Phonetic quantity as a social marker in urban Finland Swedish

Therese Leinonen

Abstract

Quantity has been considered one of the most salient social markers in Finland Swedish. Extra-long consonants in stressed V:C syllables have been characteristic for the speech of upper class Swedish in Helsinki. In contrast, rural Finland Swedish dialects are characterized by very short consonant durations in V:C sequences.

The aim of this paper is to study phonetic quantity in Swedish spoken in three cities in Finland: Helsinki, Turku and Vaasa. Spontaneous speech data from 40 speakers in two age groups were analyzed. The results show significant differences across the cities. V:/C ratios are significantly lower in Helsinki and Turku than in Vaasa. In Turku, there is a significant difference between older and younger speakers: the older speakers have similar values to the Helsinki speakers, whereas many of the younger speakers have values resembling those of Vaasa speakers.

Vowel duration is the most important cue for discriminating between V:C and VC: sequences in Helsinki and Turku. In Vaasa, on the other hand, vowel duration as well as consonant duration is used for differentiating V:C and VC: sequences.

1. Introduction

Swedish is described as a language with so-called complementary quantity, which means that every stressed syllable has either a long vowel /V:(C)/ or a short vowel followed by a long consonant /VC:/ (Elert 1964: 39). This quantity system is the result of the quantity shift, which took place during the Late Old Swedish period (starting around 1300, Riad 1992: 235ff.). Four types of sequences could occur in stressed syllables in Proto-Nordic: VC, V:C, VC: and V:C:. The quantity shift resulted in the loss of VC and V:C: sequences.

Dialects in different parts of the language area were affected differently by the quantity shift. All Swedish dialects in Finland, except for the ones on the Åland Islands, preserved VC sequences (Sw. *kortstavighet*, Ahlbäck 1971: 10). Also the extra-long sequences V:C: were preserved, at least partly, in many rural Swedish dialects in Finland; it was completely lost only in the westernmost parts of the Finland Swedish dialect area (i.e. Åland, western Åboland and southern Österbotten; Ivars 1988: 67).

The Finland Swedish standard language spoken in urban areas lacks phonemic V:C: sequences. VC sequences do occur in the urban varieties, but in a much more restricted set of words than in the rural dialects. These are mainly stressed instances of words that are normally unstressed

(functions words) and certain types of loanwords (Kiparsky 2008, Reuter 1980).

In contrast to Swedish, Finnish is a typical quantity language which allows two degrees of length for most sounds of the language in stressed as well as un-stressed positions. The Swedish settlement in Finland dates back to the 13th century (Reuter 1992: 102), which means that Swedish and Finnish have existed side by side for centuries, and they have influenced each other mutually. According to Reuter (1992: 108), the preservation of syllable types in Finland Swedish that have been lost in Sweden is "partly an archaic feature, but probably it is even more a result of influence from Finnish". Swedish spoken especially in the areas surrounding Helsinki has several prosodic resemblances with Finnish (Kim 2006, Reuter 1992: 109). Additionally, Finland Swedish has only minimal qualitative differences between long and short vowel pairs, whereas Standard Swedish has qualitative differences in addition to the quantitative difference (Reuter 1992: 108). The vowels in Finland Swedish have positions intermediate between Finnish vowels and Standard Swedish vowels in the F1–F2 acoustic space (Reuter 1971).

Not only phonemic quantity, but also phonetic segment durations show large variation in the Swedish language area in Finland. Long durations of voiceless obstruents in V:C sequences have been characteristic for higher-class Swedish in Helsinki and Turku, and quantity has been considered one of the most salient social markers in Finland Swedish (Reuter 1982: 197). In contrast to the long C durations in Helsinki and Turku, rural Swedish dialects in Finland are characterized by very short durations of voiceless obstruents in V:C sequences, significantly shorter than the corresponding durations in Swedish dialects in Sweden (Schaeffler 2005: 80–82, Reuter 1982: 194).

Acoustic studies of phonetic quantity in Finland Swedish have focused mainly on highly educated speakers from Helsinki (Reuter 1982) or on rural dialects (Schaeffler 2005). No comparable studies of quantity across different urban varieties of Finland Swedish exist so far. The aim of this paper is to study phonetic quantity in V:C sequences in three cities: Helsinki, Turku, and Vaasa. These three cities are the commercial and cultural centers of the three regions Nyland, Åboland, and Österbotten. The study is restricted to V:C sequences where the consonant is an intervocalic voiceless obstruent. Intervocalic consonants were chosen, because segmentation is facilitated, especially in spontaneous speech in varying recording environments, when the start of the following vowel can be used to identify the end of the consonant. The results will be more comparable

to those by Reuter (1982), who studied quantity in disyllabic as well as monosyllabic words, than to those by Schaeffler (2005) who only studied monosyllabic words. Reuter (1982: 89) found longer durations for voiceless obstruents in disyllabic than in monosyllabic words.[1]

2. Previous research

The discussion here focusses on V:C sequences where the consonant is a voiceless obstruent. The voiceless obstruents are the ones that have been observed to show large variation between urban and rural varieties of Finland Swedish on the one hand, and between Swedish spoken in Finland and Sweden on the other. Voiced consonants are generally shorter than voiceless ones in Swedish (Reuter 1982: 88, Elert 1964: 148), and the regional variation for voiced consonants is small compared to that of the voiceless ones. It might also be worth mentioning that voicing is not contrastive in Finnish and that Finnish obstruents are voiceless[2].

Quantity is a feature with strong social and geographic associations in Finland Swedish. Long intervocalic obstruent durations in V:C sequences have been characteristic for higher-class Swedish in Helsinki and Turku. As shown by Reuter (1982: 113–131) the variable is socially stratified in Helsinki, so that there is a correlation between perceived C duration and socio-economic status: the higher the socio-economic status the longer C durations. Reuter also found that speakers who grew up in Helsinki had longer C durations than speakers who had moved to Helsinki (from other parts of Finland) as grown-ups. Moreover, women had longer C durations than men in Reuters study in all social groups and among native Helsinki speakers as well as immigrants. The study, thus, shows the common sociolinguistic pattern that women use prestige forms to a higher extent than men. The variable can be regarded a sociolinguistic stereotype in Labov's (1972) terms, since it is often commented on and imitated by laymen (af Hällström-Reijonen 2010, Reuter 1982: 100).

One of the speakers in the current paper mentions the feature during the interview. The speaker is a young man from Helsinki and he is talking

1 The V:/C ratios in Schaeffler's (2005) study, especially the ones from Österbotten, exceeded the ones of disyllabic as well as monosyllabic words in Reuter's (1982) study of Helsinki Swedish.

2 Except for /d/ which occurs mostly under consonant gradation (weakening) and is phonetically "half-way between a plosive (and hence obstruent) and a flap-like resonant" (Suomi et al. 2008: 33). The consonants /b/ and /g/ do occur in loanwords but are not produced with voicing by all speakers of Finnish.

about what kind of Swedish is being spoken by his study mates at the college:

> *dä: fösö:kä: mensjå: (.) pra:ta me:r de hä nårma:la (.) helsingfårs språ:ke men de e ju int de dä: "ha:tta: må:ssa:" de hä grani grejjen (.) u:tan de e ju de e ju ganska så hä: (.) plein*[3]
> 'There people try to speak more kind of the normal Helsinki speech but, it is not this "hatar måsar" (Eng. *hate seagulls*) this Grani thing, but it is more kind of plain.'

