The voicing contrast in Fenno-Swedish stops

Catherine Ringen[a]* and Kari Suomi[b]**
[a] Department of Linguistics, 564 EPB, University of Iowa, Iowa City, Iowa, USA
[b] Phonetics, Faculty of Humanities, P.O.B. 1000, 90014 University of Oulu, Finland
* Names of the authors in alphabetical order.
** Corresponding author

Abstract
This paper presents the results of an investigation of the fortis-lenis contrast in Fenno-Swedish stops. The data show that in utterance-initial position, the two-way contrast is realised as a contrast between unaspirated and prevoiced stops. Word-medially and finally, the contrast is that of voiceless, unaspirated stops and fully voiced stops. Fenno-Swedish is thus another Germanic language, like Dutch, Afrikaans and Yiddish, that has a contrast between plain unaspirated and prevoiced stops in utterance initial position. The stop contrast in Fenno-Swedish differs from that of Central Standard Swedish, in two ways: (i) In Central Standard Swedish the contrast is between aspirated and prevoiced stops in utterance initial position, whereas it is between voiceless, unaspirated stops and prevoiced stops in Fenno-Swedish; in medial and final position, one series of stops in Central Standard Swedish is voiceless (aspirated, preaspirated or unaspirated) and the other series is fully voiced, whereas in Fenno-Swedish, one series is voiceless and unaspirated and the other voiced, i.e., in Fenno-Swedish there is no aspiration. (ii) Variation was observed for some Fenno-Swedish speakers in the production of /b d g/, with many tokens being completely voiceless and overlapping phonetically with tokens of /p t k/, whereas there is no overlap between the two stop categories in Central Standard Swedish. Stops in Finnish are voiceless and unaspirated. The fact that the aspirated stops have been lost in Fenno-Swedish, and the fact that there are voiceless occurrences of /b d g/ tokens in Fenno-Swedish suggests influence from Finnish. Fenno-Swedish thus appears to pattern with other languages influenced by a language with a different system of laryngeal contrasts.

Keywords: aspiration, Central Standard Swedish, Fenno-Swedish, Finnish, laryngeal features, prevoicing,Voice Onset Time

Table of contents

1. Introduction

In this paper we present the results of our investigation of the fortis-lenis contrast in the speech of twelve speakers of Fenno-Swedish in utterance-initial, medial and utterance-final stops. We use the terms *fortis* and *lenis* and the notations /p t k/ and /b d g/ as convenient labels to denote two stop series, possibly contrasting along what has traditionally been called the voicing dimension, without committing ourselves in any way to the exact phonetic nature of the stop series. Thus, although we use the term “fortis stops” (or e.g. “/p/”) in a number of languages, these terms and notations do not imply a claim that these (sets of) stops are phonologically or phonetically identical across the languages mentioned. In most varieties of Finnish, there is no contrast along the voicing dimension, and the stops occurring in such varieties are here referred to as *fortis* stops because they are typically voiceless, unaspirated. Hence, although no exact phonetic interpretation is implied by the term *fortis*, it does imply a lesser degree of voicing than the term *lenis*. In the text, we try to assess, in the light of empirical evidence, the phonetic properties of fortis and lenis stops in Fenno-Swedish.

Fenno-Swedish is a minority language that has been spoken, primarily in coastal areas of mainland Finland, since the Middle Ages: in Uusimaa/Nyland, Turunmaa/Åboland and Pohjanmaa/Österbotten (Ostrobothnia), see Figure 1. Mainland Finland (including any islands not belonging to Ahvenanmaa/Åland) is officially bilingual; Åland is an autonomous, monolingually Swedish-speaking province of Finland. It is unclear whether or not the Swedish spoken on Åland should be considered to be Fenno-Swedish. However, since we had no speakers from Åland, we do not take a stand on this question.

Figure 1. Map of Finland showing the areas in which Fenno-Swedish is spoken. 1 = Uusimaa, 2 = Turunmaa, 3 = Åland, 4 = Pohjanmaa. For Åland see the text.

For many centuries, mainland speakers of Fenno-Swedish have had, at most, few contacts with the Swedish spoken in Sweden, whereas today most of them have many contacts with Finnish. Since the 1970’s, both Finnish and (Fenno)-Swedish have been compulsory subjects at school for all children. According to the 2006 census, 5.1% of the inhabitants of mainland Finland were registered as native speakers of (Fenno)-Swedish. Except where Saami is an official language (in the four northernmost municipalities), there are four types of municipalities in Finland: monolingual Finnish or monolingual (Fenno)-Swedish, and bilingual municipalities with either Finnish or

(Fenno)-Swedish as the majority language. The proportion of mainland inhabitants living in monolingual (Fenno)-Swedish municipalities in 2006 was about 0.3%. Accordingly, most Fenno-Swedish speakers understand and speak Finnish, and many of them are functionally fully bilingual in Fenno-Swedish and Finnish.

1.1. Background

Given that the Baltic Sea separates most of Sweden from Finland and the fact that ancestors of Fenno-Swedish speakers first settled in Finland about eight centuries ago, it is not surprising that there are now many segmental and prosodic differences between Central Standard (CS) Swedish (the variety spoken in and around Stockholm) and Fenno-Swedish, see Kuronen & Leinonen (2000; 2008). Leinonen (2004a) convincingly argues that, today, CS Swedish has no influence on Fenno-Swedish pronunciation. One of the segmental differences between the two varieties, mentioned by Kuronen & Leinonen (2000), is that fortis stops at the onset of a stressed syllable are aspirated in CS Swedish, but unaspirated in Fenno-Swedish; for numerical data see Leinonen (2004b). Kuronen & Leinonen do not specify how the lenis stops are produced, except to say that the two sets of stops are distinguished by voicing. Recently Helgason & Ringen (2008) have described the results of an investigation of the fortis-lenis stop contrast in CS Swedish. They report that in utterance initial position the two-way stop contrast is between post-aspirated stops and prevoiced stops. Word-medially and finally, the contrast is between fully voiced stops and variably pre-aspirated or unaspirated stops. Since CS Swedish has prevoiced stops as well as aspirated stops, and Fenno-Swedish is known to have a two-way laryngeal contrast in stops, but to have no aspirated stops, it is not unreasonable to surmise that Fenno-Swedish might be another Germanic language, like Dutch, Yiddish, Afrikaans and CS Swedish, with prevoiced word-initial stops. Hence, one reason for undertaking this study was to determine what the phonetic nature of the Fenno-Swedish stop contrast actually is.
Finnish fortis stops are unaspirated at the onset of a stressed syllable and elsewhere (Suomi 1980). Leinonen (2004b) suggests that Fenno-Swedish voiceless/fortis stops are often pronounced as they are in Finnish.[1] He notes that indirect support for this suggestion is provided by Reuter (1977: 27) who states that "the [Fenno-Swedish] voiceless stops *p*, *t* and *k* are wholly or partially unaspirated […]. Despite this, they should preferably be pronounced with a stronger explosion than in Finnish, so that one clearly hears a difference between the voiceless stops and the voiced *b*, *d* and *g*" (translation by KS). As Leinonen points out, an implication of this normative exhortation is that speakers of Fenno-Swedish often pronounce the voiceless/fortis stops in the same way as do speakers of Finnish. But empirical evidence is needed to determine if indeed this is the case. To our knowledge there has been little discussion of the phonetic realization of Fenno-Swedish lenis stops or of influence from Finnish in the literature. Historically, Finnish lacks a laryngeal contrast in the stop system, the basic stops being /p t k/, which are voiceless unaspirated. For more details of the Finnish stop system see section 4.2. below.

