University of Groningen

An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects

Leinonen, Therese

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Leinonen, T. (2010). *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects.* s.n.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license. More information can be found on the University of Groningen website: https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment.

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 25-09-2023

Chapter 5

Acoustic measures of vowel quality

As described in § 2.4.3 acoustic measures of vowel pronunciation are generally influenced by the anatomy/physiology of the speaker. The largest differences related to anatomy/physiology can be found between men and women and children, but also within these three groups speaker dependent variation is found. This is a problem in dialectological and sociolinguistic research, since researchers are mainly interested in the socially and geographically conditioned variation and would like to disregard variation related to anatomy.

As mentioned in § 2.4.4 a large number of methods has been developed that attempt to normalize for the speaker-specific variation in formant measurements. However, these normalization procedures are successful only to some extent. The main finding is that normalization procedures can and should be applied only to data sets that are fully phonologically comparable, that is, have the same mean value and standard deviation (Disner, 1980; Adank, 2003). This means that normalization procedures can only be used when all speakers have the same phoneme system or at least share the same stable point vowels.

In some variationist studies the problem of speaker variability is solved by averaging over a number of speakers for each variety. This can be done if the groups are large enough and if the share of men and women is equal across all groups. If not, anatomical/physiological differences will bias the results.

In the Swedish language area there are dialects with very deviant vowel systems (see § 2.3.3). Moreover, all the vowels in the data set for the current thesis were not recorded by all speakers (see § 4.3). This makes the use of standard normalization procedures impossible. Because of the differing number of men and women per variety and per vowel, averaging over all speakers per variety would also not normalize for the variation related to speaker anatomy/physiology.

Principal component analysis (PCA) of band-pass filtered vowel spectra has been shown by Jacobi (2009) to be a suitable method for large-scale language variation studies (see § 2.4.2). One advantage of this method, introduced by Plomp, Pols, & Van de Geer (1967), is that it can be fully automated without leading to errors of the kind made by formant tracking algorithms. Because the method does not require manual correction of the results it can easily be applied to large amounts of data. Still, this method is also sensitive to speaker variability and to some extent to the amount of noise in the recordings (Jacobi, 2009, 59–63). Jacobi (2009, 63–66) related the vowel pronunciations of each speaker to his or her point vowels and could, thus, reduce the speaker and recording specific variation. This could be done because the Dutch point vowels—/a/, /i/, /u/—are assumed to be stable across the whole language area.

PCA of band-pass filtered vowel spectra was chosen as a measure of vowel quality for the present thesis. Because the data comprises nearly 1,200 speakers it was essential to choose a method which can be automated to a higher extent than formant measurements. Since the Swedish point vowels are not stable across all dialects, Swedish does not offer the opportunity to use point vowels to reduce speaker-dependent variation. However, using PCA on Bark-filtered vowel spectra offers an opportunity to eliminate the largest source of speaker dependent variation: the one caused by anatomical/physiological differences between men and women. The method is described in § 5.1. In § 5.2 principal components of Bark-filtered spectra are compared with formants, and the principal components are interpreted in relation to formant frequencies.

5.1 Principal component analysis of Bark-filtered vowel spectra

The method chosen for assessing vowel quality for this thesis comprises two steps: Bark filtering and principal component analysis (PCA). The method is described in the following sections, and in § 5.1.7 a short summary of all the steps of the analysis is given.

5.1.1 Bark filtering

Using the Praat[1] software vowel spectra were filtered with Bark[2] filters up to 18 Bark with a window length of 13 ms. Each pass band had a bandwidth of one Bark and adjacent filters overlapped at −3 dB (Jacobi, 2009, Fig. 3.3). Following Jacobi (2009, 55) 18 Bark was chosen as the highest frequency. The frequency range up to ca. 18 Bark is where the first three formants of vowels are found. Higher frequencies are not used by listeners for identifying vowels but mainly show speaker-specific variation. Table 5.1 shows the mid-frequencies in Hertz of the 18 Bark filters.

---

[1] Praat: phonetic software. Version 5.1. By P. Boersma and D. Weenink, University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>

[2] A number of different algorithms have been proposed for modeling the Bark scale. Praat uses the conversion formula $7 \times \ln(Hz/650 + \sqrt{1 + (Hz/650)^2})$. See also the description of the Bark scale in § 2.4.2 (p. 28).

Table 5.1. Mid-frequencies of the 18 Bark filters in Hertz

Bark	Hz	Bark	Hz	Bark	Hz
1	93	7	764	13	2031
2	188	8	915	14	2357
3	287	9	1086	15	2732
4	392	10	1278	16	3163
5	505	11	1497	17	3657
6	628	12	1746	18	4228

Measurements were made at nine points in time within every vowel segment starting at 25% of the total vowel duration and ending at 75% of the vowel duration, that is, at $\frac{4}{16}$, $\frac{5}{16}$, $\frac{6}{16}$, $\frac{7}{16}$, $\frac{8}{16}$ (=center), $\frac{9}{16}$, $\frac{10}{16}$, $\frac{11}{16}$ and $\frac{12}{16}$.

The Bark-filtered spectra were level-normalized. Normalization was done for every 13 ms sample so that the levels add up to 80 dB.

5.1.2 Principal component analysis

The filter bank representation of vowels described in the previous section can be reduced to articulatory meaningful components by means of principal component analysis (PCA) (Jacobi, 2009, 42). A PCA of the Bark-filtered vowel spectra was carried out with the statistical software package SPSS[3].

PCA is a data reduction technique that aims at reducing a larger number of variables into a smaller set of components. It enables the researcher to identify which variables in a data set show similar patterns of variation and whether the variables can be divided into relatively independent subsets. Based on a variance-covariance matrix or a standardized correlation matrix of the observed variables, variables that correlate with each other are combined into components, so that the total amount of data can be reduced. The first principal component (PC) explains as much as possible of the total variance in the data set, the second PC as much as possible of the variance still left, etc.

The analysis produces a set of *loadings* and a set of *scores* for each extracted component. Loadings can be interpreted as correlations between the original variables and the components and can be used to calculate scores for each object based on the original variables. The scores can be interpreted as such, or they can be used as input to further analyses replacing the larger number of original variables for each object and thus reducing the data set. Thorough descriptions of PCA can be found in statistical handbooks, for example Field (2005) or Tabachnik & Fidell (2007).

5.1.3 Computing loadings based on point vowels

The PCA of the acoustic data in this thesis was based on a variance-covariance matrix. Following Jacobi (2009) the loadings of the PCA were calculated using only

---

[3] SPSS version 16.0 for Windows. SPSS Inc.

Table 5.2. Sample of the four point vowels (measured at the temporal mid-point) of a number of speakers used in the analysis phase of the PCA for computing loadings. bf = Bark filter

objects		*variables*			
speaker	vowel	bf$_2$ (dB)	bf$_3$ (dB)	...	bf$_n$ (dB)
speaker$_1$	a	58.60	61.10	...	46.37
	æ	65.57	65.25	...	64.85
	i	67.80	72.10	...	68.49
	u	63.09	67.44	...	56.73
speaker$_2$	a	61.71	62.41	...	41.85
	æ	71.94	73.89	...	58.98
	i	69.23	72.27	...	56.01
	u	67.96	70.28	...	37.85
...	...	...	...	...	...
speaker$_n$	a	59.97	61.52	...	52.51
	æ	54.37	53.34	...	57.02
	i	65.15	68.73	...	59.00
	u	65.16	66.34	...	38.82

Table 5.3. Sample of data to be reduced by the PCA. Each of the 19 vowels of every speaker is represented by the intensities (in dB) in a number of Bark filters (bf) measured at nine sampling points within the vowel segments.

objects		*variables*			
speaker	vowel$_{point}$	bf$_2$ (dB)	bf$_3$ (dB)	... ...	bf$_n$ (dB)
speaker$_1$	dis_1	66.69	71.96	...	64.31
	dis_2	66.79	71.82	...	63.48
	$dis_{...}$	...	...	...	...
	dis_9	72.62	75.57	...	68.85
	$disk_1$	73.24	75.72	...	64.99
	$disk_{...}$	...	...	...	...
	$disk_9$	72.87	75.17	...	65.57
	...	...	...	...	...
	typ_1	67.26	73.04	...	63.85
	$typ_{...}$	...	...	...	...
	typ_9	74.03	75.45	...	62.56
...	...	...	...	...	...
speaker$_n$	typ_1	72.71	75.28	...	52.36
	$typ_{...}$	...	...	...	...
	typ_9	75.51	75.83	...	50.67

point vowels (that is, vowels with the most extreme values for the two first formants). This is done in order to weigh all articulatory dimensions equally and to make sure that the PCA accounts for all possible variation in the vowel space. Table 5.2 shows a sample of the data used as input in this initial phase of the PCA. The point vowels of the speakers serve as objects in the analysis, and the intensities of the Bark filters as variables.

For the Swedish data the vowels transcribed as [iː], [æː], [ɑː]/[aː] and [uː] in the database were chosen for computing the loadings in the initial phase of the PCA. Two variants were allowed for /ɑː/, since the pronunciation in Standard Swedish is [ɑː], but we did not want to exclude that even more open vowels were added to the analysis. Standard Swedish [æː] has a more extreme F1 than Swedish [ɑː] (see Table 5.11, p. 83) and was therefore also used as a point vowel.