Grani refers to the city Kauniainen (Sw. Grankulla) about 15 km to the west from the center of Helsinki. Kauniainen has been attractive to high earners because of low tax rates and has a posh reputation. The speaker is imitating what he calls the "Grani thing" and is using two words with V:C sequences with voiceless obstruents. He is using a nasal voice when imitating and is over exaggerating the length of the intervocalic consonants. The imitation is clearly depreciative and according to the speaker a more neutral pronunciation is used at his school. The speaker (He.ym.3) is, indeed, using a quite neutral pronunciation himself without extremely low or extremely high V:/C ratios (see Figure 2 and Table 4).

Helsinki Swedish pronunciation has traditionally had high status in the Finland Swedish community. It has been used as a norm, for example, by news readers in Finland Swedish public broadcasting. In recent years, however, there seems to have been a change in which variety is considered standard or 'the best language' by Finland-Swedes. Ivars (2003) and Östman & Mattfolk (2011) report that Helsinki Swedish is not regarded as Standard Swedish by people outside the Helsinki region (especially in the province Österbotten). Stenberg-Sirén & Östman (2012) and Stenius (2012) have noted decreasing use in media of features typically associated with Helsinki, including long intervocalic voiceless obstruents in V:C sequences.

Many researchers have tried to explain quantity patterns in Finland Swedish by contact with Finnish. The C durations in V:C sequences in Central Standard Swedish (CSS) are longer than in Finnish V:C sequences, and speakers with Finnish as their native language tend to perceive Swedish V:C sequences as V:C: (Hakulinen 1979: 28). Some researchers have accounted the long durations of voiceless obstruents in V:C sequences in Helsinki speech to interference among bilinguals (Hakulinen 1979: 28, Itkonen 1965: 261). Swedish was the only administrative language in

3 The phonetic notation is a rough notation with Swedish letters. It is a standard for Finland Swedish dialects used, e.g., in FO (vol. I, pp. XIV–XV, XVII).

Finland until the mid 19th century and the dominating language in higher education and administration until the beginning of the 20th century. Therefore, long C durations in V:C sequences could be seen as a substratum effect as Finnish native speakers when learning Swedish interpreted the Swedish V:C sequences as V:C:.

Comparable measurements of CSS and Helsinki Swedish are, however, scarce, and it remains unclear to what extent Helsinki Swedish C durations are in fact longer than in CSS. Measurements from different studies are usually not directly comparable due to differing experimental settings and different segmentation rules applied. At least under certain circumstances, lengthening of postvocalic short consonants can be observed also in CSS. Elert (1964: 166) reports that "in utterances consisting of isolated words, or lists thereof, all phonemes are lengthened, especially the consonants following after a stressed vowel. Even phonemically short consonants following long vowels have longer duration than the vowels". Helgason, Ringen & Suomi (2010) conducted experiments with six speakers from Stockholm as well as twelve Swedish-speaking Finns. The Finland Swedish speakers were from all three regions Nyland, Åboland, and Österbotten[4]. Eight of the speakers from Finland were reported to produce very short intervocalic stops in V:C sequences (rural type), while four speakers were reported to produce significantly longer intervocalic stops (urban type). The mean V:/C ratios from the experiment show that the rural Finland Swedish pronunciation (M = 2.39, N = 8) differs more from Stockholm Swedish (M = 1.22, N = 6) than the urban Finland Swedish pronunciation does (M = 1.00, N = 4). Inter- and intra-speaker variation was not reported, however, and it seems unlikely that the difference between Stockholm speakers and the urban type of Finland Swedish speakers would be statistically significant. Nonetheless, the results by Helgason et al. (2010) correspond to those of Reuter (1982: 108–112), who also reported slightly higher V:/C ratios for two speakers from Stockholm than for speakers from Helsinki.

The exceptionally short C durations in rural Swedish dialects in Finland have been explained by Finnish influence, too (Itkonen 1965: 261–262). V:C sequences in rural Finland Swedish dialects resemble Finnish V:C sequences more durationally than they resemble V:C sequences in CSS or Swedish dialects in Sweden. But the short C durations can also be explained language-internally. Disregarding any possible Finnish influence, Schaeffler (2005: 121–127) put forward a constraint-based account of

4 Geographic provenance was not exploited as an explanatory factor, however.

phonetic variation in segment durations in V:C and VC: sequences in Swedish dialects. According to Schaeffler, the durational variation is closely connected to the phonological systems. Dialects which have, or have until recently had, a 4-way quantity system (with VC, V:C, VC:, and V:C: sequences in stressed syllables) distinguish vowel length as well as consonant length well. In dialects with a 3-way system, where V:C: sequences are lost, C in V:C "has lost a durational constraint, as a C: after a V: does not exist", which allows longer C durations after V:. In the 3-way system, as in the 4-way system, C: in VC: is still long because it contrasts VC and VC:. In the final state, the 2-way system, VC is lost which allows overall shorter consonant durations, since the only contrast left is the one between V:C and VC:, which can be distinguished solely by vowel length.

Kiparsky (2008) proposed a stratal optimality theoretic analysis of quantity in varieties of Finland Swedish. Similarly to Schaeffler, he concluded that the short C durations in rural dialects are due to contrast preservation, since consonant duration contrasts V:C and V:C: sequences in these dialects. Intervocalic obstruents in V:C sequences in Helsinki Swedish were analyzed as geminates by Kiparsky.

Previous instrumental analyses of quantity in varieties of Finland Swedish have relied on carefully elicited data with the target sequences pronounced in isolated words or in short carrier sentences (Helgason et al. 2010, Schaeffler 2005, Reuter 1982). Factors like context, speech rate, and emphasis can be well controlled for in this kind of experiments. However, the real variation in spontaneous speech might not show up when speakers are reading a word list, and factors like prepausal lengthening may influence the results (cf. Elert 1964: 166). Especially when it comes to a socially sensitive feature, speakers might also be influenced by a formal experimental setting. Reuter (1982: 127) concluded from a perceptual evaluation of a socially stratified data set of Helsinki Swedish spontaneous speech, that markedly long and markedly short intervocalic C durations only occurred occasionally in spontaneous speech and most speakers produced intermediate C durations most of the time.

Studies by Heldner & Strangert (2001) on Swedish and by Pind (1999) on Icelandic (which has complementary quantity like Swedish) have shown that vowel/consonant ratios are not constant but tend to depend on speech rate and emphasis. Both studies showed that the least durationally varying segment is V in VC: sequences. When VC: sequences are lengthened, due to slower speech rate or stronger emphasis, it is the long consonant which is lengthened. In V:C sequences both the vowel and consonant are lengthened when speech rate decreases, but the variation in V:

duration is somewhat larger than in C duration. Hence, the phonemically long segment is lengthened the most in both sequence types, maximizing the quantity contrast between the long and the short segment in the sequence as well as between the two sequence types. Engstrand & Krull (1994) showed that speakers of the quantity languages Finnish and Estonian tend to preserve durational correlates of quantity to a greater extent in spontaneous speech than speakers of Swedish do.

An additional linguistic factor affecting segment length is vowel height. There is a universal tendency in the world's languages of close vowels being intrinsically shorter than open vowels (Lehiste 1970: 18); which has been demonstrated for CSS (Elert 1964), Finnish (Lehtonen 1970), and Helsinki Swedish (Reuter 1982). When it comes to consonants, voiced consonants are generally shorter than voiceless consonants (Elert 1964: 148), but otherwise intrinsic consonant duration seems to vary across languages (Lehiste 1970: 27–30).