Swedish is a quantity language in which quantity is contrastive only in stressed syllables. In most varieties, a short vowel in the stressed syllable must be followed by a consonant that is long (or by two different consonants). Conversely, if a long vowel in the stressed syllable is followed by a consonant, the consonant is always short. That is, there is a complementary durational relationship between the stressed-syllable vowel and the following consonant; a vowel in a stressed open syllable must be long. For example, in *väg* [veːg] 'road', the vowel is long and the following consonant is short (V:C), whereas in *vägg* [vegː] 'wall' the vowel is short and the following

[1] He says nothing about lenis stops except that some are devoiced when adjacent to voiceless stops.

consonant is long (VC:). These restrictions obtain for CS Swedish. This is also the dominant pattern in Fenno-Swedish, but there are also deviations in Fenno-Swedish that will be discussed below.

1.2. Influence of language/dialect contact on VOT

The speakers in our study are all fluent in Finnish, as was determined during the recording sessions. They are in contact, on a daily basis, with a language in which there is no aspiration and in which prevoicing is not often heard. Hence, we might expect that VOT in their Fenno-Swedish speech would have been influenced by Finnish in a way that is similar to what has been suggested in other studies of bilingual speakers or in cases of extensive dialect/language contact. In a recent study of Dutch, van Alphen & Smits (2004) show that prevoicing is only present in 75% of the productions of lenis stops by their subjects when reading a wordlist, a surprising result for a language which is supposed to contrast prevoiced stops with short-lag VOT stops in word initial position. One suggestion made by van Alphen & Smits is that Dutch is undergoing a sound change that may be caused or enhanced by the large influence from English through the educational system and the media. There are, however, also dialects of Dutch with aspiration and no prevoicing (Marc van Oostendorp, p.c.), so it may be that van Alphen & Smits' speakers have been in contact with speakers of these Dutch dialects, as well as with English and German speakers. In a study of Canadian French, Caramazza & Yeni-Komshian (1974) observed substantial overlap between the VOT distributions of lenis and fortis stops: only 42% of the lenis tokens were produced with prevoicing, while all fortis stops were produced without aspiration. The authors argued that the Canadian French VOT values are shifting as a result of the influence of Canadian English.

There is evidence that speakers are very sensitive to the VOTs they are exposed to. Nielsen (2006, 2007) has shown that speakers of American English produced significantly longer VOTs in /p/ after they were asked to imitate speech with lengthened VOTs in /p/ and the increased aspiration was generalised to new instances of /p/ (in new words) and to the new segment /k/. Sancier & Fowler (1997) showed that the positive VOTs in the speech of a native Brazilian Portuguese speaker were longer after an extended stay in the United States and shorter again after an extended stay in Brazil, and the authors concluded that the English long-lag stops influenced the amount of positive VOT in the speaker's native Brazilian Portuguese.

Since CS Swedish has both prevoicing and aspiration, whereas Finnish has no aspiration and only occasional prevoicing, we might expect that the differences in the laryngeal contrast in Finnish would have resulted in changes on the laryngeal contrast in Fenno-Swedish. Thus, another motivation for the present study was to determine whether there is evidence of influence on the laryngeal contrast in Fenno-Swedish that can be attributed to language contact and/or bilingualism.

2. Method

2.1. Subjects

We recorded 12 native speakers of Fenno-Swedish (6 males and 6 females) in Turku, Finland. The ages of the male speakers varied between 22 and 32 years, the female speakers were between 24 and 48 years old. Fenno-Swedish was the first language and the language of education of the subjects, as well as the language of both of their parents. The speakers came from all of the three main areas in which Fenno-Swedish is spoken on the Finnish mainland noted above. The speakers were paid for participating in the experiment.

2.2. Materials and recording

The speakers read a word list with 60 target words with both fortis and lenis stops (see the Appendix). Of the words without stop clusters in Helgason & Ringen (2008), we used all except five words (*bad*, *bebis*, *bibel*, *räd*, *tabbe*) but replaced them with six words not used by Helgason & Ringen (*bröd*, *bytt*, *kabel*, *nubbe*, *snabel*, *säd*). Native speakers (who did not participate in the experiment) confirmed that all target words are commonly known to speakers of Fenno-Swedish. The target words occurred twice in the list, with six filler words added to the beginning of the list. The recordings took place in an anechoic chamber at the Centre for Cognitive Neuroscience of Turku University. The words were presented to the subjects on a computer screen. The subjects received each new word by clicking the mouse and were instructed to click only when they had finished uttering a target word. The subjects were instructed to speak naturally. The microphone was placed in front of the subject at a distance of about 30 cm. Subjects' productions were recorded directly to hard disk (22.5 kHz, 16 bit) using high quality equipment.

2.3. Measurements

The onset (offset) of voicing was determined visually from broad-band spectrograms and oscillograms by identifying the onset (cessation) of periodic oscillation taken to indicate vocal fold vibration. VOT was measured, in utterance-initial stops, as the interval from post-silence onset of voicing to the onset of stop release (negative VOT), or from the onset of stop release to the onset of voicing (positive VOT). When both onsets occurred simultaneously, VOT = 0 ms. In medial stops, closure duration, VOT (= duration of voiceless portion after the offset of the occlusion) and VoiDur (= duration of voicing during the occlusion), were measured, and VdProp (= proportion of the voiced part of the total duration of the occlusion) was computed from closure duration and VoiDur. If voicing had ceased during the occlusion, VOT was defined as the interval from the onset of stop release to the onset of voicing; if voicing had not ceased during the occlusion and continued throughout the release burst to the following vowel, VOT was defined as 0 ms. For a stop of the latter kind, duration of voicing during the occlusion was equal to closure duration, and proportion of voicing during occlusion was defined as 100%. In final stops, closure duration, VoiDur and VdProp were measured in the corresponding manner.

3. Results

The grouping variables were place of articulation (labial, alveolar or velar), speaker sex (female or male), quantity (short or long consonant), speaker identity (12 speakers) and, in some comparisons, phonological voice (lenis or fortis). In what follows, we use the abbreviations Place, Sex, Quantity, SpeakerId and Voice. Place, Quantity and Voice were between-items factors, Sex was a between-subjects factor and SpeakerId a random factor. Univariate ANOVAs were performed on each dependent variable, with Tukey post hoc tests, where applicable. However, it was not possible to use Sex and SpeakerId as grouping variables in the same analysis as these variables are not orthogonal.

We pay little attention to the quantity opposition, as such, unless it interacts with the voicing characteristics of the stops. This is because the quantity opposition in Fenno-Swedish is a robust one (and it always had a highly significant effect in the analyses in which it was relevant), and hence Quantity is simply taken as a variable that may have an effect on voicing. Similarly we usually take the lenis-fortis contrast for granted by looking at these categories separately, but in studying tokens exhibiting phonetic overlapping between these categories, we analyse the categories together using the grouping variable Voice.

If, in reporting the results of statistical tests, a potential main effect and interaction is not mentioned, it means that the effect did not reach significance.

3.1. Utterance-initial stops

For utterance-initial stops, by necessity only VOT was measured (as it is impossible from acoustic records to determine e.g. the duration of the occlusion of stops preceded by silence). In order to be able to make relevant cross-linguistic comparisons with earlier studies, we wanted to examine precisely utterance-initial stops, not utterance-internal word-initial stops in which the situation is known to be different.

3.1.1. Utterance-initial lenis stops

For each speaker, 20 tokens of bilabial stops, 6 alveolar stops and 8 velar lenis stop tokens were analysed, or 34 tokens for each speaker.[2] Thus, altogether, 34 x 12 = 408 tokens of initial lenis stops were analysed. Only the variable SpeakerId had an effect on VOT [$F(11, 24) = 9.64$, $p < 0.001$], meaning that there were significant differences among the speakers. A great majority of the 408 lenis tokens (355 or 87%) had negative VOT, i.e. prevoicing, but the speakers differed among themselves with respect to the duration of the prevoicing and the proportion of tokens without prevoicing. Pairwise comparisons of the speakers with respect to the mean VOT values of all initial lenis tokens indicated that the speakers were statistically grouped into five homogeneous subgroups; speakers within a given subgroup are statistically non-distinct from each other, while the subgroups are statistically distinct from each. The speaker differences are summarised in Table 1.

