice by the participants as the category of the stop, i.e. fortis or lenis.

Table 3 presents an overview of the responses on the auditory classification task. Danish words with intervocalic alveolars were removed from all analyses (as in the production experiment). In total (pooling the classification of native and non-native sounds), participants answered unbiased. Items were given an extreme score and were judged to be clearly fortis or lenis an equal number of times. Roughly a quarter of the items received a score in between the extremes, showing less confidence in the choice. As the total mean scores were close to zero on average the participants' choices were equally balanced.

In prevocalic position, both Danish and Swedish participants perceived more words to contain fortis stops than words to contain lenis stops, whereas in intervocalic position this pattern is mirrored. For words with stops in intervocalic position, Danish participants were unsure in 44.3% of the cases, whereas for words with stops in prevocalic position they were more sure, with only 15.3% of choices between the extremes. Swedish participants, on the other hand, were almost equally confident for words with stops in prevocalic and intervocalic position. The 0-score, indicating complete uncertainty, was chosen by Swedish participants in 3.3% of the cases (5.3% in prevocalic and 2.0% in intervocalic position) and by Danish participants in 4.7% of the cases (1.9% in prevocalic and 6.2% in intervocalic position).

To find out how accurately the participants perceived which word the speaker intended (i.e. the orthographically presented word the speaker read), we looked at the accuracy of scoring. To this end, we transformed the confidence scale into a bipolar scale, leaving out the 0-scores. With a mixed-effects model with restricted maximum likelihood, we compared the accuracy (dependent variable) by four fixed effects: the language of the participant (2 levels: Danish vs. Swedish), the language of the item (2 levels: Danish vs. Swedish), the position of the sound in the word (2 levels: prevocalic vs. intervocalic) and the category of the intended stop (2 levels: fortis vs. lenis). As a first analysis revealed that there were no significant differences between words and non-words, we collapsed the words and non-words in our further analyses. Listener and item are crossed random effects, and, again, we employ a significance level of 5%. Initially, the random effect of speaker was included, but this effect turned out to be redundant, and was therefore removed from the analysis. We report the actual degrees of freedom (rounded to the nearest integer) that were used in the statistical test. All effects and interactions turned out to be significant. The accuracy differs significantly for language of the participant ($F(1,88) = 46.26$, $p = .000$), language of the item ($F(1,29) = 20.5$, $p = .000$), position ($F(1,22) = 87.82$, $p = .000$), and for category ($F(1,22) = 5.58$, $p < .05$). Interactions are significant for language participant*language item ($F(1,6974) = 230.38$, $p = .000$), language participant*position ($F(1,6971) = 158.67$, $p = .000$), language participant*category ($F(1,6967) = 5.19$, $p < .05$), language item*position ($F(1,29) = 63.23$, $p = .000$), language item*voice ($F(1,29) = 53.37$, $p = .000$), position*category ($F(1,22) = 59.89$, $p = .000$), language participant*language item*position ($F(1,6969) = 47.92$, $p = .000$), language participant*language item*category ($F(1,6965) = 24.35$, $p = .000$), language participant*position*category ($F(1,6968) = 62.05$, $p = .000$), language item*position*category ($F(1,29) = 32.02$, $p = .000$) and language participant*language item*position*category ($F(1,6966) = 144.65$, $p = .000$). Below, we will discuss the most important results.

First of all, there is a significant effect of the language of the participant. This suggests an asymmetrical perception, which was the starting point of our study: Danish participants were expected to understand Swedish better than Swedish

Table 4
Mean accuracy scores on the auditory classification task (0 = incorrect, 1 = correct).

Language of participant	Language of item	Position	Category[a]	EM Means
Danish	Danish	Prevocalic	Lenis	0.980
			Fortis	0.965
		Intervocalic	Lenis	0.669
			Fortis	0.403
	Swedish	Prevocalic	Lenis	0.743
			Fortis	0.920
		Intervocalic	Lenis	0.614
			Fortis	0.567
Swedish	Danish	Prevocalic	Lenis	0.748
			Fortis	0.958
		Intervocalic	Lenis	0.888
			Fortis	0.144
	Swedish	Prevocalic	Lenis	0.782
			Fortis	0.933
		Intervocalic	Lenis	0.924
			Fortis	0.997

[a] See footnote a Table 2.

participants understand Danish. However, the mean accuracy on the comprehension task is lower for the Danish participants ($M = 0.73$) than for the Swedish participants ($M = 0.80$), contrasting with our expectations. The lower accuracy of the Danish participants is caused by their 'incorrect' classification of native Danish words with intervocalic stops (which is no surprise, since the items in the two-way choice are homophonous) as well as of Swedish words with intervocalic stops. Swedish participants, on the other hand, do not correctly classify Danish words with intervocalic stops, but they do correctly classify native Swedish words with intervocalic stops. A complete overview of the accuracy scores is given in Table 4.

In section 3.2, we formulated several predictions for the comprehension of Danish and Swedish words with prevocalic and intervocalic stops on the basis of the acoustic properties of the sounds measured in our production experiment. These predictions are valid if listeners base their classification of non-native sounds on the acoustic properties of the sounds in their native language only. Here, we compare these predictions to the accuracy scores of the participants in our perception experiment.

With respect to *prevocalic* stops we formulated three predictions. The first prediction was that classification of words with native prevocalic stops is excellent for both Swedish and Danish. This prediction is largely confirmed by our perception results, although the native comprehension of Swedish words with prevocalic lenis stops is unexpectedly low. We do not have an explanation for the latter result. Secondly, it was predicted that Swedish classification of Danish words with prevocalic fortis stops[9] is moderately good, but that Swedish classification of Danish words with prevocalic lenis stops is at chance level. In our perception experiment, it was found that words with prevocalic lenis and fortis stops are all perceived moderately accurately. As predicted, the perception of words with lenis stops is less accurate than the perception of words with fortis stops. However, in contrast to our prediction, the identification of words with lenis stops is well above chance level. A final prediction about prevocalic stops was that Danish classification of Swedish words with prevocalic stops is poor, as all words with prevocalic stops would be classified as the variant with the lenis stop. Our results contradict this prediction: Danish participants perceive Swedish prevocalic stops very accurately. The words with lenis stops are perceived less accurately than the words with fortis stops; however, this pattern is visible in Swedish native classification as well.