In order to study the variation of the socially marked variable 'intervocalic voiceless obstruent duration in V:C sequences' in natural context, spontaneous speech data was used for this paper.

3. Data

The data comprises sociolinguistic interviews recorded in 2005–2007 in the project *Spara det finlandssvenska talet* ('Document the Finland Swedish speech'), carried out by the Society of Swedish Literature in Finland (Ivars & Södergård 2009, 2007). Each speaker was recorded during approximately an hour, and a sample of around 20 minutes per speaker has been transcribed for a spontaneous speech database. The recordings were made with a lapel microphone and a Hi-MiniDisc recorder. The recordings were done at 44.1 kHz sample rate and 16-bit amplitude resolution.

The study includes 40 speakers representing three cities and two age groups (see Table 1). The older speakers were born 1921–1953 (M = 1942, SD = 8.5), and the younger speakers were born 1974–1990 (M = 1983, SD = 4.1). All speakers grew up in the city they represent and in most cases their parents, or at least one parent, have a local background, too.

The number of speakers per site in the speech database depends on the number of inhabitants. Consequently, the number of speakers is the largest from Helsinki and fewest from Vaasa.

The voiceless obstruents that occur in inter-vocalic position in Swedish are /f/, /k/, /p/, /t/, /s/ and /ʃ/. Of the possible consonants, /f/ and /ʃ/ are very infrequent and occur only in loanwords and were therefore not

segmented. After the data had been segmented, it turned out that as many as 17 speakers lacked tokens of the consonant /p/. Because of this, all instances with /p/ as intervocalic consonant were left out of the analysis[5]. The consonants left in the analysis were /k/, /t/ and /s/.[6]

Table 1: Number of speakers per site, age group and gender

	old		young		
city	female	male	female	male	total
Helsinki	4	4	4	5	17
Turku	3	2	4	4	13
Vaasa	3	3	2	2	10
total	10	9	10	11	40
	19		21		

Four levels of prominence are usually distinguished in Swedish prosodic notation: no stress, secondary stress, primary stress/accent and focus accent (Bruce 2012: 183–185, Bruce 1994). Distinguishing between the four prominence levels in spontaneous speech has, however, turned out to be difficult even to trained phoneticians (Strangert & Heldner 1995). A decision was made to include only tokens representing the two highest levels of prominence. Hence, all tokens in the analysis have primary stress (according to the authors perception), but the words were not labeled for possible focal accent, since deciding between the two highest levels in the prosodic model has turned out to be especially problematic (Strangert & Heldner 1995). Segmentation was carried out using the segmentation rules applied in the SweDia project and described by Schaeffler (2005: 60–61). However, Schaeffler studied quantity in monosyllabic words and found it appropriate to mark the end of stops with the initiation of the closure burst. Because the present study concerns disyllabic words, silent phase as well as burst was included in the duration of stops.

The recordings were usually made in a rather silent room in the participant's home. Still, the recordings are of somewhat different quality, some including more background noise than others. Prior to segmentation, the

5 /p/ is intrinsically shorter than other stops in many languages (Lehiste 1970), e.g. in Swedish (Elert 1964).

6 Frequent words are, e.g., different tense forms of the verbs *läsa*, *försöka*, *sluta* and *prata*, and nouns like *saker*, *dator* and *huset*.

recordings were cleaned using the noise reduction function in Audacity®. Some of the recordings have reverberation which can make it hard to identify the transition from vowel to stop. Only tokens with distinct transitions between segments were segmented. Segmentation and acoustic measurements were done in Praat (Boersma & Weenink 2010).

4. Predictors of V:/C ratio

The V:C sequences were subjected to mixed-effects linear regression analysis in order to test which linguistic and extra-linguistic factors influence segment durations. The dependent variable of the analysis was V:/C ratio, i.e. relative segment durations. Linguistic fixed-effect factors tested were *vowel height*, *consonant*, and *local speech rate* (quantified as the total duration of the V:C sequence). Extra linguistic fixed factors were *city* and *age*. *Speaker* was treated as a non-repeatable random factor in the model.

The analysis was conducted using the `lmer()` function of the `lme4` library (Bates, Maechler & Bolker 2011) in `R` (R Development Core Team 2011). The base-line model included only the random-effect factor speaker as predictor of V:/C ratio. Additional factors were included step-by-step, and the increase in goodness of fit was measured by the log-likelihood ratio test (Baayen 2008: 253).

Table 2: Fixed effect coefficients of the mixed-effects regression model fitted to log(V:/C ratio)

	Estimate	Std. Error	*t*-value
Intercept	0.288	0.090	3.190
log(V:C duration)	0.169	0.092	1.832
consonant = *stop*	-0.139	0.026	-5.345
vowel height	-0.041	0.016	-2.641
city = *Turku*	-0.154	0.132	-1.165
city = *Vaasa*	0.362	0.126	2.880
age = *young*	-0.105	0.113	-0.933
log(V:C duration): *Turku*	0.490	0.142	3.444
log(V:C duration): *Vaasa*	0.294	0.147	1.998
Turku: *young*	0.511	0.173	2.948
Vaasa: *young*	0.153	0.188	0.814

The dependent variable V:/C ratio was log-transformed prior to the analysis in order to obtain normal distribution (One-sample Kolmogorov-Smirnov tests, V:/C ratio $D = 0.096$, $p < 0.001$; log(V:/C ratio) $D = 0.034$, $p = 0.409$). Likewise, the numeric predictor speech rate was log-transformed to reduce the effect of outliers, and additionally it was centered because it turned out that a by-speaker random slope for log(V:/C duration) was needed in the model (cf. Baayen 2008: 254–255).

The total number of tokens in the analysis was 647. The average number of tokens per speaker was 16.2 ($SD = 6.5$, $min = 6$, $max = 41$).

Table 2 shows the estimates and *t*-values of the fixed effects of the model. An absolute *t*-value higher than 2 indicates statistical significance at the 5% level (Baayen 2008: 248). The significance of the non-speaker-related main effects was additionally tested by adding by-speaker random slopes to the model. All main effects remained significant also in models including the corresponding by-speaker random slope. The only random effect which improved the model significantly was the by-speaker random slope for log(V:C duration), which was, accordingly, kept in the final model.

4.1 Linguistic factors

The linguistic factors consonant and vowel height both improved the goodness of fit of the model significantly. An initial analysis showed no significant difference between the consonants /k/ and /t/, which were consequently combined into one factor level, so that the factor consonant had only two factor levels corresponding to manner of articulation: fricative (/s/) and stop (/k/ and /t/).

Table 2 shows a significant main effect for consonant. The level fricative is mapped to the intercept and the model shows a negative estimate for the level stop, which indicates that the V:/C ratio is lower for stops than for the fricative. The effect size is, however, small (–0.139). A lower V:/C ratio could mean either shorter V: duration or longer C duration, or both. The mean duration of vowels before /s/ is somewhat longer ($M = 136$ ms, $SD = 45$, $n = 148$) than before stops ($M = 125$ ms, $SD = 45$, $n = 499$), while the consonant duration is equal (/s/: $M = 103$ ms, $SD = 29$; stops: $M = 103$ ms, $SD = 33$). Given the fact that a durational difference has to be at least 10% to be perceivable to humans (Rietveld & Van Heuven 2009: 221), the effect of consonant on the V:/C ratio is of merely theoretical interest.