Homogeneous subsets					Data for individual speakers				
1	2	3	4	5	M_All	M_Neg	No_N-Neg	M_N-Neg	M_Len-For
AS					-120	-145	3	11	158
ML					-119	-119	0	-	158
TB	TB				-106	-106	0	-	129
	JW	JW			-83	-83	0	-	112
		LS	LS		-70	-72	2	2	95
		LC	LC		-61	-75	4	10	89
		TO	TO		-59	-63	1	0	88
			ME	ME	-50	-75	8	16	74
			JA	JA	-42	-72	8	15	77
			KL	KL	-40	-52	3	6	79
				SL	-28	-49	11	13	57
				ML	-26	-52	13	11	61
All speakers					-67.0	-80.1	53	7.0	98.1

Table 1. Results for the utterance-initial lenis stops across the three places of articulation. The first five columns: The statistically homogeneous subsets into which the twelve speakers, here represented by their initials, were grouped according to their mean VOT across all initial lenis tokens; speakers within a given subgroup are statistically non-distinct, while the subgroups are statistically distinct from each other. In the remaining five columns data are given for individual speakers. M_All = mean VOT (in ms) computed across all tokens, M_Neg = mean VOT (in ms) computed across the negative values only, No_N-Neg = number of non-negative tokens, M_N-Neg = mean VOT (in ms) computed across the non-negative VOT values only, M_Len-For = mean difference of the VOTs computed across all initial lenis tokens and those across all initial fortis tokens by a given speaker. For more explanation see the text.

Thus, Table 1 shows e.g. that speaker TB did not statistically differ from speakers AS, ML and JW, but speakers AS and ML differed from speaker JW. Each subset consists of at least two speakers and eight of the twelve speakers belong to two subsets. The speakers were grouped into the

[2] For purposes of comparison, we wanted our target words to match those use by Helgason & Ringen (2009) as closely as possible. Helgason & Ringen decided to balance with respect to quantity, and hence were unable to balance the initial lenis stops with respect to place.

subgroups on the basis of M_All, but M_Neg shows a similar trend of diminishing absolute values from the top of the Table to the bottom. Note that this trend in M_Neg is independent of the number and VOT value of non-negative tokens produced by a speaker, because only the tokens with negative VOT were included in computing this variable. Consequently, both M_All and M_Neg show that, from the top to bottom of the Table, there is a shortening of prevoicing. In addition to this trend, there is the trend that speakers who have relatively short prevoicing also tend to produce tokens with non-negative VOT more often than do speakers who have relatively long prevoicing; speaker AS, with longest prevoicing, but with three non-negative tokens, is the most conspicuous counterexample to this latter trend.

As the mean duration of prevoicing in the lenis stops decreases in Table 1 from top to bottom, so does the mean distance between the lenis and the fortis tokens (column M_ Len-For) along the VOT continuum. For speakers in the upper part of Table 1 there is a wide margin between the contrasting sets of stops, whereas for speakers in the lower part the margin is narrower. This difference, ranging from 158 ms to 57 ms, is largely due to the differences in the extent of prevoicing of the initial lenis stops that were prevoiced, as the (non-significant) inter-speaker differences in the mean VOT values of the initial fortis stops were small, with a range only 16 ms long. In fact, there is a strong negative correlation between the variables M_Neg and M_Len-For [$r = -.983$, p (two-tailed) = 0.01].

3.1.2. Utterance-initial fortis stops

For each speaker, VOT was measured in 6 labial, 6 alveolar and 14 velar fortis stop tokens, thus altogether in 26 x 12 = 312 tokens of initial fortis stops. Only place had an effect [$F(2, 30) = 53.27$, $p < 0.001$]. Post-hoc tests indicated that /k/ had a longer mean VOT (48 ms) than /p/ (20 ms) and /t/ (23 ms) that did not differ from each other statistically. It is noteworthy, in view of the large inter-speaker variability in the initial lenis stops, that there was no such variation in the initial fortis stops.

3.1.3. Comparison of initial lenis and fortis stops

A histogram of all initial stops, collapsed into bins 5 ms wide, is shown in Figure 2. There were 53 lenis tokens that spanned the short-lag 0 ms to 36 ms VOT range (i.e., the lenis tokens that were not prevoiced). Within this same range, there were 128 fortis tokens. On the other hand, all initial tokens with a negative VOT were lenis stops, and there were no tokens of either category in the range from -24 ms to -1 ms. To assess whether or not there was complete overlapping of the lenis and fortis categories within the 0 ms to 36 ms VOT range, an ANOVA was run on these 181 tokens. The only effects observed were those of Voice [$F(1, 32) = 38.18$, $p < 0.001$] and Place [$F(2, 32) = 13.40$, $p < 0.001$]. There were no interactions. Post hoc analyses showed that, within both Voice categories, each place of articulation differed from the others. The means were 7 ms, 13 ms and 16 ms for /b/, /d/ and /g/, respectively, and 17 ms, 21 ms and 31 ms for /p/, /t/ and /k/, respectively. Thus, even within the short-lag VOT range in which there was occasional overlapping between the lenis and fortis tokens, the two sets were significantly different.

In addition, the PASW Statistics (Release 18.0.0) Discriminant Analysis, with its default settings, was used to classify the stops in the 0 ms to 36 ms VOT range, with Voice as the grouping variable. With only VOT as the independent variable, the percentages of correctly classified cases in the 0 ms to 36 ms VOT range were 69.8% for the lenis and 74.2% for the fortis stops (overall mean 72.9%). With both VOT and Place as independent variables, the corresponding percentages were 75.5%, 76.6% and 76.2%. The clear improvement of correct classification, especially of the lenis stops, following the addition of Place as an independent variable, shows that when Place information is taken into account, there is less overlapping than when such information is ignored.

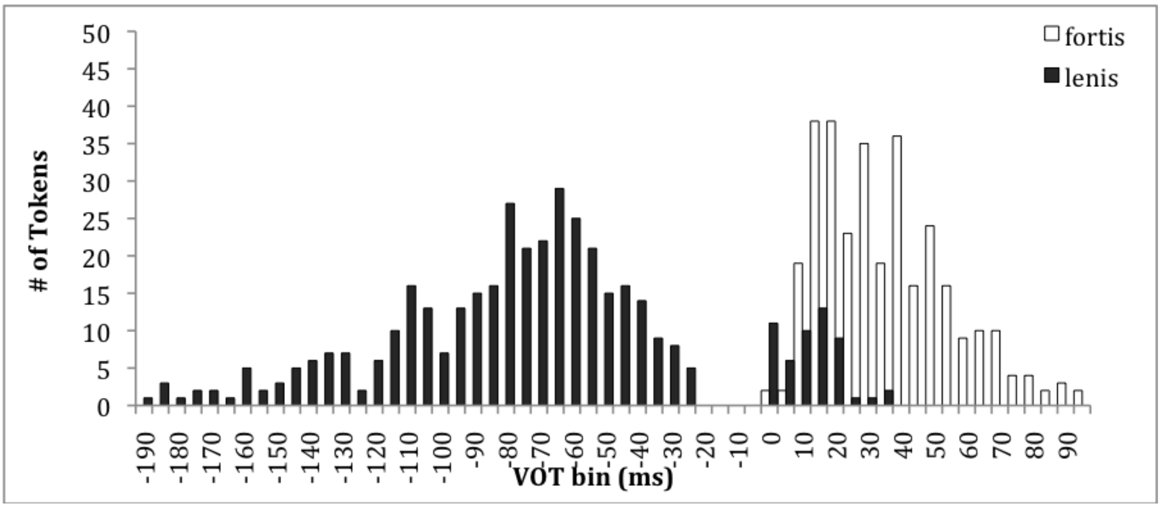

Figure 2. Histogram of the initial stops. Individual tokens have been collapsed into bins 5 ms wide.