Now let us turn to the predictions that were formulated with respect to *intervocalic* stops. First, it was predicted that Swedish classification of native words with intervocalic stops is excellent. This prediction is confirmed by our results. Second, Danish classification of native words containing intervocalic stops is predicted to be at chance level based on acoustic properties only. Indeed, Danish classification of native words with intervocalic stops is almost at chance level, although the results do show that Danish listeners are more inclined to choose the word form with the orthographic intervocalic <bb> or <gg> when hearing the Danish homophones. Since the Danish homophones contain an intervocalic lenis stop, it is likely that participants choose the words containing <bb> or <gg> because these lenis stops are invariably

[9] Remember that the terms 'fortis' and 'lenis' are used as phonological terms. We do not take this to mean that, phonetically, the Danish prevocalic aspirated or affricated voiceless stop is a prototypical fortis stop.

written as such in prevocalic position. Clearly, the participants do not choose the <bb> or <gg> variant all the time (even though they could, the words being homophones), but this orthographic factor may have had a subtle effect. A second possible explanation may have to do with an effect of lexical properties of the experimental items on the choice the participants make. One such property would be frequency of occurrence.

Returning to our predictions on intervocalic stop perception, a third prediction was that Swedish classification of Danish words containing intervocalic stops is poor, as most Danish stops will be perceived as lenis. This prediction is confirmed by our perception results, where we found that the vast majority of the Danish intervocalic stops was scored as lenis, by Swedish participants. The difference between the behavior of the Danish and Swedish participants with respect to the Danish intervocalic stops is interesting. Danish participants show a (near-) guessing pattern, which is as expected and even correct, since the items are homophones. Swedish participants, however, judge all Danish intervocalic stops to be lenis; an acoustical decision.

A final prediction was that Danish classification of Swedish words containing intervocalic stops is either poor or good, depending on what is a stronger factor: the dissimilarity between the non-native Swedish sound and the native Danish sound, or the phonetic salience of the contrast between the two Swedish sounds. Our perception experiment revealed that Danish participants scored only slightly above chance level when classifying Swedish intervocalic stops. This suggests that participants do not categorize words with non-native sounds based on the strength of the contrast in the non-native language, but possibly based on the characteristics of their native sounds.

Summarizing, some of the predictions we formulated for native and non-native perception of Danish and Swedish stops on the basis of the acoustic properties of the sounds were confirmed in the perception experiment. However, in other cases the predictions were contradicted by the results of the perception experiment. In the following section, we try to shed more light on the process of cross-language sound perception by investigating which phonetic correlates were actually used by the participants in our perception experiment. We do this by performing a mixed models analysis of the perception data of Experiment 2.

4.3. Phonetic correlates in perception

The goal of a mixed models analysis of the perception data, with the phonetic correlates as dependent variables, is to find out which correlates are actually used by listeners in native and non-native stop perception in Danish and Swedish, and how well these correlates predict the listeners' categorization of sounds as fortis or lenis. This differs from our analysis in section 3, where we looked at the acoustic correlates present in the produced sounds. To determine which correlates are used by listeners, we focus on the comprehension of words based on the listeners' judgments. So, whereas in the analysis of the production data in section 3 we based properties of the fortis–lenis contrast on the *intended*[10] words by the speaker, in the analysis in this section we base the properties of the fortis–lenis contrast on the *perceived* words by the listener.

We ran a separate analysis for each phonetic correlate, as we did in the analyses of the production data in section 3. The fixed factors in each analysis are the perceived category (fortis or lenis), position (prevocalic or intervocalic) and language of production (Danish or Swedish). We used an aggregate model of the results of the perception experiment, collapsing the performance of the listeners and items, since these factors proved to be redundant in the first run of the analyses. More importantly, keeping the listeners as a separate factor would increase the chance of an alpha-error, because it artificially increases the weight of the values of the phonetic correlates and greatly decreases variation, since each separate value would be counted 40 or 37 (participants)*2 (perceived categories) times. Collapsing the factors means that the values that enter the analysis represent the mean of the acoustic correlate for the specific combination of factors (e.g. Danish participants categorizing a Swedish intervocalic stop as pronounced by a particular speaker as lenis).

Because there are eight different correlates, two languages and three fixed factors, we will summarize the results for convenience. Significant differences between levels of the variables 'position' and 'language' are not necessarily interesting, since the values are bound to the range within which the sound is produced. The significant interactions with the perceived category, however, give an insight into the choices the listeners make and at which point they differentiate between fortis on the one hand and lenis on the other hand in the different languages and positions. The results will be presented separately for the Danish and the Swedish listeners, and for each group of listeners a distinction will be made between the perceived categories in prevocalic and intervocalic position.

[10] With the intended words, we mean the words the speaker read in the production experiment. For the Danish intervocalic stops this means that two words (e.g. <lækker> and <lægger>) are treated as separate words, although acoustically they are homophones. The perceived words are categorized in the same way, since the participants were asked to choose between two orthographically presented categories (e.g. <lækker> and <lægger>).

4.3.1. Danish listeners

For the Danish listeners, we find significant interaction effects on duration of release (correlates 1 and 5) for position*language ($F(1,12) = 17.14$, $p = .001$), position*category ($F(1,12) = 5.34$, $p < .05$), language*category ($F(1,12) = 10.89$, $p < .01$) and position*language*category ($F(1,12) = 11.53$, $p = .005$). For intensity of release (correlates 2 and 6) there are significant effects for position ($F(1,12) = 16.84$, $p = .001$) and an interaction for position*language ($F(1,12) = 7.29$, $p < .05$). For rise of the following vowel (correlates 3 and 7) there is a significant effect of position ($F(1,12) = 16.50$, $p < .005$). For VOT (correlate 4), which is measured in prevocalic position only, we find a significant effect of language ($F(1,4) = 13.41$, $p < .05$) and category ($F(1,4) = 180.6$, $p = .000$). For decay of the preceding vowel (correlate 8), which is measured in intervocalic position only, we find a significant effect for language ($F(1,4) = 7.78$, $p < .05$). Duration of the preceding vowel (correlate 9, intervocalic only) is significantly different by language ($F(1,4) = 21.36$, $p = .01$). Occlusion duration (correlate 10, intervocalic only) is significantly different by language ($F(1,4) = 399.00$, $p = .000$) and category ($F(1,4) = 161.45$, $p = .000$), with an interaction of language and category ($F(1,4) = 198.47$, $p = .000$). Intensity of the voice bar (correlate 11, intervocalic only) is significantly different for language ($F(1,4) = 20.03$, $p < .05$), category ($F(1,4) = 37.50$, $p < .005$) and language*category ($F(1,4) = 26.91$, $p < .01$).