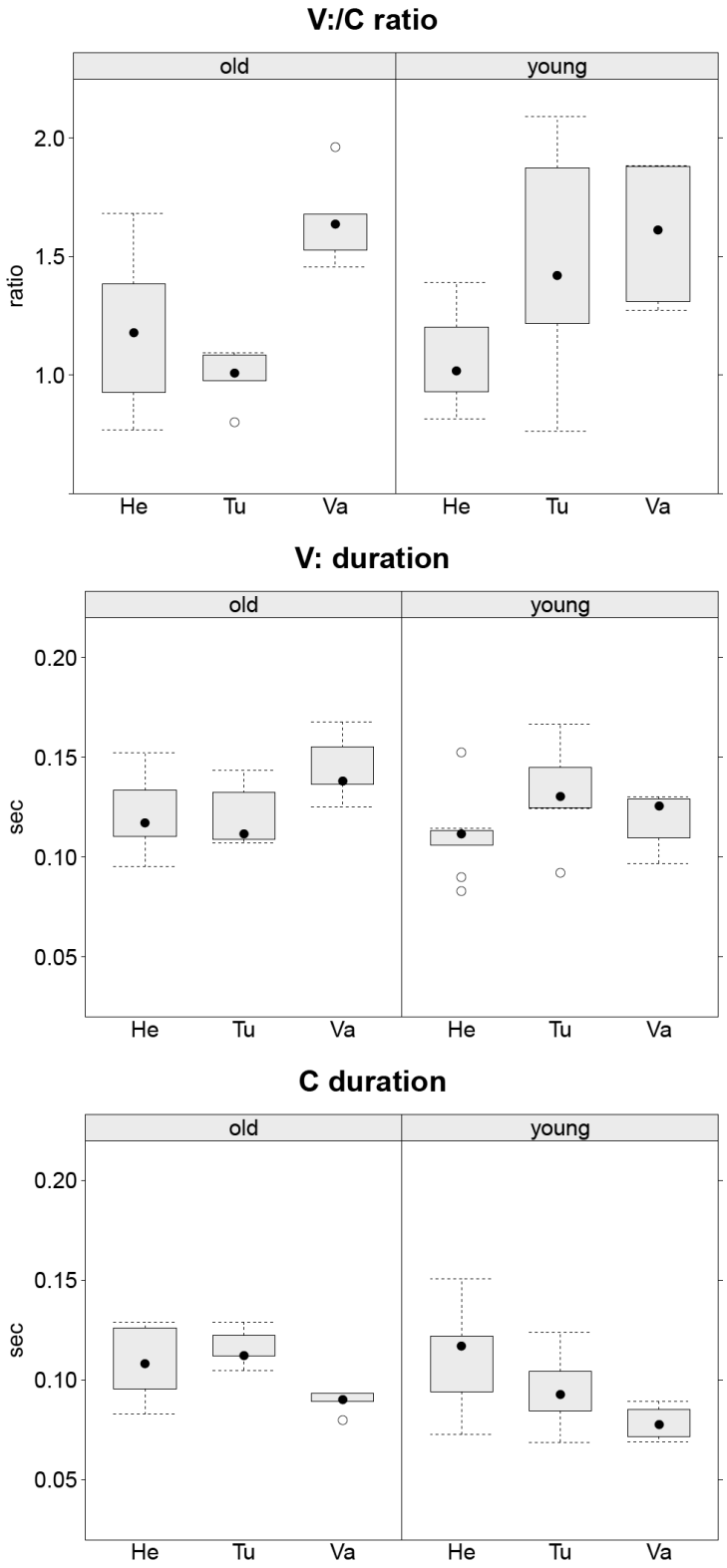

Figure 1: *Boxplots of V:/C ratio, V: duration, and C duration per city and age group. Median values per speaker were used for drawing the boxplots. For number of speakers per speaker group, see Table 1. (He = Helsinki, Tu = Turku, Va = Vaasa*

Vowel height was treated as a numeric factor with values 1 (= close: /i:/ /y:/ /o:/ /u:/; *n* = 197), 2 (= mid: /e:/ /ø:/ /o:/; *n* = 330), and 3 (= open: /a:/; *n* = 120). The effect of vowel height is significant, but even smaller than for consonant. Open vowels are longer than close vowels, which is in accordance with previous research (Lehiste 1970: 18). The average V: durations in the data set were 121 ms (*SD* = 42) for close vowels, 125 ms (*SD* = 48) for mid vowels (the contrast between close and mid vowels is not significant), and 143 ms (*SD* = 40) for the open vowel. C duration did not vary across the three degrees of openness of the preceding vowel.

4.2 Social factors

Of the three cities in the data set, Helsinki was mapped to the intercept. The contrast between Helsinki and Turku is not statistically significant, but Vaasa has a significantly higher V:/C ratio. There is also a significant interaction between city and age group. There is no significant effect of age in Helsinki and Vaasa, but in Turku younger speakers have significantly higher V:/C ratios than older speakers.

Figure 1 shows boxplots of the three measures V:/C ratio (no log-transformation), V: duration and C duration. The boxplots were drawn based on median values per speaker. A V:/C ratio of 1 means that V: and C in V:C sequences are equally long. Many speakers in Helsinki and Turku have data centered around a V:/C ratio of 1, and some speakers even have median values < 1, indicating that V: duration is shorter than C duration. The V:/C ratios in Vaasa are well over 1, showing that V: is clearly longer than the C. The values of many young speakers from Turku are more similar to the Vaasa values than to those of Helsinki and older Turku speakers. The variation is particularly large among young speakers from Turku. The large difference between older and younger speakers in Turku suggests an ongoing change.

The two lower boxplots in Figure 1 show that V: duration as well as C duration contributes to the large differences in V:/C ratio across the speaker groups.