These analyses show that, although there was overlap of the two sets of stops in the positive VOT range 0 – 36 ms, the two sets were not phonetically identical. On average, the lenis tokens within this range exhibited reliably shorter mean positive VOTs than the fortis tokens within the same range. Consequently, we conclude that the tokens of lenis stops with non-negative VOTs are also phonologically distinct from the fortis stops, they are renditions of /b d g/ stops as are those tokens with prevoicing (negative VOT). That is, speakers of Fenno-Swedish who produce occasional tokens of initial lenis stops with non-negative VOTs maintain, by and large, the lenis-fortis contrast although, in these tokens, the contrast is rather precarious. Thus it is not the case that the fortis-lenis contrast has been completely neutralised in some lenis tokens; the contrast has only become less robust (or, conceivably, it is based on phonetic parameters not investigated in this study).

3.2. Medial stops

For medial stops, closure duration, VOT (= duration of voiceless portion after the offset of the occlusion), VoiDur (= duration of voicing during the occlusion) and VdProp (= proportion of the voiced part of the total duration of the occlusion) were measured; however, the last parameter was meaningful for the lenis stops only, as the fortis stops were essentially all voiceless. If voicing continued throughout the occlusion into the following vowel, VOT was defined as 0 ms.

3.2.1. Medial lenis stops

For each speaker, 8 labial lenis stops, 8 alveolar stops and 8 velars stops were measured. Quantity was fully counterbalanced. An ANOVA was first performed on the closure durations with Sex, Quantity and Place as fixed factors. Sex had an effect [$F(1, 60) = 6.16$, $p < 0.05$], as did Quantity [$F(1, 60) = 628.88$, $p < 0.001$] and Place [$F(2, 60) = 4.24$, $p < 0.05$], and there was an interaction between Quantity and Place [$F(2, 60) = 4.24$, $p < 0.05$]. As can be seen in Table 2, the closure durations were always longer for the female speakers than for the male speakers. The expected quantity difference is obvious. The Place effect and the Quantity x Place interaction are due to the fact that in the long consonants the order of magnitude was /b/, /d/ > /g/, in the short consonants the order was /b/ > /g/ > /d/.

The means for VoiDur are also shown in Table 2. Sex had no effect ($F < 1$). Quantity had an effect [$F(1, 60) = 180.79$, $p < 0.001$] as did Place [$F(2, 60) = 9.57$, $p < 0.001$]. The Place effect is due to /b/ having longer VoiDur's than /d/ and /g/.

Sex very narrowly failed to have a significant effect on VdProp [$F(1, 60) = 3.92$, $p = 0.052$], as did Place [$F(2, 60) = 2.92$, $p = 0.062$]. Thus the male speakers tended to have larger VdProp's than the female speakers, and /b/ tended to have larger VdProp than /d/ and /g/.

		Long consonants			Short consonants		
		/b/	/d/	/g/	/b/	/d/	/g/
Female speakers	Closure	201	199	179	74	48	59
	VoiDur	168	120	136	73	43	41
	VdProp	85	63	77	99	91	69
	VOT	2	11	19	0	4	15
Male speakers	Closure	176	180	161	70	46	53
	VoiDur	165	130	146	69	40	47
	VdProp	95	78	93	99	89	91
	VOT	0	5	7	1	4	9
Mean	Closure	188	189	170	72	47	56
	VoiDur	166	125	141	71	42	44
	VdProp	90	71	85	99	90	80
	VOT	1	8	13	0	4	12

Table 2. The closure durations, VoiDur's (duration of voicing during the occlusion), VdProp's (proportion of the voiced part of the total duration of the occlusion) and VOTs of the medial lenis stops. VdProp's are given as percentages of closure duration, the other measures are given in ms.

Finally, Sex had an effect [$F(1, 60) = 4.88$, $p < 0.05$] on VOT as did Place [$F(2, 60) = 12.45$, $p < 0.001$]. The female speakers had longer VOT's than the male speakers, and /b/ had the shortest and /g/ the longest VOT.

Turning now to speaker differences, interesting among these are differences in VdProp because VdProp in effect overlooks any differences in speaking rate; moreover, there were no speaker differences in occlusion duration ($F < 1$) nor in VoiDur ($F < 1$). SpeakerId, as a random factor, had an effect on VdProp [$F(11, 60) = 6.76$, $p < 0.001$], and pairwise comparisons of the mean VdProp values of all medial lenis tokens indicated that the speakers were statistically grouped into three homogeneous subgroups as shown in Table 3 which also contains information on the distribution of the individual VdProp means, pooled across short and long stops and across the three places of articulation. In addition, Table 3 shows the speaker-specific means of VOT, similarly pooled, and the bottom row shows the overall means.

Since in intervocalic stops in which voicing continues throughout the occlusion the release burst is also usually voiced, but voiceless in stops in which voicing ceases during the occlusion, it would be expected that the VdProp values and the VOT values would correlate negatively with each other, and they did [$r(288) = -.581$, p (two-tailed) = 0.01]. As can be calculated from column "No_100" (with 24 tokens per speaker), the proportion of fully voiced intervocalic lenis stops was 66% across all speakers, ranging individually from 17% (speaker JA) to 100% (speakers JW and TO).

Homogeneous subsets			Data for individual speakers					
1	2	3	M_All	No_Zero	No_18-50	No_51-92	No_100	M_VOT
JA			53	4	7	9	4	9
KL			53	6	3	9	6	18
SL	SL		68	2	4	9	9	15
	ML	ML	84	0	2	12	10	9
	LC	LC	85	0	3	8	13	8
		AS	93	0	1	4	19	2
		TB	94	0	0	9	15	10
		LS	98	0	0	2	22	5
		ME	98	0	0	3	21	2
		MN	99	0	0	1	23	1
		JW	100	0	0	0	24	0
		TO	100	0	0	0	24	1
All speakers			85.4	12	20	66	190	6.7

Table 3. Results for the medial lenis stops across the three places of articulation. The first three columns: The homogeneous subsets into which the speakers were grouped according to their mean VdProp values computed across all medial lenis tokens. In the remaining six columns data are given for individual speakers. M_All = mean VdProp (in %) computed across all tokens, No_Zero = number of tokens with zero VdProp, No_18-50 = number of tokens in the VdProp range 18-50%, No_51-92 = number of tokens in the VdProp range 51-92%, No_100 = number of tokens in which VdProp was 100%, M_VOT = mean VOT.

3.2.2. Medial fortis stops

For each speaker, 10 labial stops, 8 alveolar stops and 8 velars stops were measured. Four of the labials, four of the alveolars, and two of the velars were preceded by a short vowel. In most varieties of Swedish (including CS Swedish) any consonant following a long vowel in a stressed syllable would be phonetically short. But in Fenno-Swedish a fortis stop following a long vowel is not always phonetically short. Eight of our speakers systematically produced phonetically short medial fortis stops after long stressed vowels, roughly in line with the pattern in CS Swedish, but four speakers (JA, KL, ME and ML, three of whom are female) systematically produced phonetically long fortis stops after long vowels. For example, *baka* ‘to bake’ was produced in one of two ways: either as [baːka] or as [baːkːa]. The [baːkːa] type of pronunciation is a well-known characteristic of Fenno-Swedish, and the [baːka] – [baːkːa] variation is determined by both regional and sociolinguistic factors. Since the present paper is about voicing rather than quantity, we postpone a detailed discussion of these quantity patterns to a comparison of the durational realisation of quantity in Fenno-Swedish, CS Swedish and Finnish (Helgason, Ringen & Suomi, in preparation). With respect to voicing, the results for the medial fortes were very straightforward: VoiDur (with a mean of 2 ms) exhibited no significant main effects or interactions. Hence, the medial fortes were always voiceless during occlusion. For the mean VOT values see the next section.