Some dialects show such strong deviation from Standard Swedish that not all point vowels could be found in the set of 19 words. For example, the South Swedish dialects have strongly diphthongized pronunciations of the long close vowels /iː/ and /uː/. Because point vowels were not available for all of the speakers, a subset of speakers was used for calculating the loadings. All point vowels were available for 230 women. The number of men with all point vowels was a bit larger, so out of these speakers 230 men were picked randomly, in order to include as many men as women in the analysis. Using 230 men and 230 women for computing the loadings of the PCA, means that a great number vowel spaces with different speaker-dependent sizes are included. Based on this, scores can be calculated also for the vowels of the speakers that were not included in the initial subset. Only the central measurement point was used for calculating the loadings. For each of the 230 men and 230 women average levels of the Bark filters of all occurrences of the point vowels were calculated.[4]

Figure 5.1 shows the mean intensities per Bark filter for the point vowels of the 230 men and 230 women. In these figures the characteristic spectra of the point vowels can be identified. The [iː] vowel has a very low first formant resulting in an intensity peak in the lowest frequency area, approximately at Bark filters 2 and 3. The second formant of [iː] is very high, at approximately 13–15 Bark. The open vowel [æː] has a high F1 at 6–7 Bark. The spectral peaks resulting from F2 and F3 of [æː] can also be seen clearly at 11–12 respectively 14–15 Bark. F1 and F2 of the vowel [ɑː]/[aː] are very close to each other giving a broad peak at 5–9 Bark, while the first two formants of [uː] are very close to each other as well, but at a much lower frequency (3–7 Bark). The lines of men and women are very similar in these graphs but the peaks of the spectra are consistently at a somewhat higher frequency for women than men. This can be illustrated by shifting the frequency scale with 1 Bark, which is done in Figure 5.2. In these figures the lines of men and women follow each other almost perfectly.

The biggest difference between the spectra of men and women in Figure 5.2 seems to be that the women in general have less intensity at the highest frequencies,

---

[4] $L_{average} = 10log_{10}[(\frac{1}{n})\sum_{i=1}^{n} log^{-1}(L_i/10)]$

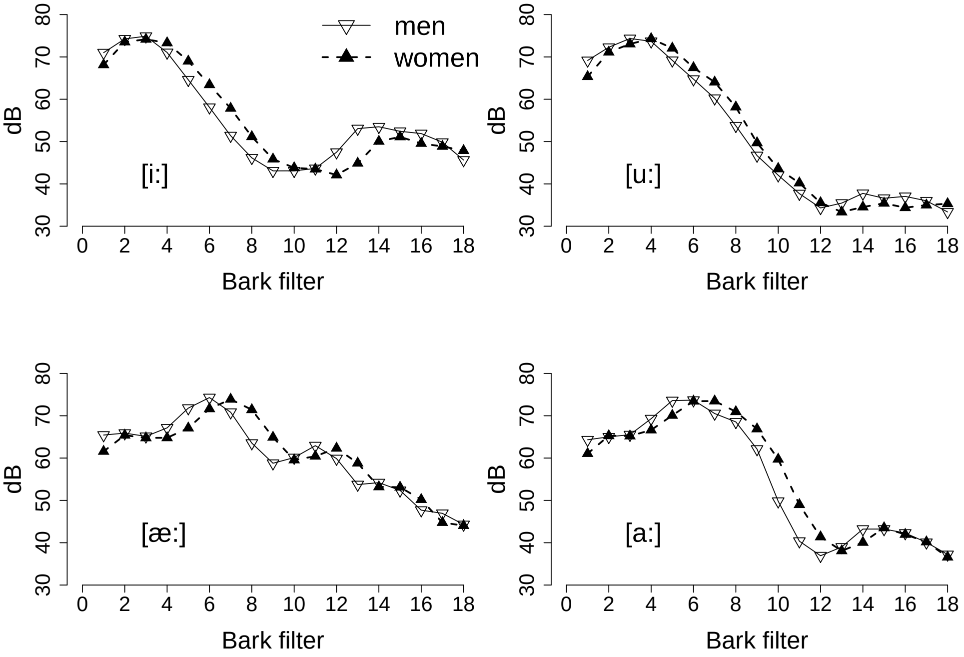


Figure 5.1. Average decibel levels of the point vowels in the band-pass filtered frequency regions.

which can be most clearly seen in the case of [iː]. The vocal folds of men and women produce a different kind of pulse. The duration of the open portion of a fundamental period is relatively longer in women's voices, and due to the relatively longer pulse by women than by men the higher harmonics are weaker in women's speech (Rietveld & Van Heuven, 2009, 341). This leads to a steeper spectral tilt for women than for men. Measured with four rather broad frequency bands Sluijter & Van Heuven (1996) found that at 0–500 Hz female voice have 1 dB greater intensity than male voices, while at 500–4000 Hz the intensity is 2–3 dB weaker in female voices than in male voices. The greater intensity of men than of women at the highest frequencies in Figure 5.2 is therefore attested also in previous studies.

Van Nierop, Pols, & Plomp (1973) showed that the systematic differences in the spectra of vowels produced by men and women lead to similar results when carrying out separate PCAs for the two groups. The main latent variables related to vowel pronunciation are present in the acoustic data of both men and women, but the information is found at a higher frequency in the female data. This fact was used in the current thesis for normalizing for the systematic differences in the acoustic data of men and women related to anatomy/physiology: separate PCAs were carried out on the male data and the female data. The scores of the two separate PCAs can subsequently be used as comparable measures of vowel quality and can be used

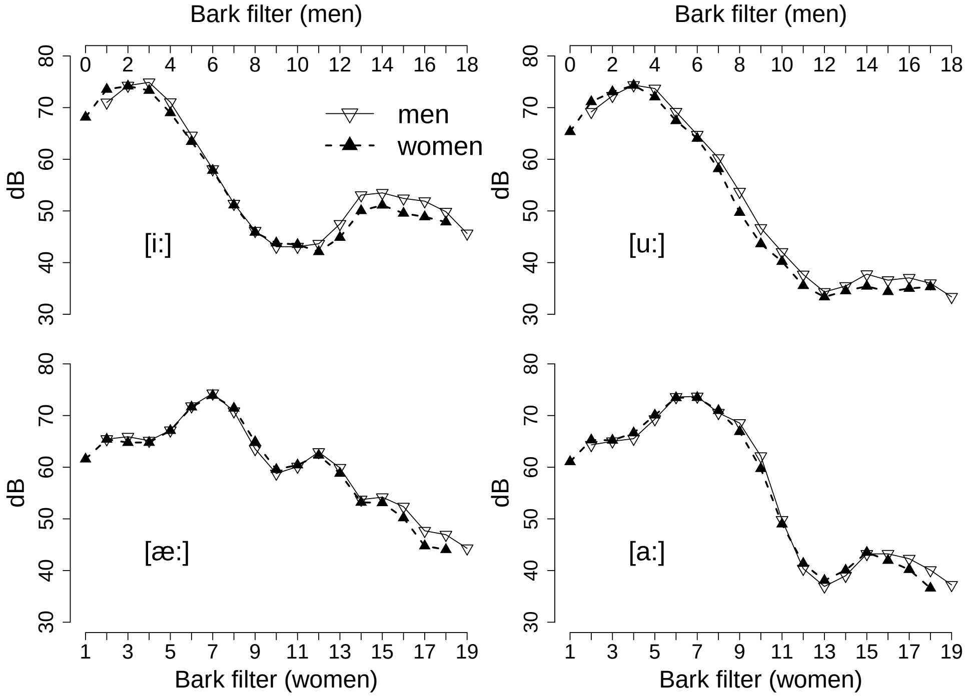


Figure 5.2. Average decibel levels of the point vowels in the band-pass filtered frequency regions. Scales of men and women shifted with 1 Bark.

in analyses of dialectal variation. Bark filters 2–17 were used as variables in the PCA of the male data and Bark filters 3–18 in the female analysis. Using a set of Bark filters shifted with 1 Bark as variables in the two analyses means that the frequency area analyzed in the two PCAs contains the same information related to vowel pronunciation. Both analyses comprised 920 objects in the initial phase (4 vowels × 230 speakers). In § 5.1.5 the effect of analyzing the female and male data separately is examined more closely.

Jacobi (2009) carried out one PCA, which included both female and male speakers. She combined the two lowest Bark filters (1 and 2) in order to “prevent strong variance caused by the speakers’ varying fundamental frequency”. The mean fundamental frequency of Swedish speakers has been shown to be 188 Hz for women and 116 Hz for men (Pegoraro-Krook, 1988). Thus, the fundamental frequency is mainly represented in the first Bark filter for men and in the second for women. These Bark filters are not important for identifying vowels. When analyzing data from men and women separately the lowest Bark filters which represent the fundamental frequency can be left out.