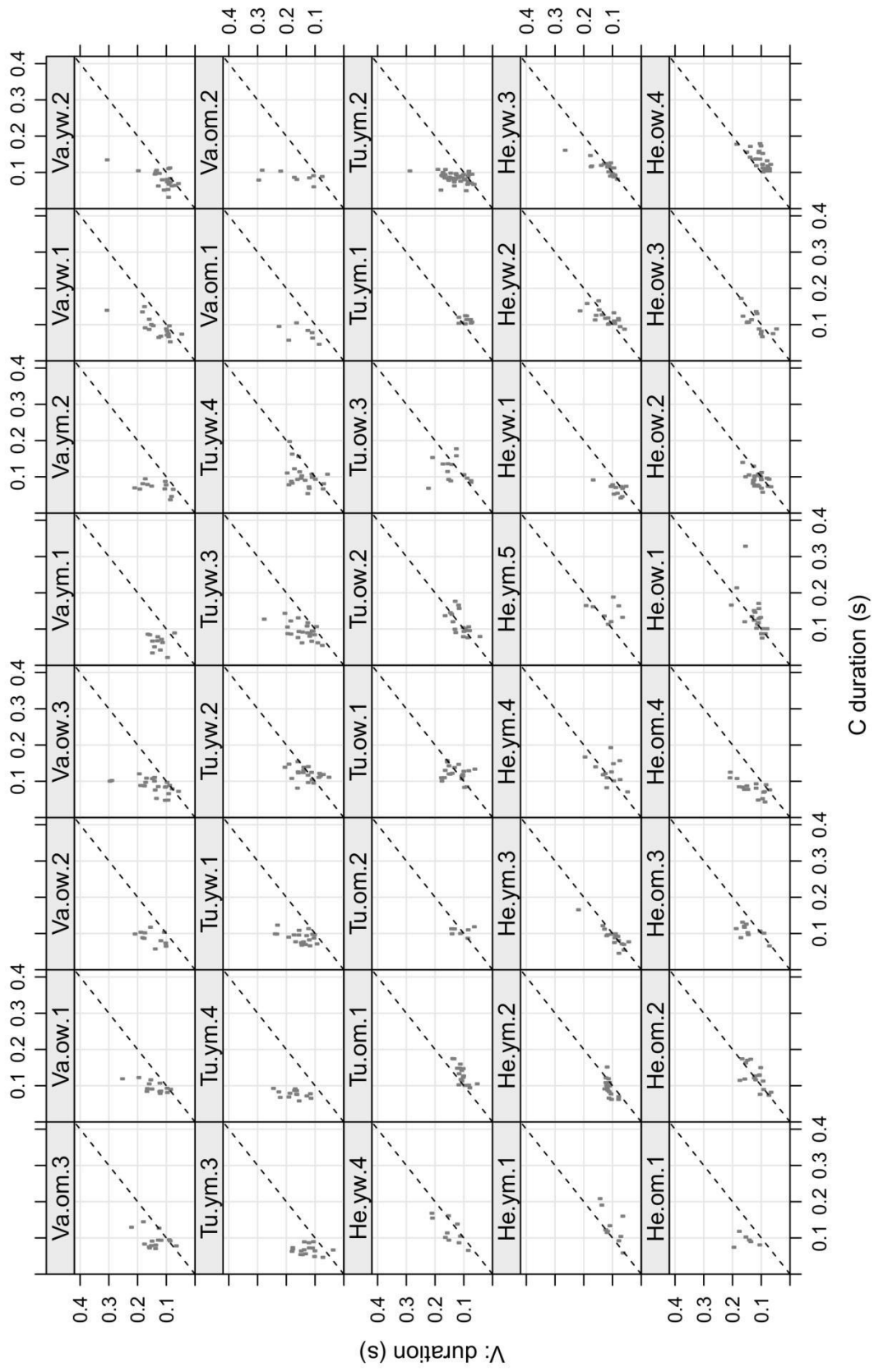

Figure 2: Scatter plots of V: duration and C duration plotted for each speaker separately. The dashed line indicates a V:/C ratio of 1.

4.3 Local speech rate

The model includes the factor log(V:C duration), with a significant interaction with city. V:/C ratios increase more with decreasing local speech rate in Turku and Vaasa than in Helsinki. The difference between the three cities is, hence, not only a difference in average V:/C ratio, but also a significant difference in how V:/C ratio is influenced by slower speech rate or emphasis. Speakers in Turku and Vaasa tend to make a clearer durational distinction between the long and the short segment at slower local speech rate than at faster speech rate, while speakers from Helsinki do not.

A by-speaker random slope for log(V:C duration) improved the model significantly, which means that the effect of local speech rate on V:/C ratio varies significantly across speakers. Figure 2 shows scatter plots of C duration plotted against V: duration for all 40 speakers separately. The dashed line in the plots indicates a V:/C ratio of 1, e.g. equally long V: and C. The data of many speakers (such as Tu.ym.4, He.om.4 and most of the data from Vaasa) fall above this line, indicating that V: is always longer than C.

However, there are also speakers (such as Tu.om.1 and He.ow.4) for which most of the data points are below the line, which means that the consonant is actually longer than the vowel in phonemic V:C sequences. The differences in the effect of speech rate are also evident. Speakers such as Tu.ym.2 and Va.om.2 clearly increase V: duration when the total duration of the sequences increases, while others (such as He.ow.1) vary C duration more than the V: duration. For a few speakers (such as He.ym.3), the data is centered around the dashed line, indicating that vowel and consonant are equally long and both increase equally much with decreasing speech rate, keeping the V:/C ratio constant.

4.4 Random intercepts

The only social factors in the fixed-effect part of the model are city and age. Additional social factors of interest, based on previous literature, would have been gender and socio-economic status. Because of the small number of speakers in each group, these variables could not be tested quantitatively. When it comes to socio-economic status, the database, moreover, does not provide enough information for all speakers.

One way to study random intercepts is to use the random intercepts from the full model with all the fixed effects displayed in Table 2. Another way is to completely leave the social grouping factors out of the model and let the random intercepts include all speaker-related variation including

the effect of city and age. Both methods are of interest. Table 3 shows the random intercepts of the full model (the one in Table 2). The table is grouped according to city and age, since the variation according to these two factors is taken care of by the fixed factors of the model. The speakers are sorted according to increasing random intercept within each group. In practice, this means that the speakers with the highest V:/C ratios within each group are on the bottom of each list.

Table 3: Random intercepts of the mixed-effects regression model in Table 2

Hf.ym.5	-0.319	Ab.ym.1	-0.408	Va.yw.1	-0.181
Hf.ym.1	-0.192	Ab.yw.2	-0.381	Va.yw.2	-0.163
Hf.ym.4	-0.159	Ab.yw.4	-0.062	Va.ym.2	0.155
Hf.yw.2	-0.069	Ab.yw.3	-0.011	Va.ym.1	0.223
Hf.ym.3	0.063	Ab.ym.2	0.056		
Hf.yw.3	0.082	Ab.yw.1	0.125		
Hf.ym.2	0.164	Ab.ym.4	0.310		
Hf.yw.4	0.188	Ab.ym.3	0.342		
Hf.yw.1	0.239				
Hf.ow.4	-0.399	Ab.om.1	-0.165	Va.om.3	-0.102
Hf.ow.1	-0.165	Ab.ow.1	-0.015	Va.ow.1	-0.099
Hf.om.2	-0.157	Ab.ow.2	0.029	Va.ow.2	-0.033
Hf.ow.3	-0.031	Ab.om.2	0.070	Va.ow.3	-0.009
Hf.ow.2	0.041	Ab.ow.3	0.109	Va.om.2	0.033
Hf.om.3	0.074			Va.om.1	0.176
Hf.om.1	0.269				
Hf.om.4	0.372				

Table 4 displays random intercepts from an analysis identical to the one in Table 3 except for the fact that the factors city and age are not included in the fixed part of the model. The speakers are sorted according to increasing random intercept, so any patterns relating to the factors city, age or gender are only a result of the ranking of the random intercepts.

Most speakers from Helsinki are found at the top of Table 4, while most speakers form Vaasa are found at the bottom of the table. The older Turku speakers are found among the Helsinki speakers at the top of the table, while some young Turku speakers are at the bottom of the table. This only repeats the results reported by the fixed factors in the original model. The interesting part about Table 4 are the speakers who go against these main grouping patterns.