In *prevocalic* position, we see an important role for VOT. The interaction effect of language of production and perceived category is significant, where VOT is higher for Danish sounds perceived as fortis ($M = 77$, SE = 6) than for Swedish sounds perceived as fortis ($M = 43$, SE = 6), and is higher for Danish sounds perceived as lenis ($M = 36$, SE = 6) than for Swedish sounds perceived as lenis ($M = 15$, SE = 6). The mean VOT for the perceived Swedish fortis category is very close to the Danish lenis category (see Table 5, top half, for the Danish listeners' means of the phonetic correlates in prevocalic position). When we compare these results to the production data, we see that the mean values of VOT in the categories perceived by the Danish listeners are very close to the mean values of the categories produced by the Danish speakers. In prevocalic position, this seems to be the case for the other correlates as well, which is not surprising, given the high accuracy scores of the Danish participants.

An interesting effect with duration of release is that the pattern for non-native sounds is the mirror image (with lenis ($M = 25$, SE = 4) being longer than fortis ($M = 18$, SE = 4)) of the pattern for native sounds (with lenis (M = 22, SE = 4) being shorter than fortis (M = 53, SE = 4)). It is unlikely that an important factor in native categorization would yield the opposite result when applied to non-native sounds, which leads us to believe that duration of release is not used by Danish listeners to categorize Swedish prevocalic stops. Rather, it is likely that Danish listeners base their categorization on VOT, for which the patterns are the same in the two languages. However, the problem remains that when the categorization is near-perfect, the values for all acoustic correlates are native-like (since they are not independent) and it is impossible to see which correlate listeners base their categorization on. To address this issue one would have to design a task with resynthesized speech.

In *intervocalic* position, Danish listeners seem to level out any contrast between fortis and lenis that was available in the production data of Swedish speakers (see Table 6, top half, for the Danish listeners' means of the phonetic correlates in intervocalic position).

Table 5
Mean values of the phonetic correlates and standard errors (SE) for the perceived and produced categories in prevocalic position in Danish and Swedish.

Phonetic correlates: EM means (SE)			Danish stimuli		Swedish stimuli	
			Perceived	Produced	Perceived	Produced
Danish listeners	Duration of release	Fortis[a]	53 (4)	53 (4)	18 (4)	17 (4)
		Lenis	22 (4)	22 (4)	25 (4)	28 (4)
	Intensity of release	Fortis	−17.6 (2.6)	−17.5 (3.7)	−17.3 (2.6)	−16.9 (3.7)
		Lenis	−19.8 (2.6)	−19.9 (3.7)	−12.2 (2.6)	−12.7 (3.7)
	Rise of following vowel	Fortis	36 (6)	35 (5)	35 (6)	36 (5)
		Lenis	44 (6)	44 (5)	27 (6)	28 (5)
	VOT	Fortis	77 (6)	76 (9)	43 (6)	49 (9)
		Lenis	36 (6)	37 (9)	15 (6)	15 (9)
Swedish listeners	Duration of release	Fortis	47 (4)	53 (4)	17 (4)	17 (4)
		Lenis	22 (4)	22 (4)	26 (4)	28 (4)
	Intensity of release	Fortis	−18.1 (2.9)	−17.5 (3.7)	−16.7 (2.9)	−16.9 (3.7)
		Lenis	−20.0 (2.9)	−19.9 (3.7)	−13.8 (2.9)	−12.7 (3.7)
	Rise of following vowel	Fortis	36 (6)	35 (5)	34 (6)	36 (5)
		Lenis	43 (6)	44 (5)	29 (6)	28 (5)
	VOT	Fortis	69 (6)	76 (9)	43 (6)	49 (9)
		Lenis	37 (6)	37 (9)	15 (6)	15 (9)

Durations are in milliseconds, intensities in decibels (see footnote b Table 2).

[a] See footnote a Table 2.

Table 6
Mean values of the phonetic correlates and standard errors (SE) for the perceived and produced categories in intervocalic position in Danish and Swedish.[a]