After computing the loadings of a PCA with a subset of objects, scores can be computed for the full data set. Also the second phase of the PCA was carried out separately for men and women. Table 5.3 shows a sample of the full data set used

Table 5.4. Result of data reduction by PCA. The original variables (that is, the intensities in a number of Bark filters) have been reduced to scores on two PCs. The objects of the analysis are the 19 vowels of all speakers measured at nine sampling points.

speaker	vowel_{point}	PC1	PC2
speaker_1	dis_1	-0.99	1.75
	dis_2	-1.20	1.61
	$dis_{...}$	...	...
	dis_9	-2.15	1.62
	$disk_1$	-1.56	1.76
	$disk_{...}$	...	...
	$disk_9$	-1.83	1.58
	...	...	...
	typ_1	-1.02	1.79
	$typ_{...}$	...	...
	typ_9	-1.82	1.70
...	...	...	...
speaker_n	typ_1	-102	0.98
	$typ_{...}$	...	...
	typ_9	-1.31	0.96

as input in the second phase of the PCA. The objects of the full data set are the 19 vowels (see § 4.2.1) of all (male/female) speakers measured at nine sampling points. Based on the loadings from the initial phase of each PCA scores for all vowels of all male/female speakers were computed. Table 5.4 displays a sample of the scores which are the output of the PCA. In the data set that has been reduced by means of PCA each object is described by scores on two extracted components (PCs) instead of by the original variables. These scores are the measures of vowel quality.

5.1.4 Rotating the solution

Because the first PC explains as much as possible of the total variance in the data set, most variables will have relatively high loadings on the first PC and smaller loadings on the remaining components. This can make interpretation of components difficult, since the original variables are not necessarily unambiguously connected to only one of the extracted components. This characteristic can be changed by using rotation techniques.

The most commonly used rotation technique is varimax, which rotates the axes so that the variables correlate maximally with only one component. Figure 5.3 shows the loadings of the two first PCs based on the male and female point vowels in unrotated solutions. The plots show clouds of variables that are not centered along any of the axes, but the variables correlate with both the first and the second PC. These are typical cases where rotation could lead to more easily interpretable solutions.

Figure 5.4 shows the loadings of PCAs of the same data with varimax rotation. Contrary to Figure 5.3, the clouds of variables are now centered along the x- and y-axes and only a few variables correlate highly with both PCs. Interpretation of

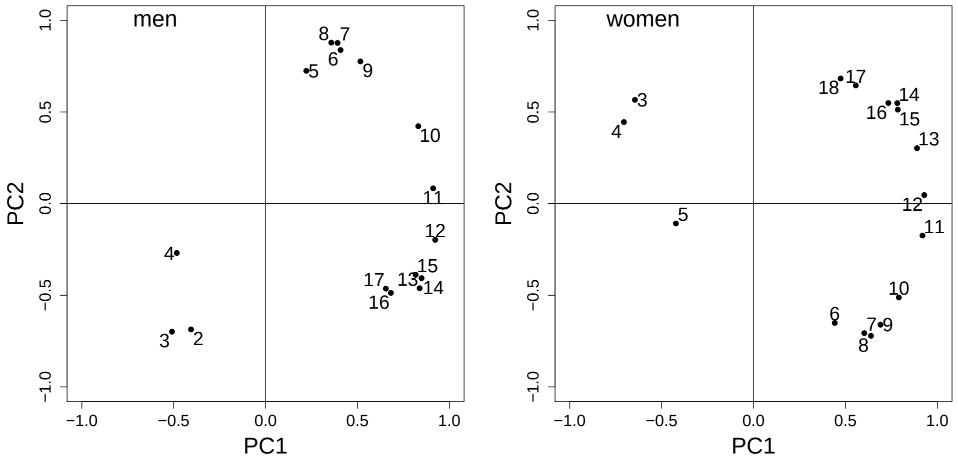


Figure 5.3. Loadings on the two first PCs (men to the left, women to the right). The numbers indicate the number of the Bark filters (men 2–17 Bark, women 3–18 Bark).

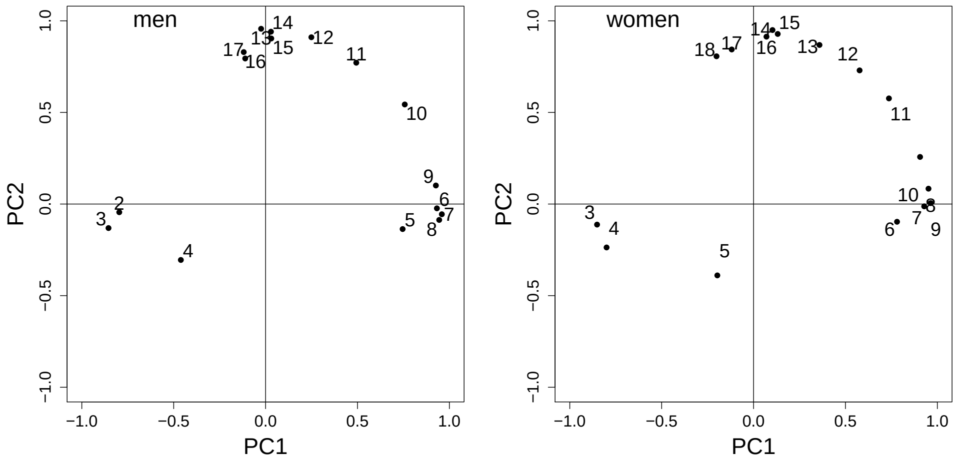


Figure 5.4. Loadings on two PCs extracted with varimax rotation (men to the left, women to the right). The numbers indicate the number of the Bark filters (men 2–17 Bark, women 3–18 Bark).

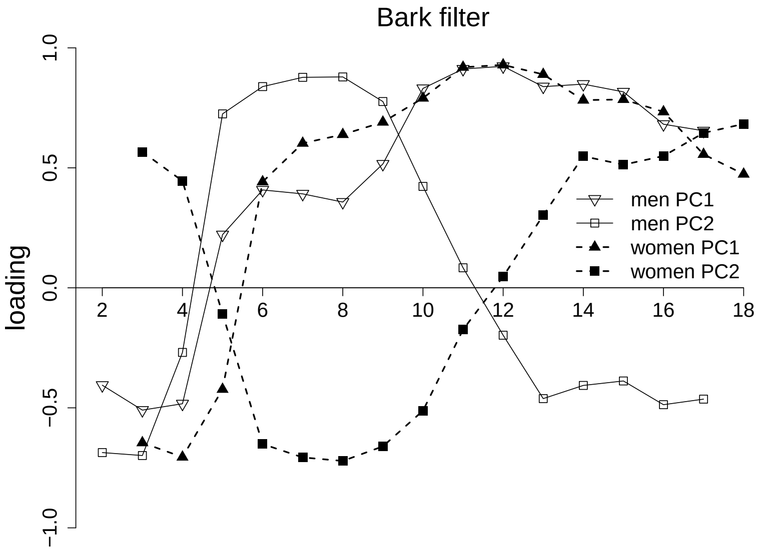


Figure 5.5. Loadings of the male and female PCAs (2 PCs extracted without rotation) plotted against the frequency scale.

the components becomes easier because it is evident which variables correlate with each of the PCs.[5]

Varimax also tends to equalize the proportion of variance explained by the components, by taking variance from the first component and distributing it among the later ones (Tabachnik & Fidell, 2007, 638). Because of this all components will be affected by the number of components extracted when using varimax.

One further difference between the two pairs of plots in Figures 5.3 and 5.4 is that the configurations of the male and female analyses are mirrored around the x-axis in the unrotated solutions in Figure 5.3; the highest band-pass filters are found in the fourth quadrant in the male solution but in the first quadrant in the female solution, while the middle frequencies (Bark filters 6–10) are in the first quadrant in the male solution and in the fourth in the female solution. After applying varimax rotation the solutions are much more similar for men and women (Figure 5.4). Van Nierop et al. (1973) found that solutions based on male data and female data are comparable and only need to be rotated in order to overlap each other. Varimax seems to offer a standard solution for rotation so that the configurations of male and female data have the same orientation. Therefore, varimax was chosen as rotation technique for extracting the PCs for this thesis.

The difference between non-rotated solutions and solutions using varimax can be seen even more clearly when plotting the loadings along the frequency scale. Figure 5.5 shows the loadings of the two first PCs of the male and female data extracted without rotation, while Figure 5.6 shows the loadings of the varimax solutions. While

---

[5] For an interpretation of the PCs, see § 5.2.3. See also § 5.2.1, where both the unrotated and the rotated solutions are compared to formant measurements.

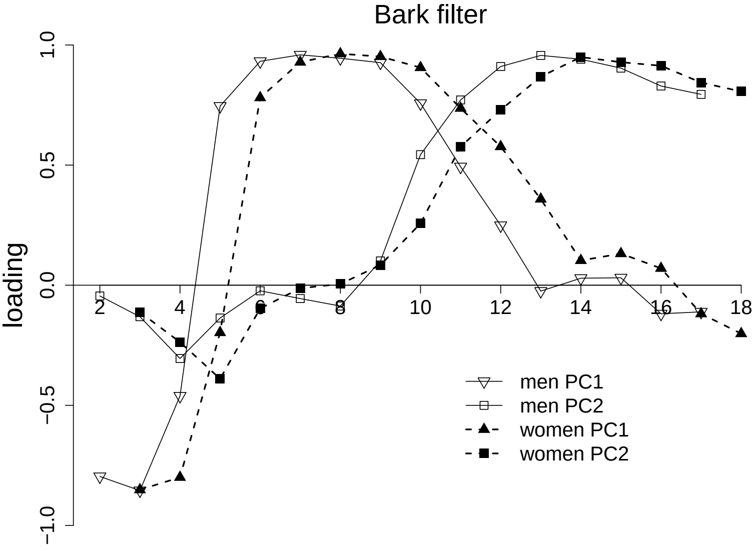


Figure 5.6. Loadings of the male and female PCAs (2 PCs extracted with varimax rotation) plotted against the frequency scale.