Table 4: Random intercepts of a regression model without city and age as fixed effects

city	age	gender	nr	intercept
Helsinki	young	man	5	-0.552
Helsinki	old	woman	4	-0.512
Turku	old	man	1	-0.425
Helsinki	young	man	1	-0.411
Helsinki	young	man	4	-0.373
Turku	young	man	1	-0.316
Helsinki	old	woman	1	-0.280
Helsinki	old	man	2	-0.273
Helsinki	young	woman	2	-0.271
Turku	old	woman	1	-0.257
Turku	young	woman	2	-0.241
Turku	old	man	2	-0.214
Turku	old	woman	2	-0.171
Helsinki	old	woman	3	-0.164
Helsinki	young	woman	3	-0.126
Helsinki	young	man	3	-0.119
Turku	old	woman	3	-0.087
Helsinki	young	man	2	-0.049
Helsinki	old	woman	2	-0.046
Helsinki	old	man	3	-0.036
Helsinki	young	woman	4	-0.015
Turku	young	woman	4	0.035
Turku	young	woman	3	0.123
Vaasa	young	woman	2	0.125
Vaasa	young	woman	1	0.128
Helsinki	young	woman	1	0.131
Vaasa	old	woman	1	0.150
Helsinki	old	man	1	0.180
Vaasa	old	man	3	0.202
Turku	young	man	2	0.205
Vaasa	old	woman	2	0.234
Turku	young	woman	1	0.250
Vaasa	old	woman	3	0.267
Vaasa	old	man	2	0.310
Helsinki	old	man	4	0.317
Vaasa	young	man	2	0.417
Vaasa	old	man	1	0.420
Turku	young	man	4	0.462
Vaasa	young	man	1	0.491
Turku	young	man	3	0.492

Table 3 shows that there seems to be some gender related variation. In Helsinki, the women seem to have the lowest V:C ratios among older speakers, but among the young speakers it is the other way around. The values are only comparable within each group in Table 3, but in Table 4 we see that young male speakers and older female speakers from Helsinki indeed have quite similar values.

Among the young Turku speakers, the top of the list in Table 3 is dominated by women and the bottom of the list by men. When looking for these speakers in Table 4, the young men from Turku are found at the bottom among the Vaasa speakers, while the women have intermediate values. The young women from Helsinki and Turku actually have quite similar values and are found somewhere in the center of the list in Table 4, while it is the young men from these two cities who make the groups so different. The young men from Helsinki have very low V:/C ratios while the young men from Turku have very high V:/C ratios. One of the young male speakers from Turku (Tu.ym.1) clearly breaks this pattern, but the available background data about this speaker does not include any information which would shed light on why he behaves differently than the rest of his group.

The factors gender and socio-economic status should be tested quantitatively with a larger data set than the one that has been available for this paper. The qualitative analysis of the random intercepts suggests that these factors, which were found to be significant for C duration in Helsinki by Reuter (1982), are still of some importance also in the 21st century. The most striking difference in comparison to Reuter's results is that young men in Helsinki, and espcially men with low educational level, have lower V:/C ratios than young women. The older generation of Helsinki speakers in this paper behaves more in conformance with Reuter's results.

5. V:C ~ VC: distinction

Figures 1 and 2 show that for some speakers C duration is longer than V: duration in phonemic V:C sequences. The question then arises how V:C sequences are distinguished from VC: sequences. In contrast to Central Standard Swedish[7], Finland Swedish has only minimal qualitative dif-

7 Pairs of long and short vowels have a qualitative difference in addition to the durational difference in CSS. Quantity still seems to be the major cue for distinguishing between long and short vowels, with vowel quality providing an additional cue on-

ferences between long and short vowels, which means that duration is the only available cue for distinguishing long and short segments (Reuter 1992: 108). Hence, a durational difference is expected between V:C and VC: sequences.

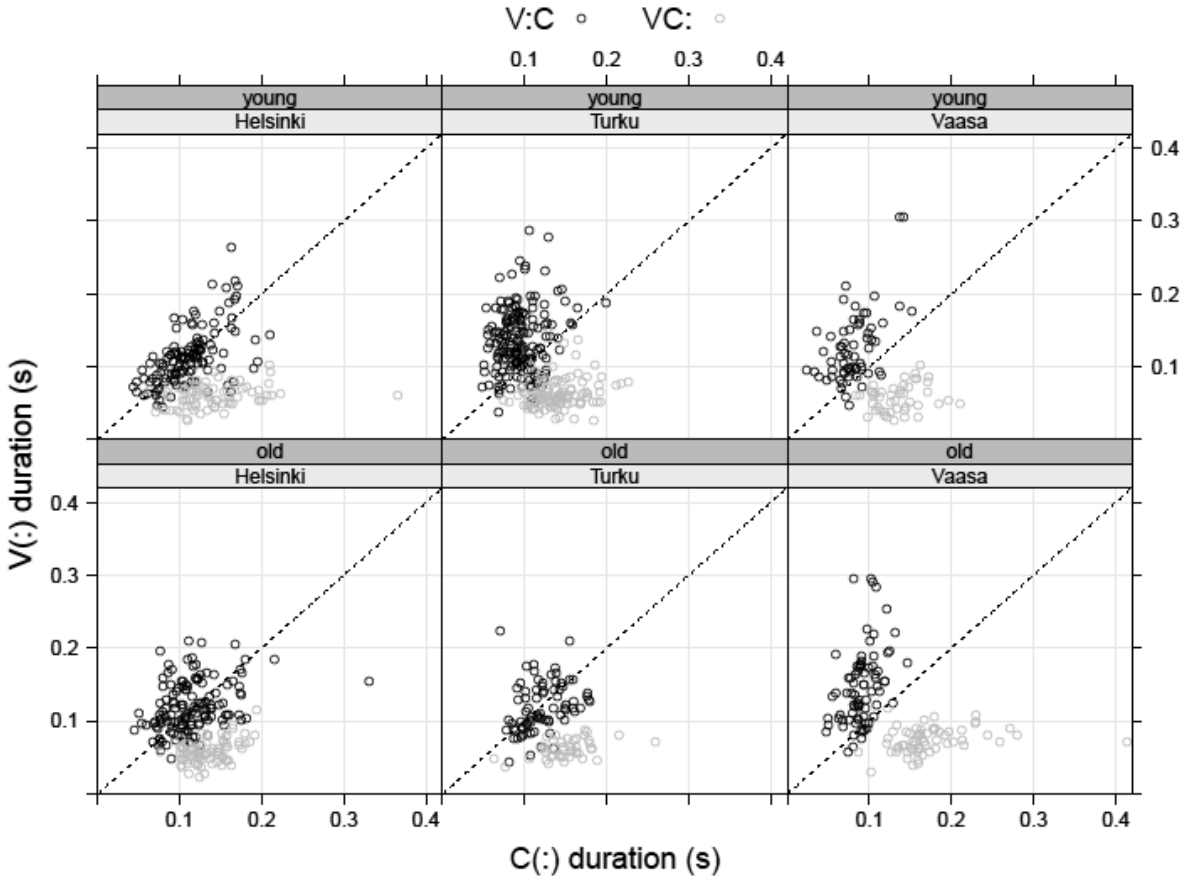

Figure 3: Scatter plots displaying the duration distributions of vowels and consonants in V:C (black) and VC: (gray) sequences in the six speaker groups. There are evident differences in the separation of the two sequence types across the speaker groups.

Figure 3 displays scatter plots of V:C and VC: sequences for the three cities and two age groups. There are some obvious differences between the speaker groups. The older speakers from Vaasa have the clearest durational separation between V:C and VC: sequences, while there is more overlap between the two types of sequences for other groups of speakers. Older

ly for some vowel pairs but not all (Behne et al. 1996, Hadding-Koch & Abramson 1964).

speakers from Vaasa seem to use vowel duration as well as consonant duration to distinguish the two sequence types, while only vowel duration seems to be used by the older Helsinki and Turku speakers.