3.2.3. Comparison of medial lenis and fortis stops

On the whole, with respect to the presence of voicing during the occlusion, the lenis and fortis medial stops differed from each other clearly and as expected. Except for the voiceless (or almost completely voiceless) tokens, the lenis stops were usually robustly voiced during the occlusion, while the fortis stops were voiceless. As expected, the two sets of stops also differed from each other with respect to mean VOT: Voice had an effect on VOT [$F(1, 66) = 53.18$, $p < 0.001$] as did Place [$F(2, 66) = 26.27$, $p < 0.001$]. The means were 1 ms, 6 ms and 13 ms for /b/, /d/ and /g/, respectively, and 10 ms, 18 ms and 25 ms for /p/, /t/ and /k/ respectively.

While in 288 tokens (92%) of the fortis stops VoiDur was 0 ms, in the remaining 24 tokens VoiDur varied from 9 ms to 61 ms, a range also occupied by 92 lenis tokens (32% of all lenis tokens). These figures, based on a single parameter, suggest rather extensive phonetic overlapping of the two phonological categories. But the whole picture is different. A discriminant analysis, with closure duration, Place, Quantity, VoiDur and VOT as independent variables and again using default settings, correctly classified 99.4% of the fortis stops and 92.7% of the lenis stops (mean = 96.2%). With only VoiDur as the independent variable, the corresponding figures were 99.7%, 75.3% and 88.0%. The overall greater success of the classification exploiting all four measured parameters over the classification exploiting just one parameter is not surprising: what is somewhat surprising is the fact that the lenis stops benefitted so much more from the addition of the full set of parameters, while the percentage of correct classification of the fortis stops was minutely *higher* when only VoiDur was exploited.

3.3. Utterance-final stops

For the utterance-final stops, only closure duration and VoiDur were measured; VdProp was computed from these. In utterance-final position (in contrast to utterance-internal, word final position), it is of course impossible to measure VOT reliably from acoustic records.

3.3.1. Final lenis stops

For each speaker, 8 final lenis stops were analysed at each of the three places of articulation. In each place, half of the stops were short and the other half were long. The closure durations of final lenis stops are shown in Table 4. Sex had an effect [$F(1, 60) = 14.06$, $p < 0.001$] as did Quantity [$F(1, 60) = 183.58$, $p < 0.001$]. The female speakers always had longer closure durations than the male speakers, and Quantity had the predictable effect.

	Long consonants			Short consonants		
	/b/	/d/	/g/	/b/	/d/	/g/
Female speakers	157	183	174	86	85	87
Male speakers	141	158	131	78	68	64
Mean	149	170	152	82	77	76

Table 4. Mean closure durations (in ms) of the final lenis stops

Again, there was much variation in the degree of voicing, as captured by VoiDur and VdProp. In the following we analyse the variation using VdProp only. A circumstance that motivated this decision is that Quantity had no effect on VdProp [$F < 1$]; that is, any variation in VdProp is not due to Quantity, but reflects genuine inter-speaker differences. Sex had an effect [$F(1, 60) = 14.76$, $p < 0.001$], as did Place [$F(2, 60) = 9.25$, $p < 0.001$]. VdProp was always larger for the male speakers (mean = 83%) than for the female speakers (64%), and /d/ had smaller (58%) VdProp than /b/ (81%) and /g/ (80%), which did not differ from each other. But these computations across all speakers are partly meaningless in view of the large variability among the speakers and also within some of the speakers. SpeakerId as a random factor had a robust effect on VdProp [$F(11, 60) = 7.98$, $p < 0.001$], see Table 5.

Homogeneous subsets			Data for individual speakers					
1	2	3	Mean	0-20	21-40	41-60	61-90	100
SL			44	4	7	8	3	2
JA			46	3	9	7	4	1
LC			52	1	5	10	6	2
KL	KL		63	5	1	3	6	9
ML	ML		64	1	4	5	12	2
TO	TO	TO	75	2	2	2	9	9
TB	TB	TB	76	0	0	6	13	5
MN	MN	MN	77	0	0	6	9	9
	LS	LS	90	0	0	2	6	16
	AS	AS	94	0	0	2	2	20
		JW	98	0	0	0	2	22
		ME	99	0	0	0	1	23
All speakers			73.1	16	28	51	73	120

Table 5. VdProp results for the final lenis stops across the three places of articulation. The first three columns: The homogeneous subsets into which the speakers were grouped according to their mean VdProp values computed across all final lenis tokens. In the remaining six columns data are given for individual speakers. Mean = mean VdProp across all tokens, x-y = number of tokens in the range x%-y%, 100 = number of tokens in which VdProp was 100%.

The inter-speaker variation can be seen by comparing the speakers at the top and those at the bottom of Table 5. The three speakers at the top had low mean VdProp's, and the majority of their tokens were within the mid range 21 – 60%, with very few tokens in the fully voiced category. The speakers at the bottom in turn had nearly maximum means because a large majority of their tokens were in the fully voiced category. The three speakers at the bottom exhibited little intra-speaker variation, while the six speakers at the top all exploited the total range of variation available. Speakers TO, TB and MN had highly similar means, but their tokens were distributed differently along the semi-continuum, and similarly for speakers KL and ML.

3.3.2. Final fortis stops

For each speaker, 8 bilabial, 12 alveolar and 8 velar fortis stops were analysed in final position. Half of the bilabial and velar stops were short and half were long; 8 of the alveolar stops were long and 4 were short. Sex had an effect on closure duration [$F(1, 60) = 11.69$, $p = 0.001$] as did Quantity [$F(1, 60) = 168.98$, $p < 0.001$]. As can be seen in Table 6, the female speakers always had longer closure durations than the male speakers; also the numerical mean differences between the three places are shown despite the fact that they failed to reach statistical significance.

The VoiDur of final fortis stops was invariably zero (as was, consequently, VdProp). That is, the final fortis stops were always fully voiceless. In the final lenis tokens, on the other hand, there were seven tokens (2%) with zero VdProp. Thus, in the final position, there was very little phonetic overlap of the lenis and fortis stops, even when potentially relevant variables such as place of articulation and rate of articulation are overlooked.

	Long consonants			Short consonants		
	/p/	/t/	/k/	/p/	/t/	/k/
Female speakers	251	246	234	150	139	142
Male speakers	217	230	204	119	116	120
Mean	234	238	219	135	128	131

Table 6. Mean closure durations (in ms) of final fortis stops.

3.4. Lenis instability

Great statistically significant variation was observed in the degree of voicing of the lenis stops in each position investigated (while there was no such variation in the fortis stops). To get an overview of the different speakers' profiles in this respect, the speakers were given points according to their positions relative to the other speakers in Table 1 (the initial position), Table 3 (the medial position) and Table 5 (the final position). Thus speaker AS who was topmost in Table 1 was given 1 point for the initial lenis stops, ML, the next highest speaker in Table 1 was given 2 points etc., and ML, at the bottom of Table 1, was given 12 points. For the medial and final positions points were given in the same way except that the speaker at the top of the relevant Table received 12 points and the speaker at the bottom received 1 point. In this way a given speaker received 1 – 12 points for each position, with a larger number of points indicating a higher degree of voicelessness. The points thus given and their sum total for each speaker are shown in Table 7 in the order of magnitude of the sum totals.

	Initial	Medial	Final	Sum total
SL	11	10	12	33
JA	9	12	11	32
KL	10	11	9	30
LC	6	8	10	24
MN	12	3	5	20
ML	2	9	8	19
TO	7	1	7	15
TB	3	6	6	15
LS	5	5	4	14
ME	8	4	1	13
AS	1	7	3	11
JW	4	2	2	8

Table 7. The individual speakers' profiles based on each speaker's position, relative to the other speakers, with respect to the degree of voicing of the lenis stops in the initial position (Table 1), in the medial position (Table 3) and in the final position (Table 5), and the sum total of the points received at each position. The larger the numbers, the more voiceless the lenis tokens were on average. For further explanations see the text.