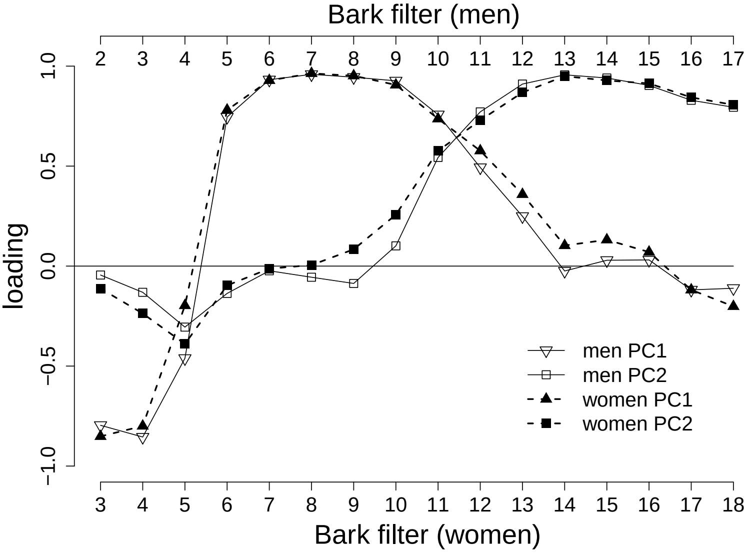


Figure 5.7. Loadings of the male and female PCAs (2 PCs extracted with varimax rotation) plotted against a shifted frequency scale.

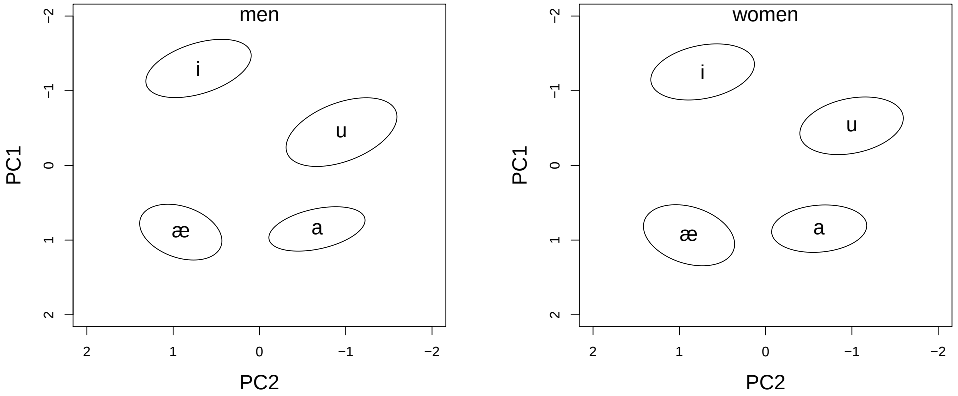


Figure 5.8. Scores of the point vowels [iː], [æː], [ɑː]/[aː] and [uː] in the PC2/PC1 plane with varimax rotation. The plots of the male data (left) and female data (right) are similar to each other and resemble the IPA vowel quadrilateral. One standard deviation ellipses.

Figure 5.5 displays quite different curves for men and women, the curves of the components in Figure 5.6 look very similar for both sexes. In the varimax solutions the loadings of men and women only seem to be placed differently on the frequency scale.

In Figure 5.7 the frequency scale has been shifted so that instead of using the same scale for men and women, the loadings of the male analysis are plotted on a scale ranging from 2 to 17 Bark, while the loadings of the female analysis are plotted on a scale from 3 to 18 Bark (which corresponds to the contiguous Bark filters used as variables in the two analyses). This figure shows that the curves are indeed almost identical for men and women after applying varimax, which suggests that the same information is extracted in both PCAs. Because of the anatomical/physiological differences between men and women this information can be found on average 1 Bark higher in the female data.

The effect of applying varimax can also be visualized by plotting the scores assigned to the point vowels by the PCA. The scores are the result of the data reduction, and can be used as measures of vowel quality for each segment. The scores were estimated with the regression method, which produces scores with a mean of 0 and a standard deviation of 1 for each PC (Tabachnik & Fidell, 2007, 650). Figure 5.8 shows the scores of the male and female point vowels in the varimax solution.[6] Just

---

[6] Ellipses are drawn by applying PCA once more, but this time separately for each vowel with the acoustic PCs as variables. The major and minor axes of the ellipses are the two first PCs of the data and the longest axis, hence, shows the direction that explains most of the variance (Harrington, 2010, Ch. 6). Assuming that the data is normally distributed an ellipse with a radius of 1 standard deviation covers 39.4% of the data points, while an ellipse with a radius of 2 standard deviations covers 86.5% of the data points. All ellipse plots in this thesis were drawn using the

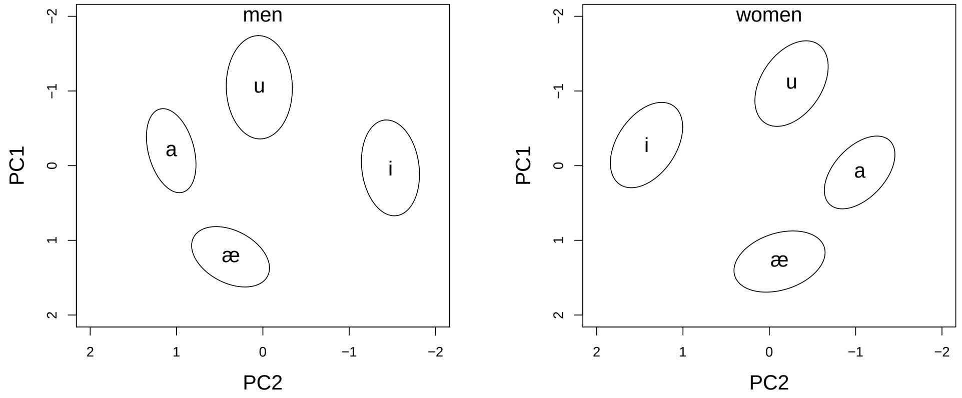


Figure 5.9. Scores of the point vowels [iː], [æː], [ɑː]/[aː] and [uː] in the PC2/PC1 plane without applying any rotation technique. The vowel plots of the male data (left) and female data (right) show a skewed position compared to the plots in Figure 5.8 and are mirrored around the y-axis. One standard deviation ellipses.

as in a formant plot (see, for example, Figure 2.4, p. 25, and Figure 2.5, p. 30), the scores are plotted with the first component on the y-axis and the second component on the x-axis with both scales reversed, which results in configurations that resemble the IPA vowel quadrilateral. Figure 5.9, on the other hand, displays the scores of the point vowels of the unrotated PCAs. These plots do not show the familiar vowel quadrilateral with backness along the x-axis and height along the y-axis, but the vowel spaces have a rotated position in the coordinate system. Moreover, the plots of the male and female point vowels are mirrored around the y-axis.

5.1.5 The result of separate PCAs of men and women

The result of analyzing the vowel pronunciations by men and women separately can be examined by comparing the vowel scores. The plot to the left in Figure 5.10 shows the scores of the point vowels after running separate PCAs for men and women with varimax rotation. The ellipses, which have a radius of one standard deviation, fit each other almost perfectly.[6]

For comparison the plot to the right in Figure 5.10 shows the vowel scores of a PCA where men and women were included in the same analysis. Bark filters 2–18 were analyzed for all speakers, and the analysis included 1,840 objects (4 vowels × 460 speakers). Because the sexes were analyzed together the anatomical differences between men and women led to systematic differences in the scores of the PCA. Vowels produced by women were assigned systematically higher scores on PC1 than

---

`eplot()` function of the Emu library in the software package R.

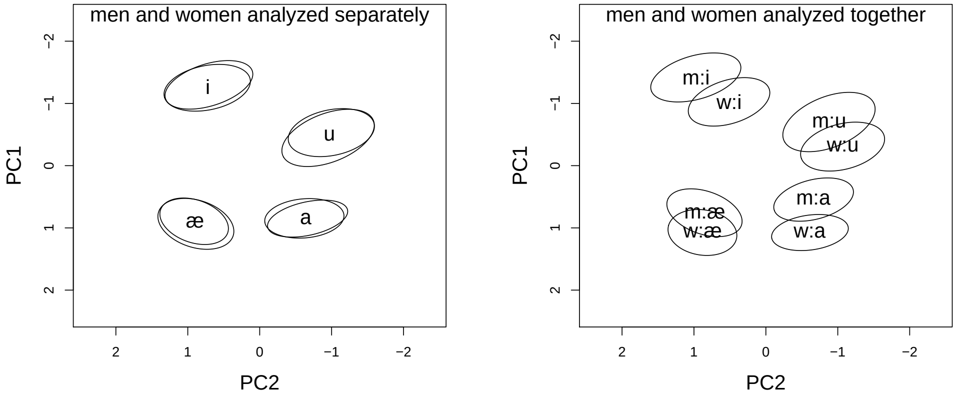


Figure 5.10. Scores of the point vowels [iː], [æː], [ɑː]/[aː] and [uː] (2 PCs extracted with varimax rotation). The plot to the left is based on separate PCAs for men and women, while the plot to the right is based on one PCA for speakers of both sexes. Separate one standard deviation ellipses are drawn for men (m) and women (w). The plot to the right fails to normalize vowel quality with respect to sex.

vowels produced by men. Separate analyses for men and women seem to normalize for the anatomical differences.