A number of logistic mixed-model regression analyses were carried out in order to find out which cues are used for distinguishing the two sequence types. The analyses were carried out separately for all six groups of speakers. The dependent variable in all analyses was the categorical variable sequence type with the two levels V:C and VC:. All analyses included a random effect for speaker. Five different models with varying independent variables were tested:

1. vowel duration
2. consonant duration
3. sequence duration, i.e. the total duration of the vowel-consonant sequence
4. vowel/consonant ratio
5. segment durations, i.e. two IVs: vowel duration and consonant duration

Somers' D_{xy} (cf. Baayen 2008: 204, 281) was used for testing how well each of the models explains the data. The results in Table 5 show that vowel duration is a very good predictor of sequence type in all six speaker groups. Consonant duration is a poor predictor in Helsinki and among older speakers in Turku, but a good predictor in Vaasa and among younger speakers in Turku.

Table 5: Somers' D_{xy} values for five different regression analyses ran separately for all six speaker groups. The statistic indicates the correlation between predicted probabilities and observed values. 0 = randomness, 1 = perfect prediction.

	age	*n* tokens			V(:)	C(:)	V(:)C(:)	V(:)/C(:)	V(:) dur+
city	group	total	V:C	VC:	dur	dur	dur	ratio	C(:) dur
Helsinki	old	220	141	79	0.954	0.606	0.465	0.990	0.993
	young	209	136	73	0.886	0.560	0.396	0.962	0.963
Turku	old	126	82	44	0.902	0.590	0.302	0.966	0.968
	young	286	183	103	0.930	0.880	0.323	0.991	0.997
Vaasa	old	155	93	62	0.905	0.978	0.344	0.999	1.000
	young	109	63	46	0.936	0.875	0.295	1.000	1.000

Sequence duration is generally a very poor predictor, indicating that there is a tendency to keep the total duration of vowel-consonant sequences constant. Sequence duration is a slightly better predictor among older speakers in Helsinki than for the other speaker groups. The older Helsinki-speakers have somewhat longer average V:C sequences than VC: sequences, due to lengthened C.

All speaker groups get slightly higher D_{xy} scores with V(:)/C(:) ratio as IV than when using only vowel duration. Models using two durational IVs (vowel duration and consonant duration) are equally good as the models using only a single relative value (vowel/consonant ratio) as IV, indicating that the relative value is the most important cue. For the Vaasa speakers, models with vowel/consonant ratio and with vowel duration + consonant duration as explanatory factors give a perfect fit of the data.

6. Discussion

The results of this study show large differences in the phonetic correlates of phonemic quantity in the three cities Helsinki, Turku and Vaasa. V:/C ratios are significantly higher in Vaasa and among younger speakers in Turku than in Helsinki and among older speakers in Turku.

Vowel duration is the main cue for the phonemic distinction between V:C and VC: sequences. All speaker groups studied use vowel duration as a phonemic cue, while consonant duration is used only by some of the speakers. This is in accordance with results by Schaeffler (2005: 125) on Swedish dialects.

According to Schaeffler's (2005) typology of phonological quantity in Swedish dialects (described in Section 2), the urban varieties of Finland Swedish can be described as having a 3-way quantity system (with VC, V:C and VC: sequences in stressed syllables). This would predict well-distinguished vowel length (high V:/V ratio), long C: durations, and no short/long effect on C duration (Schaeffler 2005: 122). This is exactly what can be observed for Helsinki speakers and older Turku speakers: vowel length is a good predictor of the V:C~VC: contrast, while consonant length is a poor predictor.[8] According to Schaeffler's constraint-based the-

8 The VC~VC: contrast has not been studied in this paper, so no conclusions can be drawn at this point about how exactly VC: sequences are distinguished from VC sequences. A small-scale study by Reuter (1982: 173–175) suggests that the total duration of VC sequences in Helsinki speech is less than 50% of V:C and VC: sequences. The duration of V in VC is on average shorter than V in VC:, and the du-

ory, Vaasa has a system with "durational overspecification" (Schaeffler 2005: 126) since both vowel and consonant durations are well-distinguished even though the system lacks V:C: sequences. This could be explained by the fact that Vaasa is surrounded by rural dialects which have, or have until recently had, 4-way systems. The durational adaptation to the new 3-way system can according to Schaeffler proceed very slowly over several generations.

In Helsinki and Turku, the older speakers show sociolinguistic patterns that are known from previous literature. V:/C ratios are lower for women than men, and there seems to be a correlation between educational level and V:/C ratio. Young speakers in these two cities behave contrary to the expectations. In Turku, especially the young men have very high V:/C ratios, as high as speakers from Vaasa. In Helsinki, the young men have lower V:/C ratios than the women.

The pronunciation of the young men from Turku is puzzling, since they behave contrary to what Schaeffler's (2005) theory predicts. If the variation between older and younger speakers in Turku is seen as apparent-time change, the system is changing from a more economical system (where vowel duration only is used for contrasting V:C and VC:) towards a durationally over-specified system in Schaeffler's terms, which would be a typologically unexpected change.

On the other hand, language change in modern western societies is mostly led by young women, not by young men (Chambers & Trudgill 1998: 61). It has also been argued that women lead language change not only by adapting to standard norms, but by actually creating prestige (Milroy, Milroy & Hartley 1994). *If* one would assume that young women are showing the most recent direction of language change, then one could assume that the pronunciation with highest prestige in the early 21st century is a rather "neutral" one. Young women in all three cities have less extreme V:/C ratios than other speakers in the same cities. Admittedly, the data set is quite small for drawing conclusions about gender-related variation. The hypothesis that a pronunciation with intermediate V:/C ratios is gaining status would, still, be supported by the impression of the speaker cited in the Section 2 and by Stenberg-Sirén & Östman (2012) and Stenius (2012) who have noted decreasing use in public media of the extremely low V:/C ratios typically associated with Helsinki.

---

ration of C in VC is shorter than C in V:C. The main cue might, hence, be total sequence duration, maybe in combination with relative segment durations.

Quantity is and has been a feature that Finland-Swedes are sensitive to and have strong social associations with. The sensitivity to durational differences might, as suggested by Höckerstedt (2010: 61) have been enhanced by contact with the quantity language Finnish. Since extremely low as well as extremely high V:/C ratios have strong associations ("posh" resp. "rural"), an intermediate "neutral" pronunciation free from any of these associations might be developing into the most desired pronunciation. This pronunciation might also be closer to the Central Standard Swedish one.

For a better understanding of the social value of quantity in Finland Swedish it would be interesting to combine the production data with attitudinal data. The linguistic behaviour of young speakers from Turku is especially interesting, since a typologically unexpected change seems to have taken place. Qualitative, attitudinal and demographic data might provide some explanations to this unexpected change. To get a complete picture of the durational correlates of quantity, voiced consonants should be studied, too, as well as segment durations in VC sequences. From a language contact point of view it would be valuable to contrast the results with comparable measurements from Finnish and Central Standard Swedish.

Bibliography

Ahlbäck, O. (1971). *Svenskan i Finland.* Skrifter utgivna av Nämnden för svensk språkvård 15, 2nd edition. Stockholm: Läromedelsförlagen.

Baayen, R.H. (2008). *Analyzing linguistic data – A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bates, D., Maechler, M. & Bolker, B. (2011*). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-41.* http://CRAN.R-project.org/package=lme4.