Phonetic correlates: EM means (SE)			Danish stimuli		Swedish stimuli	
			Perceived	Produced	Perceived	Produced
Danish listeners	Duration of release	Fortis[b]	19 (4)	20 (3)	27 (4)	25 (2)
		Lenis	21 (4)	21 (3)	28 (4)	27 (2)
	Intensity of release	Fortis	−9.6 (2.6)	−10.6 (3.1)	−12.3 (2.6)	−10.9 (2.9)
		Lenis	−9.6 (2.6)	−8.6 (3.1)	−13.5 (2.6)	−12.3 (2.9)
	Rise of following vowel	Fortis	18 (6)	17 (4)	19 (6)	16 (3)
		Lenis	19 (6)	17 (4)	19 (6)	16 (3)
	Decay of preceding vowel	Fortis	51 (4)	51 (5)	70 (4)	73 (4)
		Lenis	58 (4)	54 (5)	70 (4)	58 (4)
	Duration of preceding vowel	Fortis	138 (7)	133 (8)	93 (7)	80 (7)
		Lenis	142 (7)	128 (8)	94 (7)	89 (7)
	Occlusion duration	Fortis	50 (4)	45 (7)	171 (4)	192 (6)
		Lenis	50 (4)	44 (7)	162 (4)	134 (6)
	Intensity of voice bar during occlusion	Fortis	−12.4 (1.7)	−12.6 (1.9)	−24.9 (1.7)	−32.3 (1.9)
		Lenis	−12.1 (1.7)	−12.0 (1.9)	−21.2 (1.7)	−11.8 (1.9)
Swedish listeners	Duration of release	Fortis	21 (4)	20 (3)	27 (4)	25 (2)
		Lenis	20 (4)	21 (3)	29 (4)	27 (2)
	Intensity of release	Fortis	−10.5 (2.9)	−10.6 (3.1)	−11.4 (2.9)	−10.9 (2.9)
		Lenis	−9.4 (2.9)	−8.6 (3.1)	−14.7 (2.9)	−12.3 (2.9)
	Rise of following vowel	Fortis	20 (6)	17 (4)	19 (6)	16 (3)
		Lenis	19 (6)	17 (4)	19 (6)	16 (3)
	Decay of preceding vowel	Fortis	60 (4)	51 (5)	76 (4)	73 (4)
		Lenis	55 (4)	54 (5)	61 (4)	58 (4)
	Duration of preceding vowel	Fortis	139 (7)	133 (8)	90 (7)	80 (7)
		Lenis	140 (7)	128 (8)	98 (7)	89 (7)
	Occlusion duration	Fortis	50 (5)	45 (7)	190 (5)	192 (6)
		Lenis	50 (5)	44 (7)	138 (5)	134 (6)
	Intensity of voice bar during occlusion	Fortis	−11.8 (1.8)	−12.6 (1.9)	−31.6 (1.8)	−32.3 (1.9)
		Lenis	−12.2 (1.8)	−12.0 (1.9)	−12.9 (1.8)	−11.8 (1.9)

Durations are in milliseconds, intensities in decibels (see footnote b Table 2).

[a] The cases in which participants did not make a choice for either word (i.e. chose ‘0’) are excluded from this analysis.

[b] See footnote a Table 2.

There is a significant interaction for language of production and perceived category for two correlates in intervocalic position, namely occlusion duration and intensity of the voice bar. In the Danish listeners’ categorization of Danish sounds, there are no acoustic differences for intervocalic stops on either correlate. This is as expected, since there is no difference. In the Danish listeners’ categorization of Swedish sounds, on the other hand, there is a contrast: lenis (M = 162, SE = 4) are shorter than fortis (M = 171, SE = 4). The same difference between the Danish listeners’ perception of Danish and Swedish sounds can be seen in the intensity of the voice bar. Although in Danish there is not contrast, the Danish listeners’ category of Swedish lenis has a significantly higher intensity (M = −21.2, SE = 1.7) than the Danish listeners’ category of Swedish fortis (M = −24.9, SE = 1.7). This means that even though in Danish the intervocalic stops have merged, the Danish participants’ categories of Swedish lenis and fortis are significantly different, especially in duration and intensity. However, if we compare the differences in occlusion duration and intensity of the voice bar of these perceived Swedish categories to those of the produced Swedish categories, the differences are very small. In other words: the contrasts produced by the Swedish speaker are largely leveled out in the categorizations made by the Danish listener.

4.3.2. Swedish listeners

For Swedish listeners, we found significant effects for duration of release (correlates 1 and 5) for the interactions of position*language ($F(1,16) = 11.26$, $p < .005$), language*category ($F(1,16) = 9.09$, $p < .01$) and position*language*category ($F(1,16) = 6.02$, $p < .05$). For intensity of release (correlates 2 and 6), there is a significant effect of position ($F(1,12) = 12.71$, $p < .005$) and a significant interaction for position*language ($F(1,12) = 5.67$, $p < .05$). For rise of the following vowel (correlates 3 and 7), there is a significant effect of position ($F(1,12) = 24.07$, $p = .000$). For VOT (correlate 4), measured only in prevocalic position, there was an effect of language ($F(1,4) = 10.37$, $p < .05$) and category ($F(1,4) = 62.84$, $p = .001$). For decay of the preceding vowel (correlate 8), which is measured in intervocalic position only, we find significant effects for category ($F(1,4) = 13.719$, $p < .05$). Duration of the preceding vowel (correlate 9, intervocalic only) is significantly different by language ($F(1,4) = 27.25$, $p < .01$). Occlusion duration (correlate 10, intervocalic only) is significantly different by language ($F(1,4) = 379.07$, $p = .000$) and category ($F(1,4) = 31.03$, $p = .005$), with an interaction of language and category ($F(1,4) = 30.57$, $p = .005$). Intensity of the voice bar (correlate 11, intervocalic only) is significantly different for language ($F(1,4) = 16.50$, $p < .05$), category ($F(1,4) = 192.71$, $p = .000$) and language*category ($F(1,4) = 426.55$, $p = .000$).

In *prevocalic* position, the Swedish listeners, like the Danish listeners, seem to strongly rely on VOT (see Table 5, bottom half, for the Swedish listeners' means of the phonetic correlates in prevocalic position). There are many significant effects and significant interactions of the eleven phonetic correlates in the perceived categories, but only with VOT the difference between fortis (Swedish: $M = 43$, SE = 6, Danish: $M = 69$, SE = 6) and lenis (Swedish: $M = 15$, SE = 6, Danish: $M = 37$, SE = 6) has the same direction in the two languages, with VOT being higher in fortis than in lenis sounds. For the Swedish listeners, like for the Danish listeners discussed earlier, the means in VOT of the perceived Swedish fortis category are close to the means of the perceived Danish lenis category. In fact, the difference between the native and non-native perceived categories is significant. In contrast to VOT, for duration of release and intensity of release we find that the patterns in categorizing non-native sounds are the mirror image of the patterns in categorizing native sounds. This resembles the patterns we found for these correlates with Danish listeners. For Swedish listeners, the duration of release of perceived Danish fortis is longer than that of Danish lenis and the intensity of release is lower, but the duration of release of perceived Swedish fortis is shorter than that of Swedish lenis and the intensity of release is higher. Again, it is conceivable that these effects merely surface because of the success of the classification based on VOT, suggesting that these correlates did not play any role in the categorization of the non-native sounds by Swedish listeners.