In order to test to what extent separate PCAs for men and women actually normalize for the anatomical differences, t-tests were carried out for all four point vowels. This was done both for the scores of the separate PCAs of men and women and for the PCA that included both men and women. The t-values and the significance levels of the t-tests are displayed in Table 5.5. The results show that when carrying out separate PCAs for men and women there are no significant differences between the scores of men and women except for on PC1 of [uː]. When running only one PCA where both men and women are included all vowels show significant differences on PC1, and [iː] and [uː] also on PC2. The separate analyses of men and women clearly leads to fewer differences between the sexes in the vowel scores.

Given that the separate analyses of vowels produced by men and women show less significant differences between the sexes than one single analysis, we expect that also the total variance caused by speaker-specific variation is reduced by the separate analyses. A visual impression is given by the plots in Figure 5.11, which show two standard deviation ellipses of the point vowels of the 460 speakers (230 men and 230 women) in the separate analyses and in the joined analysis. The size of the ellipses suggests that a large amount of individual variation is still left after both analyses. However, the variation is reduced to some extent by analyzing men and women separately, as can most clearly be seen in the cases of [iː] and [uː]; the ellipses of these two vowels show overlap when men and women were analyzed together (right plot in Figure 5.11) but no overlap when the PCA was carried out separately for the

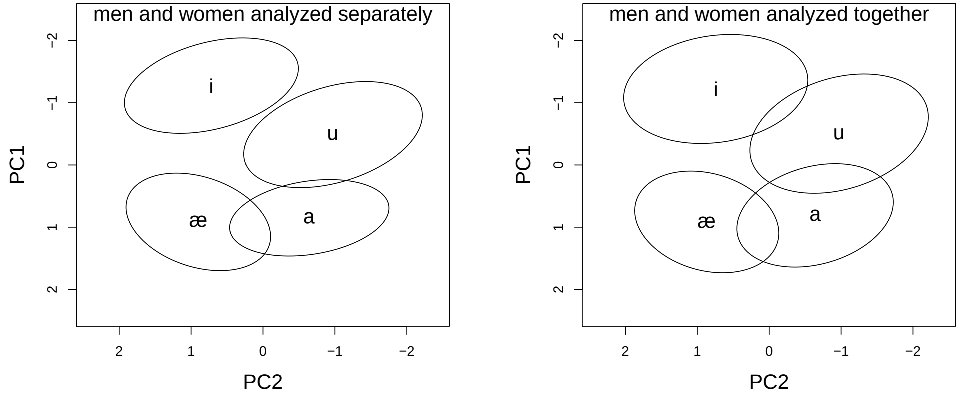


Figure 5.11. Two standard deviation ellipses of the scores of the point vowels [iː], [æː], [ɑː]/[aː] and [uː] (2 components extracted with varimax rotation). In the graph to the left, based on separate PCAs for men and women, the ellipses are smaller and show less overlap than in the graph to the right based on one PCA for speakers of both sexes. Processing male and female voices separately, hence, reduces variability.

Table 5.5. T-tests comparing the means of female and male speakers on each point vowel and on both PCs (df = 458).

		separate PCAs for men and women		one PCAs for both men and women	
vowel	value	PC1	PC2	PC1	PC2
[iː]	t	−1.3	−0.4	−10.7	8.3
	sign.	0.191	0.671	<**0.001**	<**0.001**
[æː]	t	−1.1	0.5	−9.0	−0.6
	sign.	0.254	0.608	<**0.001**	0.545
[ɑː]/[aː]	t	0.1	−0.9	−17.7	−1.0
	sign.	0.895	0.380	<**0.001**	0.318
[uː]	t	2.1	0.8	−9.7	3.3
	sign.	**0.033**	0.435	<**0.001**	**0.001**

two speaker groups (left plot in Figure 5.11).

The analyses above support the choice of carrying out separate PCAs for men and women when reducing the filter bank representation of vowels to articulatory meaningful components. In order to further investigate which sources account for the variance in the PCs multivariate analyses were carried out. The multivariate analyses of PCs as well as of formant frequencies are presented in § 5.2.2.

5.1.6 Effect of noise

Jacobi (2009, 59–63) found that the positioning of the vowel spaces of different speakers in the PC2/PC1 plane was affected by the signal-to-noise ratios in the recordings: the lower the signal-to-noise ratio (that is, the more noise), the smaller the measured vowel space in the PC plane and the higher the PC scores. Jacobi's data comprised recordings in extremely varying situations from interviews in silent environments to private conversations and broadcast recordings with music in the background.

The present data set is much more homogeneous when it comes to the recording situations than Jacobi's data. In the SweDia project attention was payed to all recordings being made in as similar circumstances as possible. The recordings were generally made in the speakers homes, and quiet rooms with as little reverberation as possible were chosen to ensure good recording quality (see § 4.1). Because no drastically varying signal-to-noise ratios were expected in the present data the effect of noise in the dialect recordings was not tested.

The speakers representing Standard Swedish, however, were not recorded in their homes, but in a studio. A studio is more silent than any home environment, which means that a higher signal-to-noise ratio could be expected for the speakers of Standard Swedish than for the dialect speakers. The speakers of Standard Swedish were not included in the initial subset used for calculating the loadings of the PCAs, but only included in the full data set for which scores were calculated. As mentioned earlier the regression technique, used for calculating the scores in the PCA, produces scores with a mean of 0 and standard deviation of 1 for each PC. To test whether the standard speakers had systematically lower scores than the dialect speakers, with a presumably lower signal-to-noise ratio, t-tests were carried out. The average scores on both PCs of the standard speakers' four point vowels ([iː], [æː], [ɑː]/[aː] and [uː]) were tested against the point vowels of the 460 dialect speakers of the initial subset. For the dialect speakers the mean of both PCs was < 0.001 as expected. For the speakers of Standard Swedish the mean of PC1 was -0.339 and the mean of PC2 -0.273. For both PCs the difference between dialect speakers and standard speakers was significant (PC1: $t(1886) = -2.30, p = 0.021$; PC2: $t(1886) = -1.86, p = 0.063$).

As found by Jacobi (2009) a lower signal-to-noise ratio seems to lead to higher PC scores. Since the PCs are not comparable across the two different recording environments, the speakers of Standard Swedish were not included in any of the statistical analyses in Chapters 6 and 7. The Standard Swedish pronunciation is only included in the maps of each vowel in Appendix C.

5.1.7 Summary of the acoustic analysis

In Figure 5.12 all the steps of the acoustic analysis of the vowel data are summarized. The acoustic analysis includes two steps which have been described above: 1) Bark filtering, and 2) data reduction of the Bark-filtered data with PCA. In a final step before analyzing dialectal variation in the vowels group averages of groups of speakers are calculated.

The Bark filtering was done at nine sampling points in each vowel segment and the Bark-filtered spectra were level-normalized to 80 dB. Since all speakers had repeated each of the 19 elicited vowels 3–5 times during the interviews, average dB levels of the Bark filters were calculated per vowel per sampling point for each speaker. After the Bark filtering and the averaging, each of the 19 vowels of each speaker is represented by the average level-normalized intensities in the Bark filters at 9 sampling points.

In the data reduction phase the data was split up according to speaker-sex. Separate PCAs were carried out for women and men with Bark filters 2–17 used as variables in the male analysis and Bark filters 3–18 in the female analysis. The PCA includes computing loadings based on point vowels, and subsequently computing scores for all of the 19 vowels. After the PCA each of the 19 vowels of each speaker is represented by two PCs at the 9 sampling points instead of by the dB levels in the frequency bands. The PCs of men and women are the output of two separate PCAs, but the scores are comparable because they represent the same latent variables in the vowel spectra.

As explained in § 4.3 three different groupings of the speakers in the data set are used in the analysis of dialectal variation in this thesis: *a)* one group per site, *b)* older and younger speakers per site, and *c)* older women, older men, younger women, and younger men per site. For the analyses in the following chapters of this thesis, arithmetic group means were calculated for PC1 and PC2 respectively for each of the 19 vowels at the nine sampling points. After this averaging the pronunciation of the 19 vowels of each speaker group is represented by two PCs at 9 sampling points.

5.2 Principal components versus formants

Since formant measurements have a much longer tradition in variationist linguistics than the use of PCs of Bark-filtered spectra, a comparison of the PCs was made with formants. This comparison was also used for finding the optimal number of PCs to extract.