Behne, D., Czigler, P. & Sullivan, K. (1996). Acoustic characteristics of perceived quantity and quality in Swedish vowels. In P. McCormack & A. Russell (eds.). *SST 1996 Proceedings*. Canberra: Australasian Speech Science and Technology Association, 49–54.

Boersma, P. & Weenink, D. (2010). *Praat: doing phonetics by computer.* [phonetic software, version 5.3]. http://www.praat.org.

Bruce, G. (1994). Prosodisk strukturering i dialog. In A. Holmberg & K. Larsson (eds.). *Svenskans beskrivning*, 20, 9–23. Lund.

Bruce, G. (2012). *Allmän och svensk prosodi.* Lund: Studentlitteratur.

Chambers, J.K. & Trudgill, P. (1998). *Dialectology*. Cambridge: Cambridge University Press, 2nd edn.

Drager, K. & Hay, J. (2012). Exploiting random intercepts: Two case studies in sociophonetics. *Language Variation and Change*, 24, 59–78.

Elert, C.-C. (1964). *Phonologic studies of quantity in Swedish: Based on material from Stockholm speakers.* Uppsala: Almqvist & Wiksell.

Engstrand, O. & Krull, D. (1994). Durational correlates of quantity in Swedish, Finnish and Estonian: Cross-language evidence for a theory of adaptive dispersion. *Phonetica*, 50, 80–91.

FO = *Ordbok över Finlands svenska folkmål* (1976–). Helsingfors: Forskningscentralen för de inhemska språken/Svenska litteratursällskapet i Finland.

Hadding-Koch, K. & Abramson, A. (1964). Duration versus spectrum in Swedish vowels: Some perceptual experiments. *Studia Linguistica*, 18, 94–107.

Hakulinen, L. (1979). *Suomen kielen rakenne ja kehitys*, 4th edition. Helsinki: Otava.

Hällström-Reijonen, C. af (2010). Ledare. *Språkbruk*, 3, 3.

Heldner, M. & Strangert, E. (2001). Temporal effects of focus in Swedish. *Journal of Phonetics*, 29, 329–361.

Helgason, P., Ringen, C. & Suomi, K. (2010). Quantity in Central Standard Swedish and Fenno-Swedish. In S. Schötz & G. Ambrazaitis (eds.). *Proceedings from Fonetik 2010*, Working Papers 54. Lund: Centre for Languages and Literature, Lund University, 47–50.

Höckerstedt, L. (2010). Svenskan på plats! Helsingforssvenskan visar vägen – födelse, liv och död. Helsingfors: Söderströms, epub edn.

Itkonen, T. (1965). Über einige Sandhi-Erscheinungen im Finnmarklappischen, Finnischen und Finnlandschwedischen. *Finnisch-Ugrische Forschungen*, 35, 218–263.

Ivars, A.-M. (1988). Närpesdialekten på 1980-talet. *Studier i nordisk filologi*, 70. Helsingfors: Svenska litteratursällskapet i Finland.

Ivars, A.-M. (2003). Lokalt och regionalt i svenskan i Finland – Tendenser i språkutvecklingen i norr och söder. In G. Akselberg, A.M. Bødal & H. Sandøy (eds.). *Nordisk dialektologi*. Oslo: Novus, 51-81.

Ivars, A.-M. & Södergård, L. (2007). Spara det finlandssvenska talet. In T. Arboe (ed.). *Nordisk dialektologi og sociolingvistik*. Århus: Peter Skautrup Centret for Jysk Dialektforskning, 202–206.

Ivars, A.-M. & Södergård, L. (2009). *Spara det finlandssvenska talet – Slutrapport*. SLS' elektronic publication archive in the Doria repository. http://urn.fi/URN:NBN:fife201004191691.

Kim, Y. (2006). Variation and Finnish influence in Finland Swedish dialect intonation. *Lund Working Papers in Linguistics*, 52, 77–80.

Kiparsky, P. (2008). Fenno-Swedish quantity: Contrast in stratal OT. In B. Vaux & A. Nevins (eds.). *Rules, Constraints, and Phonological Phenomena.* Oxford: Oxford University Press, 185–220.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, Massachusetts: MIT Press.

Lehtonen, J. (1970). *Aspects of quantity in standard Finnish.* Studia philologica Jyväskyläensia 6. Ph.D. thesis. Jyväskylä: University of Jyväskylä.

Milroy, J., Milroy, L. & Hartley, S. (1994). Local and supra-local change in British English: the case of glottalisation. *English world-wide*, 15, 1–33.

Östman, J.-O. & Mattfolk, L. (2011). Ideologies of standardisation: Finland Swedish and Swedish-language Finland. In T. Kristiansen & N. Coupland (eds.). *Standard languages and language standards in a changing Europe.* Oslo: Novus, 75–82.

Pind, J. (1999). Speech segment durations and quantity in Icelandic. *Journal of the Acoustical Society of America*, 106, 1045–1053.

R Development Core Team (2011). *R: A language and environment for statistical computing.* Wien: R Foundation for Statistical Computing. http://www.R-project.org.

Reuter, M. (1971). Vokalerna i finlandssvenskan: en instrumentell analys och ett försök till systematisering enligt särdrag. In L. Huldén & C.-E. Thors (eds.). *Festskrift till Olav Ahlbäck*. Studier i nordisk filologi 58. Helsingfors: Svenska Litteratursällskapet i Finland, 240–249.

Reuter, M. (1980). Kortstavighet i Helsingforssvenskan. In O. Ahlbäck, L. Huldén & K. Zilliacus (eds.). *Festskrift till Carl-Eric Thors*. Studier i nordisk filologi 62. Helsingfors: Svenska Litteratursällskapet i Finland, 222–233.

Reuter, M. (1982). *Kvantitet i Helsingforssvenskan.* Licentiate thesis. Helsinki: Helsinki University.

Reuter, M. (1992). Swedish as a pluricentric language. In M.G. Clyne (ed.). *Pluricentric languages: differing norms in different nations*. Contributions to the sociology of language, 62. Berlin: De Gruyter, 101–116.

Riad, T. (1992). *Structures in Germanic prosody: A diachronic study with special reference to the Nordic languages.* Ph.D. thesis. Stockholm: Stockholm University.

Rietveld, A.C.M. & Van Heuven, V.J. (2009). *Algemene fonetiek,* 3rd edition. Bussum: Coutinho.

Schaeffler, F. (2005). *Phonological quantity in Swedish dialects – Typological aspects, phonetic variation and diachronic change.* Phonum 10. Ph.D. thesis. Umeå: Umeå University.

Stenberg-Sirén, J. & Östman, J.-O. (2012). Processes of standardization through the media in minority languages: Swedish in Finland. In M. Hüning & U. Reich (eds.). *Sociolinguistic Symposium 19: Abstracts*, 183. Berlin: Freie Universität Berlin.

Stenius, Y. (2012). Språk och klass. *Åbo underrättelser*. 5 June 2012.

Strangert, E. & Heldner, M. (1995). Labelling of boundaries and prominences by phonetically experienced and nonexperienced transcribers. *Phonum,* 3, 85–109.

Suomi, K., Toivanen, J. & Ylitalo, R. (2008). Finnish Sound Structure: Phonetics, phonology, phonotactics and prosody. *Studia humaniora ouluensia*, 9. Oulu: University of Oulu.