In their categorization of native Swedish words with stops in *intervocalic* position, the Swedish listeners seem to rely on three correlates: decay of the preceding vowel, occlusion duration and intensity of the voice bar during occlusion, as for these correlates the differences between fortis and lenis categories are significant (see Table 6, bottom half, for the means of the phonetic correlates in intervocalic position). These three correlates are the same ones that contribute to the fortis–lenis contrast in production. This is not surprising, given the high accuracy of Swedish listeners in categorizing native Swedish words with intervocalic stops. For categorizing non-native Danish words, on the other hand, the Swedish listeners do not rely on any of the correlates; differences are small and not significant and all significant interaction effects are caused by significant differences in the Swedish contrast.

4.4. Discussion

In Experiment 2, we investigated listeners' comprehension of native and non-native words containing prevocalic and intervocalic stops in an auditory classification task. The results of this categorization task are discussed here, distinguishing between prevocalic and intervocalic stops.

Starting with *prevocalic* stops, in our analysis of the listeners' results on this task in section 4.2 we found that both Danish and Swedish listeners accurately perceived native and non-native words containing prevocalic stops. An analysis of the effects of the phonetic correlates in the listeners' categorization in section 4.3 showed that both Danish and Swedish listeners seem to categorize words with prevocalic stops mainly on the basis of differences in VOT. The success of non-native categorization based on VOT is unexpected, as the mean VOT of Swedish fortis stops is very close to the mean VOT of Danish lenis stops, and the differences between perceived native and non-native categories are significant. Presumably, Danish and Swedish listeners employ different category boundaries for VOT in non-native sounds than for native sounds. Since for both language groups the accuracy of perception is very high, the values of the correlates in the perceived category are very close to the values of the correlates in the produced category. Hence, the perceived and produced categories coincide almost completely.

The most surprising results surfaced in the comprehension of words containing *intervocalic* stops. Starting with the Danish listeners, we found that they performed poorly in categorizing the intervocalic stops both in their native language and in the non-native language. As expected, Swedish listeners accurately comprehended the words containing intervocalic stops in their native language, which causes the same contrast that Swedish speakers use for distinguishing between fortis and lenis sounds in production to surface in the perception analysis. When Swedish listeners had to comprehend Danish words containing intervocalic stops, however, they did not use any of the correlates they use in the categorization of sounds from their native languages.

For Danish and Swedish listeners, their patterns of comprehension of words containing intervocalic stops are remarkably different. Whereas Swedish listeners categorize the Danish words as containing lenis stops, Danish listeners seem to categorize both Danish and Swedish stops at random. It seems that Danish listeners have a slight preference for the lenis alternative, which may be due to orthographical and lexical reasons, such as frequency of the target word. Danish listeners have to base their categorization of these words containing intervocalic stops on contextual information such as word meaning, syntactic structure and discourse context. In the Danish listeners' comprehension of Swedish words containing fortis and lenis stops, accuracy is slightly above chance level. As our analysis of the perceived categories in section 4.3 shows, Danish listeners may have used the intensity of the voice bar and the occlusion duration to distinguish fortis and lenis stops in Swedish. Since these two correlates are important in the production of the Swedish intervocalic fortis–lenis contrast, this might have caused the slightly increased performance of Danish listeners on their perception of Swedish intervocalic stops as compared to the Swedish performance on Danish intervocalic stops.

5. General discussion

In this study, we compared native production and native and non-native perception of Danish and Swedish sounds to investigate the effect of cross-language sound perception on intelligibility. The question we set out to answer is whether Danish diachronic intervocalic lenition is a factor in the asymmetry in intelligibility found between Danes and Swedes in previous studies. In Danish, in contrast to Swedish, intervocalic stops are not contrastive due to the diachronic process of consonantal lenition. To answer our research question, we compared the production and perception of Danish and Swedish prevocalic and intervocalic stops.

An analysis of the intervocalic stops produced by native Danish speakers in our *production* experiment confirmed the expected lack of contrast in Danish intervocalic, originally long, stops in modern Danish. In modern Swedish, on the other hand, intervocalic fortis and lenis stops are contrastive and this contrast is phonetically very salient. The fortis and lenis stops produced by the native Swedish speakers in our production experiment could be clearly distinguished on the basis of three phonetic correlates, namely decay of the preceding vowel, occlusion duration and intensity of the voice bar during occlusion.

In our production experiment, we did not only look at intervocalic stops but also at prevocalic stops, as it is conceivable that Danish listeners may still be sensitive to the fortis–lenis contrast in intervocalic stops in a non-native language due to the presence of this contrast elsewhere in their own language. Their knowledge about the contrast in prevocalic stops in Danish may help Danish listeners in their perception of the contrast in intervocalic stops in Swedish. Analyzing the phonetic characteristics of the prevocalic fortis and lenis stop in Danish, however, we found that they were very different from the characteristics of the intervocalic stop in Danish. Also, the characteristics of the Danish stops were very different from the Swedish stops. For example, the VOT values of the Danish prevocalic lenis stop were different from the VOT values of the Swedish prevocalic lenis stop, but rather overlap with the VOT values of the Swedish prevocalic fortis stop. Because of these differences in the phonetic characteristics in the two positions and the two languages, we believe it is implausible that Danish listeners are aided in their perception of the fortis–lenis contrast in Swedish intervocalic stops by the fortis–lenis contrast in Danish prevocalic stops.

The *perception* experiment showed that the perception and categorization of native sounds was far from excellent in the absence of any disambiguating contextual information. In the production experiment, we saw that there is no contrast in Danish intervocalic stops. This was as expected and was reflected in the native perception of these intervocalic stops by Danish listeners: they showed a guessing pattern for words with these sounds. Unexpectedly, Swedish listeners had problems accurately categorizing words with prevocalic lenis stops in their native language, and partly categorized them as fortis.

To answer our research question, we considered the non-native perception of *intervocalic* stops. We assumed that the comprehension of Swedish words containing intervocalic stops by Danish listeners would be better than the comprehension of Danish words containing intervocalic stops by Swedish listeners. However, our results show that Swedish as well as Danish listeners performed poorly in their categorization of non-native stops. That is, we did not find the expected asymmetry in Danish and Swedish cross-linguistic perception, which leads us to the conclusion that intervocalic stop lenition in Danish is not likely to be part of the explanation of the asymmetry in intelligibility between Danish and Swedish.