Formants were measured for a subset of the SweDia data. Three sites were picked randomly: Ankarsrum, Malung and Skog. Formants were measured for all speakers from these sites. The data sets from Ankarsrum and Skog were complete with three speakers in each speaker group (older women, older men, younger women and younger men). From Malung, however, one young man was missing in the data set, which means that formants were measured for 18 women and 17 men in total. The formants were measured for the vowels in the stressed syllables of the words

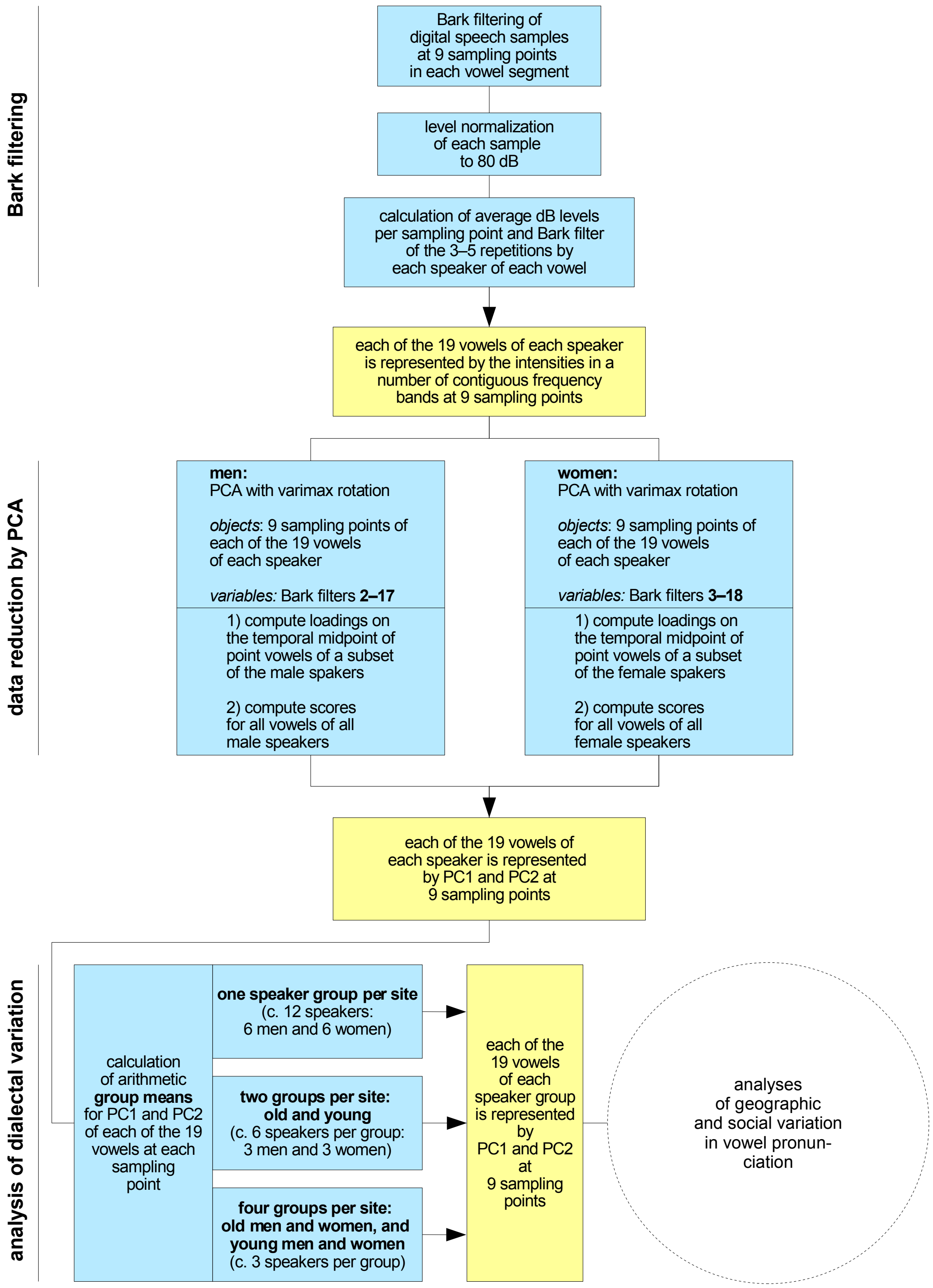


Figure 5.12. Work flow of the acoustic analysis.

described in § 4.2.1: ***dis*** /iː/, ***disk*** /ɪ/, ***dör*** /œː/, ***dörr*** /œ/, ***flytta*** /ʏ/, ***lass*** /a/, ***lat*** /ɑː/, ***leta*** /eː/, ***lett*** /e̞/, ***lott*** /ɔ/, ***lus*** /ʉː/, ***lås/låt*** /oː/, ***lär*** /æː/, ***lös*** /øː/, ***nät*** /ɛː/, ***sot*** /uː/, ***särk*** /æ/, ***söt*** /øː/, ***typ*** /yː/. In addition, the vowel /ɵ/ elicited with the word *ludd* was used. In Ankarsrum *låt* was used to elicit /oː/, while *lås* was used in Malung and Skog. One instance of every elicited vowel was measured for each of the speakers, which led to a total of 360 vowel pronunciations by women and 336 by men.[7]

The first three formants were measured in the center of each segment with the formant track function in the EMU[8] software. The window length was 25 ms and the window type Blackman. The nominal F1 value was set to 500 Hz for the male speakers and to 630 Hz for the female speakers. All measurements were inspected and errors made by the formant tracker were corrected manually.

5.2.1 Correlation with formants

Jacobi (2009, Ch. 3) found high correlations between PCs of Bark-filtered vowel spectra and formant measurements. The correlations are displayed in Table 5.6. The study was based on pronunciations of the Dutch phonemes /a/, /i/, /u/, /ɛ/ and /ɛi/ by six female and six male speakers, measured close to onset and close to offset (2,767 speech segments in total). In a second study (Jacobi, 2009, Ch. 4) including more speakers (35 female and 35 male) and more vowel tokens (12,400) and a different set of vowels (Dutch /oː/, /eː/, /ɛi/, /ɑu/ and /œy/) Jacobi found that the correlations were somewhat lower: PC1–F1 $r = 0.70$, PC2–F2 $r = 0.72$. In these studies, the formants were measured automatically without manual correction of the measurements, which means that the correlations were based on partly incorrect formant values. With corrected formant measurements we can expect to find even higher correlations. Moreover, Jacobi (2009) did not use any rotation technique to optimize the PC solution. Rotating the solution might influence the correlations between formants and components.

Table 5.6. Correlations between formants and PCs of Bark-filtered spectra found by Jacobi (2009, Table 3.2, p. 38).

	correlations			expl. var.	
PC	F1 Bark	F2 Bark	F3 Bark	%	cum. %
PC1	0.81	−0.12	0.26	65	65
PC2	−0.08	0.70	0.10	25	90
PC3	−0.19	0.05	−0.15	5	95

---

[7] Four male speakers were missing one of the target vowels, which explains the total of 336 vowels instead of the expected 340 for the male data

[8] The EMU Speech Database System. Version 2.1.1. By the Institute of Phonetics and Speech Processing, LMU Munich. <http://emu.sourceforge.net/>

Table 5.7. Correlations between formants and PCs in the unrotated solutions. Insignificant correlations ($p > 0.05$) are indicated by a hyphen.

		correlations			expl. var.	
		F1	F2	F3		cum.
sex	PC	Bark	Bark	Bark	%	%
men	PC1	0.476	0.528	−0.182	42.4	42.4
	PC2	0.696	−0.644	–	35.1	77.4
	PC3	–	−0.239	0.361	5.7	83.1
women	PC1	0.712	0.369	−0.125	49.7	49.7
	PC2	−0.506	0.731	0.239	28.4	78.2
	PC3	–	−0.234	0.399	5.2	83.3

The formant measurements of the subset of the SweDia data were correlated with the results of a number of PCA configurations in order to find the optimal PCA solution. PCA with and without varimax rotation was tested, as well as solutions with two and three extracted components. All PCAs were carried out using only point vowels in the initial phase and by analyzing data of men and women separately, as described in § 5.1. Since the PCAs were carried out separately on the vowels produced by men and women, also the correlations with formants were calculated separately for the two groups.

Because there were more front vowels than back vowels in the data, the distribution of the F2 values (and related PCs) was not completely normal. The relationship between formants and related PCs was still linear and the correlations were calculated using Pearson's correlation coefficient. All the correlations were also tested using the non-parametric Spearman's rho, which did not lead to other conclusions about the relationship between formants and PCs.

Table 5.7 shows the correlations of the three first PCs in the unrotated PC solutions with the three first formants. In both the female and the male analysis the two first PCs correlate highly with the two first formants. The correlation with F3 is smaller. Noticeable is that for both men and women PC1 does not correlate only with F1 (men: 0.476; women: 0.712) but also with F2 (men: 0.528; women: 0.369). Similarly PC2 correlates with both F1 (men: 0.696; women: −0.506) and F2 (men: −0.644; women: 0.731). When correlating F1 and F2 with each other there is no significant correlation in the female data set and a significant but very modest (0.119) correlation in the male data set. This means that both PC1 and PC2 catch variation caused by F1 and F2. This is not surprising when looking at the vowel plots of the unrotated solutions in Figure 5.9 (p. 67). The vowel plots are skewed in comparison to a formant plot, which explains that both PC1 and PC2 correlate with F1 and F2.