The profiles are somewhat crude in that the relative order of the speakers in each of the three Tables was based on just one parameter in each position. Thus the profiles do not capture, e.g., such differences in the distribution of the VOT values of the initial lenis stops that had no effect on the mean VOTs. Even so, clear differences among the speakers' profiles emerge. Speakers SL, JA and KL are close to the voiceless end of the voicing continuum in each position, while speaker JW exhibited the least overall voicelessness. Speakers may have almost the same sum total of points yet have very different profiles (e.g. ME and TB, MN and ML). So it was not the case that some speakers systematically favoured voicelessness in each of the three positions and other speakers did not (which is a conceivable scenario). Instead, the inter-speaker variability also varied across the three positions studied.

4. Discussion

In this paper, we have investigated parameters related to voicing in Fenno-Swedish stops. Let us first summarise the major results. The fortis stops were observed to be invariably voiceless unaspirated, with the typical properties of such stops in each of the three positions investigated (moderate positive VOT in utterance-initial position, voiceless occlusion in medial and final positions, and moderate positive VOT in medial position). There was no statistically significant inter-speaker variation in the production of the fortis stops in any position. The lenis stops, in contrast, while predominantly voiced, nevertheless exhibited extensive variation with some tokens exhibiting less voicing (occasional non-negative VOT in initial position, no or only partial voicing during occlusion in the other positions).

Fenno-Swedish is clearly another Germanic language, like Dutch, Afrikaans and Yiddish that contrasts fully voiced stops with voiceless unaspirated stops.

4.1. Cross-linguistic comparisons

What about other languages that have a two-way laryngeal contrast between fully voiced stops and voiceless unaspirated stops? We found that the initial fortis stops in Fenno-Swedish had positive VOTs of 20 ms for /p/, 23 ms for /t/ and 48 ms for /k/ and for intervocalic fortis stops, we found 10 ms for /p/, 18 ms for /t/ and 25 ms for /k/. Similar results are reported for Russian and Hungarian. Ringen & Kulikov (to appear) found positive VOTs in Russian initial fortis stops to be 18 ms for /p/, 20 ms for /t/ and 38 ms for /k/; for intervocalic fortis stops they found 18 ms for /p/, 18 ms for /t/ and 35 ms for /k/. Gósy & Ringen (2009) report that VOTs for Hungarian initial fortis stops were 10 ms for /p/, 16 ms for /t/ and 38 ms for /k/; for intervocalic fortis stops they report 18 ms for /p/, 20 ms for /t/ and 43 ms for /k/. Thus, with respect to fortis stops, Fenno-Swedish is similar to e.g. Russian and Hungarian.

But with respect to lenis stops, Fenno-Swedish differs from Russian and Hungarian. As noted above, we found that 87% of our speakers' initial lenis stops had prevoicing, a result very much like that reported for Dutch. Our results and those of van Alphen & Smits (2004) for initial lenis stops are different from what is reported for Russian and Hungarian; Ringen & Kulikov (to appear) report that 97.4% of the Russian initial lenis stops were prevoiced and Gósy & Ringen (2009) report that Hungarian speakers had prevoicing in *all* word-initial lenis stops. Our results for intervocalic stops are also different from those reported for Russian and Hungarian (van Alphen & Smits did not test intervocalic stops in Dutch). We found that only 66% of our speakers' intervocalic lenis stops were fully voiced; in contrast, Ringen & Kulikov report that over 97.5% of the Russian intervocalic lenis stops produced by their speakers were fully voiced, and Gósy & Ringen report that 95.5% of the lenis intervocalic stops produced by their speakers were *fully voiced.*

Other studies have found that languages with a [voice] contrast do not have robust prevoicing of initial stops. For example, in a study of Canadian French, Caramazza & Yeni-Komshian (1974) observed substantial overlap between the VOT distributions of lenis and fortis stops, with 42% of the lenis tokens produced with prevoicing, while all fortis stops were produced without aspiration. The authors argued that the Canadian French VOT values are shifting as a result of the influence of Canadian English.

We conclude that languages with a [voice] contrast may exhibit robust prevoicing (as found in Russian and Hungarian), but others may exhibit less prevoicing when speakers also speak another language without a [voice] contrast (Fenno-Swedish), when there is substantial influence from another language without a [voice] contrast and/or when the languages is undergoing a change (Dutch, Canadian French).

4.2. Influence from Finnish?

In the utterance-initial position, the mean VOTs of our speakers' fortis stops were 20 ms for /p/, 23 ms for /t/ and 48 ms for /k/. These values are clearly smaller than the ones reported by Helgason & Ringen (2008: 613) for CS Swedish, namely 49 ms for /p/, 65 ms for /t/ and 78 ms for /k/. At the same time, the Fenno-Swedish values are larger than those reported by Suomi (1980: 99) for Finnish word-initial /p/ (9 ms), /t/ (11 ms) and /k/ (20 ms) produced by ten male speakers. However, this latter difference may well be due to the fact that while the initial stops in our study were utterance-initial, the Finnish target words were embedded in a constant frame sentence. These differences are possibly due to the different elicitation method. At any rate, the Fenno-Swedish VOT values differ from the CS Swedish values in the direction of the Finnish VOT values.[3]

In the medial position the Fenno-Swedish VOTs were 10 ms for /p/, 18 ms for /t/ and 25 ms for /k/. For CS Swedish, Helgason & Ringen reported the values (here pooled across the sexes) 13 ms for /p/, 23 ms for /t/ and 31 ms for /k/, while Suomi (1980: 99) reported 11 ms for /p/, 16 ms for /t/ and 25 ms for /k/ for Finnish. Although the differences are small, it can be seen that the values are always largest in CS Swedish, and that either there is no difference between the Fenno-Swedish and Finnish values or they are negligible. What increases the distance of CS Swedish from Fenno-Swedish and Finnish is the observation by Helgason & Ringen of extensive preaspiration in CS Swedish medial and final fortis stops following a vowel in the stressed syllable; of Helgason & Ringen's six speakers, four exhibited substantial preaspiration exceeding 35 ms. Our speakers exhibited no traces of preaspiration, nor is preaspiration ever mentioned as a characteristic of Fenno-Swedish or of Finnish.

Fenno-Swedish fortis stops thus differ from CS Swedish fortis stops with respect to the duration of postaspiration and with respect to preaspiration. In all these respects Fenno-Swedish fortis stops are identical with or resemble Finnish stops. We suggest that this similarity is not accidental, but direct influence from the way fortis stops are produced in Finnish. It is noteworthy, in comparison to the lenis stops, that there was little variation in the Fenno-Swedish fortis stops, a state of affairs that also holds for CS Swedish (Helgason & Ringen 2008) and Finnish (Suomi 1980).

In the lenis stops, especially in initial and final positions, there was considerable variation in the degree of voicing (duration of prevoicing in initial position, proportion of voiced occlusion in medial and final positions). It is clear that the Fenno-Swedish speakers exhibited much more variation than the CS Swedish speakers of Helgason & Ringen. These authors report (p. 611) that, in initial position, the mean VOT for their speakers ranged from -58 ms to -132 ms while the range of means was -26 to -120 in the present study (see Table 1 above). For CS Swedish lenis stops in medial and final positions, Helgason and Ringen (2008: 614) report that "[v]oicing was robust for all subjects, and, in fact, all observed instances of stops were voiced to some degree, even for those [two] subjects who had the most frequent occurrences of partial voicelessness". Of our Fenno-

[3].See Fowler et al. (2008) for an interesting study of VOT in different groups of French-English bilinguals and monolingual speakers of English and French. A similar design would be difficult in the case of Fenno-Swedish as it is almost impossible to find monolingual speakers of Standard Fenno Swedish.

Swedish speakers three produced two to six lenis tokens (8% to 25% of the 24 lenis tokens per speaker) as completely voiceless in medial position (see Table 3), and in final position there were seven such tokens (2% of all lenis tokens). Moreover, Helgason & Ringen's Figure 6 (p. 615) shows that for all six speakers mean voicing duration in final position was either 100% or well over 80% of closure duration, while for the Fenno-Swedish speakers, these mean proportions (VdProp's) ranged from 44% to 99%, with eight speakers having a mean proportion less than 80% (see Table 5).