The Swedish listeners performed poorly on their perception of Danish stops and comprehended a vast majority of the Danish words as referring to words containing lenis rather than fortis stops. This was expected on the basis of the phonetic characteristics of their native intervocalic stops. As the intensity of the voice bar of Danish intervocalic stops is comparable to the intensity of the voice bar of Swedish lenis stops, this is likely to have caused Swedish participants to map the Danish words to their Swedish counterpart containing the lenis stop. Correlates such as occlusion duration and intensity of the voice bar during occlusion do not allow for the use of Swedish categories for the perception of Danish, because the values are too high or too low, respectively. Decay of the preceding vowel, however, has a comparable range in Danish and Swedish stops, as confirmed by the production data. This allows the Swedish listeners to use this correlate in their categorization of the Danish stops. However, although the correlate is important in the native Swedish contrast, Swedish listeners did not use this correlate in their categorization. We did not find a difference in decay between the two categories. The lack of use of decay may be a matter of acoustic salience; intensity of the voice bar is arguably a more salient acoustic correlate than decay of the preceding vowel.

The Danish listeners were barely above chance in their comprehension of Swedish words containing intervocalic stops, even though Swedish intervocalic stops show a salient fortis–lenis contrast and Swedish listeners have excellent categorization of words containing intervocalic stops in their native language. The low accuracy by the Danish listeners was against our expectations and reveals that not only Swedish listeners but also Danish listeners have problems with the categorization of sounds in the neighboring language. As a consequence, there is no evidence for an asymmetry in intelligibility of the items in our study. One of the possible explanations is that the Danish listeners are not able to perceive the contrast because the sounds are phonetically distinct from the Danish intervocalic stops. The phonetic analysis of the perception experiment showed that the fortis and lenis stops as perceived by the Danish listeners differed significantly in occlusion duration and intensity of the voice bar. This suggests that the Danish listeners picked up on the important correlates of the Swedish contrast, although the difference between the fortis and lenis categories of Swedish stops as perceived by Danish listeners is very small compared to the difference between these categories when they were produced by native Swedish speakers. The Danish listeners could perhaps have used intensity of the voice bar as a correlate of voicing in their perception of Swedish words containing intervocalic stops because of the presence of voicing during articulation of prevocalic stops in their native language. This presence of voicing (in our study only measured as a temporal correlate; VOT) is a correlate in native perception of the Swedish as well as the Danish prevocalic fortis–lenis contrast. Therefore, Danish participants may have used knowledge of their native prevocalic contrast in their perception of Swedish intervocalic stops, even though it was not successful.

Another explanation for the low accuracy in the comprehension of Swedish words containing intervocalic stops by Danish listeners is that these listeners do not pay attention to a contrast that has no role in that position in their own language. Their guessing pattern in the comprehension of Swedish words resembles the pattern of the comprehension of Danish words: at random and at chance. This is different from the behavior of the Swedish listeners concerning the Danish words containing intervocalic stops: most are classified as containing lenis stops (reflected in their selection of words containing <bb> or <gg>). The significant difference in duration and intensity between the two non-native categories arising from the classification by Danish listeners would be an argument against this hypothesis, but it cannot unequivocally be refuted on the basis of this evidence alone, since the difference is small and the accuracy, as said, is low.

In our study, we also looked at the non-native comprehension of words containing *prevocalic* stops. Listeners are less confident and less accurate in their comprehension of non-native words containing stops in this position than in their comprehension of native words with stops in the same position. However, their performance was higher than expected. Danish listeners showed the same performance in the classification of Swedish prevocalic stops as Swedish listeners. In other words, Danish listeners showed near-native categorization of Swedish prevocalic stops. With respect to the reverse situation of Swedish listeners' perception of Danish stops, Swedish listeners were found to perceive part of the Danish prevocalic lenis stops as fortis. This was as expected and can be explained by the higher VOT of Danish prevocalic stops than of Swedish prevocalic stops. However, the mean VOT values of the Danish stops as perceived by Swedish listeners are more similar to the VOT values of the stops produced by Danish speakers than that of the stops produced by Swedish speakers. These results suggest that the Danish and Swedish listeners used different boundaries for VOT in their categorization of native stops than in their categorization of non-native stops. If the listeners would have employed native category boundaries, their accuracy on the non-native language would have been much lower.

These adaptations of the boundaries for categorization of sounds in the non-native language seem to be difficult to explain by the assumptions of exemplar dynamics (Pierrehumbert, 2001, 2003) and also Best (1995, the latter based on articulatory properties), in which non-native sounds are assumed to be categorized on the basis of the similarity of this sound to the native category. The perception of intervocalic and prevocalic stops shows that, in this study, this is not the case. Rather, the categories of non-native sounds seem to be defined by different boundaries than the categories of native sounds. In contrast to bilingual language users, who show perceptual code-switching when their proficiency in the second language is above a certain level (e.g. Flege et al., 1997; Flege and Eefting, 1987; Hazan and Boulakia, 1993), the participants in our study claimed and showed to have no proficiency in the neighboring language. Nevertheless, they adjusted their categorization to the characteristics of the non-native language they perceived. It is conceivable that our participants may have come into contact with the neighboring language through television or otherwise. This contact, then, was not enough to make the participants perform well enough on an auditory translation task, but it could have resulted in stored phonetic exemplars (in exemplar dynamics' terms) and a non-native category, which could have helped categorization in the non-native language. This explanation, however, would require that new phonetic categories are established relatively easily on the basis of superficial contact with the language.