A large difference between the female and male solution is that PC2 and F2 show a positive correlation for the women, but a negative one for the men. This can be

Table 5.8. Correlations between formants and PCs in the varimax solutions with three components extracted. Insignificant correlations ($p > 0.05$) are indicated by a hyphen.

		correlations			expl. var.	
sex	PC	F1 Bark	F2 Bark	F3 Bark	%	cum. %
men	PC1	0.841	−0.443	–	39.0	39.0
	PC2	–	0.489	0.152	27.4	66.4
	PC3	0.181	0.522	−0.384	16.7	83.1
women	PC1	0.819	−0.442	−0.135	38.1	38.1
	PC2	–	0.504	0.345	25.3	63.4
	PC3	0.250	0.533	−0.326	19.9	83.3

Table 5.9. Correlations between formants and PCs in the varimax solutions with two components extracted. Insignificant correlations ($p > 0.05$) are indicated by a hyphen. These solutions have the strongest association between PC1 and F1 and between PC2 and F2.

		correlations			expl. var.	
sex	PC	F1 Bark	F2 Bark	F3 Bark	%	cum. %
men	PC1	0.880	−0.352	−0.185	41.1	41.1
	PC2	–	0.732	–	36.3	77.4
women	PC1	0.875	−0.360	−0.277	41.4	41.4
	PC2	0.152	0.744	–	36.8	78.2

compared to the mirrored scores in Figure 5.9 (p. 67) and the mirrored loadings in Figure 5.3 (p. 63).

When using varimax rotation, the number of components extracted influences all factors (see § 5.1.4). Tables 5.8 and 5.9 show the correlations of the rotated solutions when extracting three and two components respectively. In the solution with three components (Table 5.8) PC1 explains 39.0% of the variance for men and 38.1% for women. In the unrotated solutions in Table 5.7 the amount of explained variance on PC1 is considerably higher (men: 42.4%; women: 49.7%). The total amount of variance explained by three PCs, however, is the same (men: 83.1%; women: 83.3%). Because varimax maximizes the variance of the loadings within the extracted PCs, the relative importance of the PCs is equalized.

In the varimax solution with three components (Table 5.8) PC1 and F1 correlate more strongly with each other than in the unrotated solution (men: 0.841; women: 0.819). On the other hand, all three extracted components correlate considerably with F2 and there are only modest correlations between the PCs and F3. PCA does not seem to be able to completely separate the first three formants.

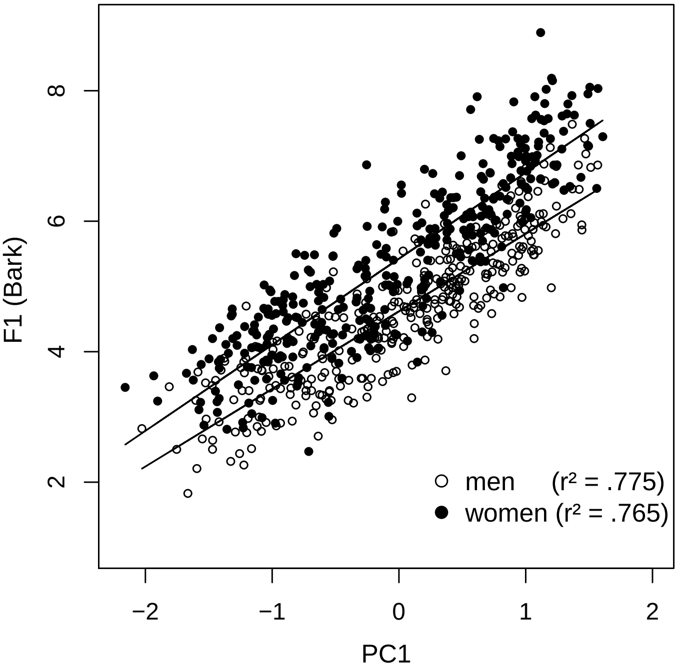
When extracting two components with varimax rotation the pairwise correlations between PC1 and F1 on the one hand, and PC2 and F2 on the other are considerably higher than in the other solutions (see Table 5.9). These correlations are also higher than the ones found by Jacobi (2009) (Table 5.6). F1 and PC1 correlate at a level of 0.88 for both men and women, and the correlation between F2 and PC2 is 0.73–0.74. PC1 also partly explains variation caused by F2 and F3.

When analyzing dialectal variation it is interesting to be able to draw articulatory conclusions about vowel pronunciation. The varimax solutions with two extracted components (Table 5.9) show the strongest direct relationship between PC1 and F1 on the one hand and PC2 and F2 on the other. As can be seen in Figure 5.8 (p. 66) this solution also gives scores that correspond to the traditional vowel quadrilateral when plotting the scores in the PC2/PC1 plane. Because this model gave the highest correlations with formant measurements, it was chosen to be used throughout this thesis for analyzing dialectal variation.

Figure 5.13 shows scatter plots of PC1 versus F1 and PC2 versus F2 of the varimax solutions with two extracted components, with separate regression lines for men and women. The linear relationship between F1 in Bark and PC1 is very strong. F2 and PC2 show a somewhat less strong relationship. The cloud is denser for high PC2/F2 values than for low values, which reflects the fact that the data comprises more front vowels than back vowels. But especially for the highest F2 values, PC2 is in some cases lower than expected. While, F1 can be well predicted from PC1, PC2 apparently includes other spectral information than F2 only (see further § 5.2.3). Because of the strong relationship between PC1 and F1, PC1 can roughly be interpreted as representing vowel height. PC2 represents vowel backness to some extent, but the articulatory conclusions based on PC2 should not be as strong as those made for PC1.

Jacobi (2009) measured high pairwise correlations between formants and PCs without using any rotation technique when extracting the PCs. One reason for this could be that the Dutch point vowels are different from the Swedish ones. Furthermore, the correlations of Jacobi (2009) are based on only five different vowels, while the current study includes 20 different vowel phonemes. This means that the correlations in the current study are based on a more varied and more continuous data set even though the total number of tokens is smaller. Because of this, the results of the two studies are not directly comparable to each other.

Even though the correlations with formant measurements are high one should still bear in mind that using band-pass filtering means that rather broadly defined frequency regions determine the representation of vowel quality. Some finer differences in formant frequencies between language varieties, possible to find by manual analysis and correction, may be lost. One advantage of using PCA of Bark-filtered spectra, however, is that the method can be completely automated, which makes it suitable for analyzing large data sets.



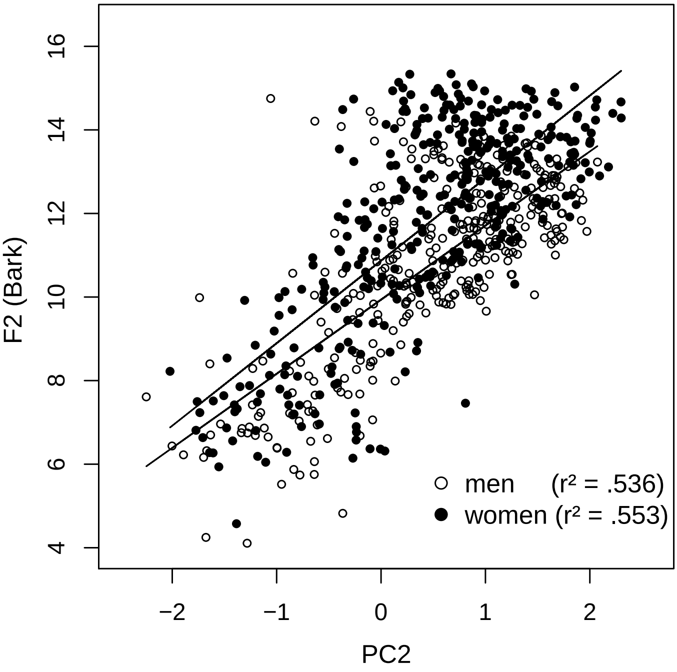


Figure 5.13. Scatter plots of PCs (2 components extracted with varimax rotation) versus formants. Regression lines are drawn separately for men and women. PC1 shows a better fit with F1 than PC2 with F2.

Table 5.10. Results of the four multivariate analyses of variance: η^2 for each significant factor ($p < 0.05$).

η^2	F1-F2 Bark	F1-F2-F3 Bark	PC1-PC2 joined analysis	PC1-PC2 sex separated
vowel	0.875	0.743	0.735	0.750
speaker-sex	0.499	0.690	0.193	0.024
site	0.038	0.054	0.062	0.066
vowel*speaker-sex	0.091	0.069	0.060	–
vowel*site	0.346	0.278	0.240	0.257
speaker-sex*site	–	0.031	0.069	0.066
vowel*speaker-sex*site	–	–	–	–

5.2.2 Multivariate analysis

One further way of comparing PCA of Bark-filtered vowel spectra with formant frequencies is to analyze to what extent the methods are able to separate different sources of variation in the data. This was done by carrying out manovas with the measures of vowel quality as dependent variables and vowel, site and speaker-sex as independent variables. Four different manovas were carried out with different dependent variables:

1. F1 and F2 measured in Bark
2. F1, F2 and F3 measured in Bark
3. two PCs extracted using varimax rotation separately for men and women
4. two PCs extracted using varimax rotation including men and women in one single PCA

In order to make the data as normally distributed as possible a few front vowels (the vowels elicited with the words *disk*, *leta*, *nät*, *särk* and *typ*) were left out of the analysis. The independent variables of the analyses were *vowel* (with 15 categories), *speaker-sex* (men and women), and *site* (Ankarsrum, Malung, Skog). The total number of speakers was 35 (17 men and 18 women), and the number of vowel tokens was 523 (due to a few instances of missing records). The significance level was estimated using Pillai's trace and the effect size was estimated by η^2, which shows the proportion of the total variance in the dependent variables accounted for by the independent variables. Table 5.10 shows the η^2 values of the significant results of the four manovas.