The Fenno-Swedish lenis tokens with long prevoicing in initial position and extensive voicing during closure in the other positions are very much like the vast majority of CS Swedish lenis tokens. But what distinguishes Fenno-Swedish from CS Swedish is the frequent occurrence, for many speakers of Fenno-Swedish, of mostly or completely voiceless tokens, with the consequent large variation or instability in the way lenis stops are produced. We suggest that this lenis instability, too, is due to influence from Finnish.

As noted above, Finnish is a language with no aspiration and in which prevoicing is rare. Standard Spoken Finnish has a type of /d/, a phoneme occurring both initially and medially, which is usually fully voiced but which, however, is not a plosive proper, rather it is something between a plosive and a flap; it is called a semiplosive by Suomi, Toivanen & Ylitalo (2008) who describe it in more detail (pp. 33-35). Under the influence of (mainly) English orthography, /b/ and /g/ are entering Finnish as separate phonemes, but not yet for most speakers, and not in all speaking situations. The number of words containing these stops is quite small in comparison to words with the corresponding fortis stops; for example, in the recent dictionary, *Suomen kielen perussanakirja,* words beginning with <b> take up eight pages, those beginning with <p> take up 213 pages. It is important to note that when words beginning with <b> or <g> are borrowed from English, Finnish speakers usually replace them with the native /p/ or /k/. Those who do not do this pronounce these consonants with prevoicing. In fact, given that Finnish /p/ and /k/ are voiceless unaspirated utterance-initially, a good way to keep words spelled with <b> and <g> distinct from /p/ and /k/ is to prevoice the former; if utterance-initial <b> and <g> were pronounced in the typical English manner (with a short-lag VOT), there would not be any distinction between them and the native Finnish stops. Consequently, then, Finnish is a language in which prevoicing is only rarely heard in true plosives.

We suggest that frequent exposure to Finnish has had an effect on the production of Fenno-Swedish lenis stops. Yet there are obviously limits to such an effect. Recall from section 3.1.3. above that even though there was some overlap of the initial fortis and lenis stops in the positive VOT range, 0 -36 ms, the two sets were not phonetically identical, except for a few outliers. Similarly in medial and final positions, speakers who produced occasional voiceless lenis tokens nevertheless maintained, by and large, the fortis-lenis contrast although, in these voiceless lenis tokens, the contrast is rather precarious. We interpret this to show that the need to maintain the contrast prevents the influence of Finnish from having full effect.

Only some of our speakers produced considerable numbers of more or less voiceless lenis tokens, while some speakers produced none. From the background information that is available to us (age, sex, place of birth), no pattern is visible that could explain the different behaviours of the speakers. Recall that different speakers produced voiceless lenis tokens in different positions: this suggests that there may be no simple explanation of the speaker differences.

Our results for the utterance-initial stops in a language in which lenis stops are predominantly prevoiced resemble the results of van Alphen & Smits (2004) for Dutch. Van Alphen & Smits observed that, overall, 25% of the Dutch lenis stops were produced without prevoicing by their 10 speakers, and, as in the present study, there was variation among the speakers: five of the speakers prevoiced very consistently, with more than 90% of their lenis tokens being prevoiced, while for the other five speakers there was less prevoicing and considerable inter-speaker variation; one speaker produced only 38% of the lenis stops with prevoicing. Van Alphen & Smits' list of target words contained words with initial lenis stops before (sonorant) consonants, as did ours. They found less prevoicing when the stops were followed by a consonant. We observed no such difference; the percentage of prevoiced tokens was the same, 87%, with and without the three words with initial lenis stops before consonants. This result is very similar to that observed for prevocalic lenis stops by van Alphen & Smits, which was 86%.

Against this background, it seems likely that Fenno-Swedish has been influenced by Finnish. While the existence and extent of the dialect and language contacts may be somewhat uncertain in the case of Dutch, it is clear that speakers of Fenno-Swedish have lived in Finland for centuries and, without doubt, most of them are in daily contact with speakers of Finnish. Kuronen & Leinonen (2011) emphasise, as does Reuter (1982), that Finnish influence on Fenno-Swedish was especially strong in eighteenth-century Turku (Åbo), at that time the administrative and educational centre of the country and where Finnish-speaking students had to learn Swedish. They also note that today there are few speakers of Standard Fenno-Swedish who cannot speak Finnish and that Fenno-Swedish has been increasingly influenced by Finnish, especially the pronunciation.

Fenno-Swedish contrasts voiced /b d g/ with voiceless unaspirated /p t k/. On the basis of our acoustic measurements and some knowledge of how these are related to glottal and supraglottal events, we conclude that the contrast in Fenno-Swedish is one of [voice] vs. no laryngeal specification. Suomi (1980: 165) concluded for the Finnish voiceless unaspirated /p t k/ that their "degree of voicing [is] completely determined by the supraglottal constrictory articulation", and that these stops have no glottal abduction or pharyngeal expansion gesture, a circumstance that leads to voicelessness of the occlusion (p. 155ff). Despite the different terminology, this amounts to concluding that the Finnish /p t k/ have no laryngeal specification as well. Thus in the two studies, thirty years apart, essentially the same conclusions were reached concerning the Fenno-Swedish and the Finnish /p t k/.

We are suggesting, then, that Fenno-Swedish is a language with a contrast between stops that are specified as [voice] and stops with no laryngeal specification.

A two-way laryngeal contrast between stops with negative VOT and short-lag VOT, as we found in Fenno-Swedish, is typologically more common than the contrast in CS Swedish between stops with negative VOT and long-lag VOT. Indeed, Lisker & Abramson (1964) did not find any language with a two-way stop contrast like CS Swedish in the languages they studied, and it has often been assumed that such languages do not exist. Helgason & Ringen (2008) and Beckman, Helgason, McMurray & Ringen (2011) have suggested that stops in CS Swedish are "overspecified", that is that one stop series is specified for [voice] and the other is specified for [spread glottis] – henceforth [sg] (the phonetic manifestation of [sg] is aspiration). If we assume that the Swedish spoken by the ancestors to today's Fenno-Swedish speakers was like CS Swedish, then it is clear where the [voice] feature of one of the one stop series came from – it has been there all along;[4] Fenno-Swedish has just lost the [sg] feature. It is hard to imagine that if the contrast in the earlier form of Swedish spoken in Finland had been between stops specified as [sg] and stops with no laryngeal specification, the present system, with a contrast of [voice], could have arisen through contact with a language *with no voiced stops*. Thus, the change that has apparently occurred in Fenno-Swedish is indirect evidence in support of the claim of Beckman et al. (2011) and Helgason & Ringen (2008) that CS Swedish has an overspecified contrast: one series is specified as [voice] and the other specified as [sg].

[4] Of course, an alternative is that the variety of Swedish spoken by the Swedes who settled in Finland had a two-way laryngeal contrast that is like that found in modern Fenno-Swedish. To our knowledge, there is no evidence that such a variety of Swedish was spoken by the early Swedish settlers in Finland.

Of course, the scenario just sketched does not explain why it was the [sg] feature that was lost rather than the [voice] feature. It is safe to assume that, at the time of the loss, Finnish had only voiceless unaspirated stops. We are assuming that the Finnish influence caused the aspirated series to become unaspirated. An alternative scenario would have been for the voiced series to become voiceless, unaspirated. We have no explanation for why aspiration was lost rather than voicing. Possibly aspiration is more susceptible to loss through contact with a language without aspiration than is voicing through contact with a language with no stop voicing. However, we have pointed out that even the voiced stop series seems to be being influenced by Finnish: the two stop series overlap to a greater extent than they do in languages like Russian (Ringen & Kulikov,to appear)

It should be noted that Norwegian appears to be changing from a system like CS Swedish to one with no voiced stops. Specifically, Ringen & van Dommelen (in preparation) have suggested that Norwegian is in the process of changing from a system like CS Swedish, with overspecified stops, to a language with only an aspiration contrast, i.e. it is losing its [voice] specification, possibly under influence of Danish which has aspirated stops but no voiced stops.