Summarizing, the diachronic lenition of intervocalic stops in Danish is not likely to be part of the explanation for the asymmetry in intelligibility between Danish and Swedish. This can be concluded from our perception results, which show that the intelligibility relation between Danish and Swedish in this study (limited to the perception of intervocalic stops) is not asymmetrical. However, this does not mean that there is no role for lenition in mutual intelligibility. The loss of the intervocalic fortis–lenis contrast in Danish clearly impairs perception of Danish by Swedish listeners. At the same time, the intervocalic long stop in Swedish seems to be an unfamiliar sound for Danish listeners, making it hard for them to categorize this sound, despite their knowledge of the native prevocalic fortis–lenis contrast. The fact that the experimental data revealed no asymmetry can in fact be attributed to these two separate factors. On the one hand, lenition in Danish is an impeding factor for Swedish listeners, and on the other hand, the differently positioned phonetic contrast leaves Danish listeners at a loss. A potentially interesting follow-up study might be an investigation of stops following long vowels in Danish and Swedish, because in that case, all the Swedish intervocalic stops would be short and post-aspirated, like the Danish stops. Additionally, it might be interesting to study the effect on perception of the diachronic lenition of Danish intervocalic alveolar stops to approximants. This lenition has led to many long vocalic sequences, which are likely to be hard to segment by Swedish listeners. This is another issue which could be an interesting topic for future research.

Results of previous studies on mutual intelligibility show that phonetic differences between languages play an important role in intelligibility (see, e.g. Gooskens, 2007). This study confirms that non-native categorization is guided by native categorization, but at the same time suggests that speakers can in some cases adjust their native category boundaries in order to fit non-native input.

6. Conclusions

This study investigated the role of intervocalic lenition in Danish in the mutual intelligibility of Danish and Swedish. Previous studies have found that Danish listeners understand Swedish speakers better than Swedish listeners understand Danish speakers. We tested the hypothesis that this asymmetry in intelligibility between Danish and Swedish may be explained by differences between the phonetic characteristics of the two languages. In particular, we hypothesized that the asymmetry may be partially explained by the diachronic process of consonantal lenition, which has neutralized any stop contrast for intervocalic stops in Danish but not in Swedish. We tested this hypothesis in two experiments: a production experiment with native speakers of the two languages, and a perception experiment investigating native and non-native perception of Danish and Swedish intervocalic stops.

Our perception experiment showed that in a controlled setting focusing on the acoustic properties of the sounds, there is no asymmetry between Swedish and Danish listeners in their cross-language perception of intervocalic stops: neither Swedish, nor Danish listeners were able to correctly categorize words containing intervocalic stops in the neighboring language. For the perception of Danish stops by Swedish listeners, this is not surprising, as our production experiment confirmed that Danish intervocalic stops are non-contrastive. However, this lack of contrast does not result in an asymmetry in intelligibility, as the perception of the acoustically salient contrast of Swedish intervocalic stops by Danish listeners was equally poor. Relating the perception results to the production results, we found that in prevocalic as well as intervocalic position, Danish and Swedish listeners adjust their native categories for non-native perception, with varying success.

Appendix. The 24 minimal pairs in Danish and Swedish used in the production task

Contrast	Voice	Danish	Swedish	English translation	
Prevocalic (cognate)	Fortis	pakker [p^{h}ɑgɐ]	packar [p^{h}akːar]	take	
		taler [t^{s}æːˀlɐ]	talar [t^{h}ɑːlar]	talk	
		kalde [k^{h}alə]	kalla [k^{h}alːa]	call	
	Lenis	bakker [b̥ɑɡ̊ɐ]	backar [bakːar]	back	
		daler [d̥æːˀlɐ]	dalar [dɑːlar]	go down	
		galle [ɡ̊alə]	galla [galːa]	gall	
Intervocalic (cognate)	Fortis	lapper [lɑb̥ɐ]	lappar [lapːar]	rags	
		retter [ʁad̥ɐ][a]	rättar [rɛtːar]	straighten	
		lækker [lɛɡ̊ɐ]	läcker [lɛkːər]	delicious	
	Lenis	labber [lɑb̥ɐ]	labbar [labːar]	paws	
		redder [ʁɛðɐ]	räddar [rɛdːar]	save	
		lægger [lɛɡ̊ɐ]	lägger [lɛgːər]	lay	
Intervocalic (non-cognate)	Fortis	ripper [ʁib̥ɐ]	klippa [klipːa]	rip	cliff
		rytter [ʁyd̥ɐ]	smitta [smitːa]	rider	infection
		tikker [t^{s}eɡ̊ɐ]	snacka [snakːa]	ticking	talk
	Lenis	ribber [ʁib̥ɐ]	klibba [klibːa]	ribs	cling
		rydder [ʁyðɐ]	smidda [smidːa]	clear	forged
		tigger [t^{s}egɐ]	snagga [snagːa]	beggar	crop
Intervocalic (non-word)	Fortis	appe [ɑb̥ə]	appe [apːə]	–	
		atte [ad̥ə]	atte [atːə]	–	
		akke [ɑɡ̊ə]	acke [akːə]	–	
	Lenis	abbe [ɑb̥ə]	abbe [abːə]	–	
		adde [aðə]	adde [adːə]	–	
		agge [ɑɡ̊ʌ]	agge [agːə]	–	

[a] Danish words and non-words with the intervocalic alveolar have not been used in the analyses.

References

Baayen, H., 2008. Analyzing Linguistic Data, a Practical Introduction to Statistics Using R. Cambridge University Press.

Basbøll, H., 2005. The Phonology of Danish. Oxford University Press.

Best, C., 1995. A direct realist view of cross-language speech perception. In: Strange, W. (Ed.), Speech Perception and Linguistic Experience: Theoretical and Methodological Issues. York Press, Timonium, MD, pp. 171–206.

Bleses, D., Basbøll, H., Vach, W., 2011. Is Danish difficult to acquire? Evidence from Nordic past-tense studies. Language and Cognitive Processes 26 (8), 1193–1231.

Bø, I., 1978. Ungdom og naboland [Youth and Neighboring Country]. Rogalandsforskning (rapport 4), Stavanger.

Boersma, P., 1998. Functional Phonology. LOT: Dissertation Series 11. LOT, Utrecht.

Cheng, C.-C., 1997. Measuring relationship among dialects: DOC and related resources. Computational Linguistics & Chinese Language Processing 2, 41–72.

Cho, T., Ladefoged, P., 1999. Variations and universals in VOT: evidence from 18 endangered languages. Journal of Phonetics 27, 207–222.

Debrock, M., 1977. An acoustic correlate of the force of articulation. Journal of Phonetics 5, 61–80.