Acknowledgements

We are grateful to Mikko Kuronen and Kari Leinonen for their very useful and informative comments on earlier versions of this paper, and to Mikko Kuronen also for help with the transcriptions in the Appendix. We also want to thank Viveca Rabb and Urpo Nikanne for assistance in recruiting subjects, Maria Ek for help with many aspects of this project, Teemu Laine, Riikka Ylitalo, and Juhani Järvikivi for technical assistance, Heikki Hämäläinen for use of the lab, Pétur Helgason for valuable discussions and assistance with our word list, Vladimir Kulikov for making figure two, Lauren Eby and Vladimir Kulikov for assistance with manuscript preparation and, finally, our subjects. The research of C. Ringen was supported, in part, by a Global Scholar Award and a Stanley International Programs/Obermann Center Research Fellowship (2007) from the University of Iowa and NSF grant BCS-0742338.

Appendix

The target words in which the stop consonants occurred, grouped according to position and stop identity. The randomized list of target words was read twice. Some words contain two target stops. All words have stress on the initial syllable. As explained in the text, the disyllabic words with a long vowel in the stressed syllable followed by a fortis stop (e.g. *baka*) have an alternative pronunciation ([baːkːa]) in addition to the pronunciation here given ([baːka]).

Utterance-initial position

/b/		/d/		/g/	
baka	[baːka]	dagg	[dagː]	gap	[gaːp]
bryggt	[brykt]	däck	[dɛkː]	gapade	[gaːpadɛ]
bröd	[brøːd]	döpte	[døptɛ]	glapp	[glapː]
bygga	[bygːa]			gubbe	[gʉbːɛ]
byggde	[bygːdɛ]				
byggt	[bykt]				
byta	[byːta]				
bytt	[bytː]				
bytte	[bytːɛ]				
bädd	[bɛdː]				

/p/		/t/		/k/	
packa	[pakːa]	tak	[taːk]	kabel	[kaːbɛl]
prat	[praːt]	tappa	[tapːa]	kapa	[kaːpa]
puck	[pʉkː]	tub	[tʉːb]	klubb	[klʉbː]
				kläcka	[klɛkːa]
				kläckt	[klɛkt]
				kläckte	[klɛktɛ]
				kub	[kʉːb]

Medial position

/b/		/d/		/g/	
gubbe	[gʉbːɛ]	föda	[føːda]	bygga	[bygːa]
kabel	[kaːbɛl]	leda	[lɛːda]	slägga	[slɛgːa]
nubbe	[nʉbːɛ]	födde	[fødːɛ]	väga	[vɛːga]
snabel	[snaːbɛl]	ledde	[lɛdːɛ]	öga	[øːga]

/p/		/t/		/k/	
gapade	[gaːpadɛ]	byta	[byːta]	baka	[baːka]
kapa	[kaːpa]	bytte	[bytːɛ]	kläcka	[klekːa]
köpa	[tsøːpa]	sköta	[ʃøːta]	läka	[lɛːka]
släppa	[slepːa]	skötte	[ʃøtːɛ]	packa	[pakːa]
tappa	[tapːa]				

Final position

/b/		/d/		/g/	
klubb	[klʉbː]	bröd	[brøːd]	dagg	[dagː]
kub	[kʉːb]	bädd	[bɛdː]	lag	[laːg]
labb	[labː]	sladd	[sladː]	väg	[vɛːg]
tub	[tʉːb]	säd	[sɛːd]	ägg	[ɛgː]
/p/		/t/		/k/	
gap	[gaːp]	bytt	[bytː]	däck	[dɛkː]
glapp	[glapː]	fat	[faːt]	puck	[pʉkː]
läpp	[lɛpː]	fött	[føtː]	tak	[taːk]
rep	[rɛːp]	lett	[lɛtː]	vrak	[vraːk]
		prat	[praːt]		
		skött	[ʃøtː]		

References

Beckman, J., Helgason, P., McMurray, B. & Ringen, C. (2011) Rate effects on Swedish VOT: Evidence for phonological overspecification. *Journal of Phonetics 39*, 39-49,

Caramazza, A. & Yeni-Komshian, G. H. (1974). Voice onset time in two French dialects. *Journal of Phonetics*, *2*, 239-245.

Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A. & Halle, P. (2008). Cross language phonetic influences on the speech of French-English bilinguals. *Journal of Phonetics 3,* 649-663.

Helgason, P. & Ringen, C. (2008). Voicing and aspiration in Swedish stops. *Journal of Phonetics*, *36*, 607-628.

Gósy, M. and Ringen, C. (2009). Everything you always wanted to know about voicing in Hungarian. Paper presented at the International Conference on the Structure of Hungarian 9, Debrecen, Hungary.

Helgason, P., Ringen, C. & Suomi, K. (in preparation). Swedish Quantity: Central Standard Swedish and Fenno-Swedish. Ms., Uppsala University, University of Oulu, and University of Iowa.

Kuronen, M. & Leinonen, K. (2000). Fonetiska skillnader mellan finlandssvenska och rikssvenska. *Svenskans beskrivning 24*. Linköping Electronic Conference Proceedings. URL: http://www.ep.liu.se/ecp/006/011/.

Kuronen, M. & Leinonen, K. (2008). Prosodiska särdrag i finlandssvenska. In M. Nordman, S. Björklund, Ch. Laurėn, K. Mård-Miettinen & N. Pilke (Eds.), *Svenskans beskrivning 29*. Skrifter utgivna av Svensk-Österbottniska samfundet 70 (161-169). Vasa.

Kuronen, M. & Leinonen, K. (2011). *Historiska och nya perspektiv på svenskan in Finland*. Tampere: Juvenes Print – Tampereen Yliopistopaino Oy.

Leinonen, K. (2004a). *Finlandssvenskt sje-, tje- och s-ljud i kontrastiv belysning*. Jyväskylä Studies in Humanities 17, Jyväskylä University. PhD Dissertation.

Leinonen, K. (2004b). Om klusilerna i finlandssvenskan. In B. Melander, U. Melander Marttala, C. Nyström, M. Thelander & C. Östman (Eds.), *Svenskans beskrivning* 26 (pp. 179-188). Uppsala: Hallgren & Fallgren Studieförlag AB.

Lisker, L. & A. S. Abramson. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384-422.

Nielsen, K. (2006). Specificity and generalizability of spontaneous phonetic imitation. In: *Proceedings of the ninth international conference on spoken language processing (Interspeech)* (paper 1326), Pittsburgh, USA.

Nielsen, K. (2007) Implicit phonetic imitation is constrained by phonemic contrast. In J. Trouvain & W. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*. Universität des Saarlandes, Saarbrücken, Germany: 1961–1964.

Reuter, M (1977). Finlandssvenskt uttal. In B. Petterson & M. Reuter (Eds.), *Språkbruk och språkvård* (pp. 19-45). Helsinki: Schildts.

Reuter, M. (1982). Kvantitet i Helsingforssvenskan. Unpublished licentiate thesis.

Ringen, C. & Kulikov, V. (to appear). Voicing in Russian stops: Cross-linguistics implications. *Journal of Slavic Linguistics.*

Ringen, C. & van Dommelen, W. (in preparation). Voicing and aspiration in Norwegian stops. Ms., University of Iowa, Iowa City, and Norwegian University of Science and Technology, Trondheim , Norway.

Sancier, M. & Fowler, C. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, 25, 421-436.

Suomi, K. (1980). Voicing in English and Finnish stops. Publications of the Department of Finnish and General Linguistics of the University of Turku 10. Ph.D. Dissertation.

Suomi, K., Toivanen J. & Ylitalo, R. (2008). *Finnish sound structure*. Studia Humaniora Ouluensia 9. URL: http://herkules.oulu.fi/isbn9789514289842.
van Alphen, A., & Smits, R. (2004). Acoustical and perceptual analysis of the voicing distinction in Dutch initial plosives: The role of prevoicing. *Journal of Phonetics, 32,* 455-491.