Delsing, L.O., Lundin Åkesson, K., 2005. Håller språket ihop Norden? En forskningsrapport om ungdomars förståelse av danska, svenska och norska [Does Language Keep the Nordic Countries Together? A Research Report on How Well Young People Understand Danish, Swedish and Norwegian]. Nordiska ministerrådet, Köpenhamn.

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

Elbro, C., 2006. Literacy acquisition in Danish: a deep orthography in cross-linguistic light. In: Malatesha Joshi, M., Aaron, P.G. (Eds.), Handbook of Orthography and Literacy. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 31–45.

Fant, G., 1973. Stops in CV-syllables. In: Fant, G. (Ed.), Speech Sounds and Features. MIT Press, Cambridge, MA, pp. 110–139.

Fischer-Jørgensen, E., 1954. Acoustic analysis of stop consonants. Miscellanea Phonetica 2, 42–59.

Flege, E.F., Eefting, W., 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. Speech Communication 6, 185–202.

Flege, E.F., Bohn, O., Jang, S., 1997. Effects of experience on non-native speakers on production and perception of English vowels. Journal of Phonetics 25, 437–470.

Flemming, E., 1995. Auditory Features in Phonology. UCLA, (PhD dissertation).

Gooskens, C., 2006. Linguistic and extra-linguistic predictors of inter-Scandinavian communication. In: van de Weijer, J., Los, B. (Eds.), Linguistics in the Netherlands, vol. 23. John Benjamins, Amsterdam, pp. 101–113.

Gooskens, C., 2007. The contribution of linguistic factors to the intelligibility of closely related languages. Journal of Multilingual and Multicultural Development 28, 445–467.

Gooskens, C., Beijering, K., Heeringa, W., 2008. Phonetic and lexical predictors of intelligibility. International Journal of Humanities and Arts Computing, 2 (1–2), 63–81.

Gooskens, C., van Heuven, V.J., van Bezooijen, R., Pacilly, J.J.A., 2010. Is spoken Danish less intelligible than Swedish? Speech Communication 52, 1020–1037.

Hazan, V.L., Boulakia, G., 1993. Perception and production of a voicing contrast by French–English bilinguals. Language and Speech 36, 17–38.

Heeringa, W., 2004. Measuring Dialect Pronunciation Differences using Levenshtein Distances. Groningen Dissertations in Linguistics (Grodil), Groningen.

Helgason, P., Ringen, C., 2008. Voicing and aspiration in Swedish stops. Journal of Phonetics 36, 607–628.

Hilton, N.H., Schüppert, A., Gooskens, C., 2011. Syllable reduction and articulation rates in Danish, Norwegian and Swedish. Nordic Journal of Linguistics 34, 215–237.

Jakobson, R., Fant, G.G.M., Halle, M., 1961. Preliminaries to Speech Analysis: The Distinctive Features and Their Correlates. MIT Press, Cambridge, MA.

Jansen, W., 2004. Laryngeal Contrast and Phonetic Voicing: A Laboratory Phonology Approach to English, Hungarian and Dutch. University of Groningen, (PhD dissertation).

Kirchner, R., 1998. An Effort-Based Approach to Consonant Lenition. UCLA, and Rutgers Optimality Archive, (PhD dissertation).

Kortlandt, F., 2003. Glottalization, Preaspiration and Gemination in English and Scandinavian. Amsterdamer Beiträge zur älteren Germanistik, vol. 58. Rodopi, Amsterdam/New York.

Ladefoged, P., Maddieson, I., 1996. The Sounds of the World's Languages. Blackwell, Cambridge, MA.

Levenshtein, V., 1965. Binary codes capable of correcting deletions, insertions and reversals. Doklady Akademii Nauk SSSR 163, 845–848.

Maddieson, I., Gandour, J., 1977. Vowel length before aspirated consonants. Indian Linguistics 38, 6–11.

Malécot, A., 1966. Mechanical pressure as an index of ''force of articulation''. Phonetica 14, 169–180.

Maurud, Ø., 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]. Nordiska rådet, Stockholm.

Moberg, J., Gooskens, C., Nerbonne, J., Vaillette, N., 2006. Conditional entropy measures intelligibility among related languages. In: Van Eynde, F., Dirix, P., Schuurman, I., Vandeghinste, V. (Eds.), Proceedings of Computational Linguistics in the Netherlands. Rodopi, Amsterdam, pp. 51–66.

Nerbonne, J., Heeringa, W., 2010. Measuring dialect differences. In: Schmidt, J.E., Auer, P. (Eds.), Language and Space: Theories and Methods. Mouton de Gruyter, Berlin, pp. 550–567.

Pierrehumbert, J., 2001. Exemplar dynamics: word frequency, lenition, and contrast. In: Bybee, J., Hopper, P. (Eds.), Frequency and the Emergence of Linguistic Structure. John Benjamins, Amsterdam, pp. 137–157.

Pierrehumbert, J., 2003. Phonetic diversity, statistical learning, and acquisition of phonology. Language and Speech 46, 115–154.

Quené, H., van den Bergh, H., 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. Journal of Memory and Language 59, 413–425.

Rischel, J., 1970. Consonant gradation: a problem in Danish phonology and morphology. In: Benediktsson, H. (Ed.), Proceedings of the International Conference of Nordic and General Linguistics. Vísindafélag Íslendinga, Reykjavík. (reprinted in JR Sound Structure in language, Oxford University Press (2009)), pp. 460–480..

Schüppert, A., Gooskens, C., 2011. Investigating the role of language attitudes for perception abilities using reaction time. Dialectologia: Revista Electrònica, 119–140.

Schüppert, A., Ziegler, J., Magnusson, K., Juul, H., Holmqvist, K., Gooskens, C. (in preparation). On-line activation of L1 orthography enhances spoken word recognition of a closely related L2. Evidence from ERP.

Slis, I.H., Cohen, A., 1969. On the complex regulating the voiced-voiceless distinction II. Language and Speech 12.3, 137–155.

Smalley, W., 1961–1964. Manual of Articulatory Phonetics. Practical anthropology, New York.

Steriade, D., 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. Ms, UCLA.

Tang, C., van Heuven, V.J., 2009. Mutual intelligibility of Chinese dialects experimentally tested. Lingua 119, 709–732.