As expected all analyses show a significant main effect of vowel. The effect size of vowel is largest for the analysis with F1 and F2 as dependent variables ($\eta^2 = 0.875$). The effect size decreases when adding F3 to the analysis ($\eta^2 = 0.743$). This can be compared with the results of Adank (2003, 99–102), who used manova to compare the effect of a number of speaker normalization procedures. Her data consisted of recordings of the Dutch vowel phonemes by 160 speakers of Standard Dutch from eight different geographic regions in the Netherlands and Flanders with an equal distribution of female and male and older and younger speakers. Adank found an effect size of $\eta^2 = 0.893$ for vowel when using two formants measured in Hertz as dependent variables, and $\eta^2 = 0.695$ with three formants as dependent variables. The best performing normalization procedure, Lobanov's (1971) z-normalization, gave a vowel effect of $\eta^2 = 0.932$ with two normalized formants as dependent variables and $\eta^2 = 0.760$ with three normalized formants as dependent variables.

The main effect of vowel is smaller in the manovas with PCs as dependent variables than when using the first two formants. The sex separated PCA ($\eta^2 = 0.750$) gives a somewhat larger effect of vowel than the PCA where men and women were analyzed together ($\eta^2 = 0.735$) and also larger than when using three formants.

The formant-based manovas show a large main effect of speaker-sex. With three formants ($\eta^2 = 0.690$) this effect is even larger than with two ($\eta^2 = 0.499$), which was also found by Adank (2003, 101–102). F3 seems to have even more sex-dependent variance than F1 and F2, which explains why the main effect of vowel decreases when adding F3 to the analysis.

In the PCs the effect of speaker-sex is much smaller than in formants. When men and women were analyzed separately there is hardly any effect of speaker-sex in the measurements of vowel quality ($\eta^2 = 0.024$), which confirms the results of § 5.1.5. But also a PCA that includes both men and women gives a smaller effect of speaker-sex than formants measured in Bark ($\eta^2 = 0.193$).

Jacobi (2009, 57) measured the area of /i – a – u/ vowel triangles of men and women based on formant measurements in Bark as well as PCs of Bark-filtered spectra. The formant measurements in Bark showed a significant difference in the sizes of the vowel spaces of men and women. The PCs, on the other hand, did not show any significant difference in the *size* of the vowel spaces of men and women; only the *position* of the vowel triangles differed between men and women in the PC plane.

Applying PCA separately to men and women sets the mean of both groups to zero, which means that there is a correction for the different positions of the vowel spaces in the PC plane. Some speaker-specific variation is still left, but the biggest factor, sex, is removed.

Using the first two formants for measuring vowel quality seems to lead to a better separation of vowels than in principal component solutions. However, formant measurements show large differences between men and women. The speaker and sex specific variation in formant measurements can be reduced by applying normalization procedures (see § 2.4.4). But successful speaker normalization procedures are generally based on the average and/or standard deviation of the vowel phonemes, which makes comparison of language varieties with different phoneme systems or different vowel centers difficult. Using PCA on Bark-filtered spectra does not remove all speaker-specific variation in vowel pronunciation, but by applying PCA separately to male and female data the systematic differences caused by the anatomical/physiological differences between the sexes can be diminished. Moreover, a PCA can be built up on point vowels of a number of speakers and subsequently applied to a larger data set. Because of this, varieties with different phoneme systems and/or varieties lacking some of the point vowels can be compared to each other. This is an important precondition when comparing Swedish dialects, some of which have phoneme systems that deviate strongly from Standard Swedish.

5.2.3 Interpreting principal components

PCA of Bark-filtered vowel spectra was shown to correlate highly with formants in § 5.2.1. Exactly how the combination of a number of pass bands can result in a configuration so similar to formants is not completely straightforward to understand.

Pols (1977) applied PCA to band-pass filtered vowel spectra and wrote (p. 49): “This mathematically well-defined factor representation is not as easy to interpret as a formant representation.” Pols (1977, 49–50) compared vowel spectra with the loadings of the PCA in order to explain the coordinate values of the vowels in the PC plane. This is also what is done in Figure 5.14; the average spectra of the point vowels are compared with the loadings of the extracted PCs. The figure is based on male speakers only, but it would look the same when based on female speakers, as is evident from Figures 5.2 (p. 61) and 5.7 (p. 65). Loadings > 0.6 and < -0.6 on the PCs are marked with a symbol. This shows that frequencies up to 10 Bark largely determine PC1, with negative correlations with the two first Bark filters and positive correlations with Bark filters 4–10. Frequencies above 11 Bark are the most important ones determining PC2.

The lines of the open vowels [æː] and [ɑː]/[aː] are similar to the one of PC1 in starting low at the two first Bark filters and having a peak at 5–6 Bark. From this follows that they will have high positive scores on PC1. The close vowels [iː] and [uː], on the other hand, show more of a mirrored curve of the loadings of PC1 at the lower frequency regions. They start high but fall soon. Because they show the mirror of the loadings of PC1 they will get negative scores on this component.

The back vowels [uː] and [ɑː]/[aː] have a low F2 at about 6–8 Bark. Because F2 is low there is only little energy in the higher frequency areas determining PC2. Accordingly, these vowels are assigned negative scores on PC2. The front vowels [iː] and [æː], on the other hand, have a high F2; [iː] around 13 Bark and [æː] around 11 Bark. Thus, the second formant of the front vowels falls in the higher frequency regions that largely determine PC2. High energy levels in this frequency area lead to high scores on PC2. This reasoning is confirmed by looking at the scores of the point vowels in Figure 5.8 (p. 66).

Some more insight can by acquired by looking at the interactions of PC1 and PC2 at the frequency areas of the formants. Table 5.11 shows the loadings of the two PCs of the male analysis. Additionally the table indicates within which Bark filter the average formant frequencies of Swedish long vowels produced by male speakers (Eklund & Traunmüller, 1997) are placed.[9]

---

[9] Since average formant frequencies are used, one should bear in mind that the actual variation spreads across the given Bark filters.

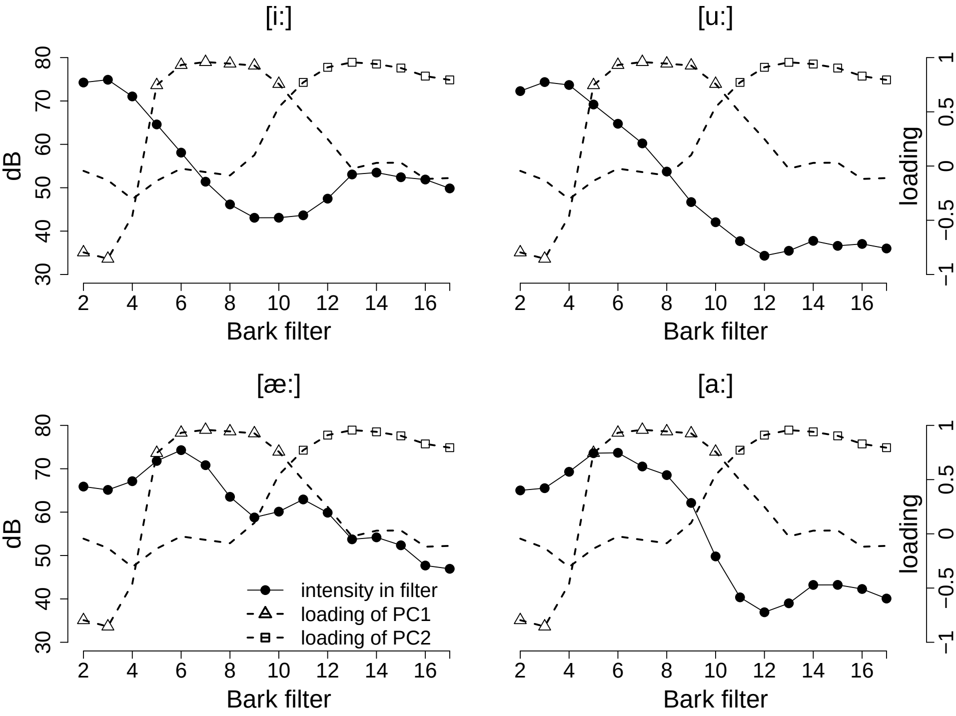


Figure 5.14. The loadings of the PCA (2 components extracted with varimax rotation; these are the same in all four graphs) and the mean intensities of each of the point vowels. For loadings between −0.6 and 0.6 the symbol is omitted. Plots are based on male data only.

The table shows that F1 of close vowels is defined by negative loadings on PC1. The frequency area of F1 of mid vowels has modest negative loadings on both extracted PCs, and F1 of open vowels is defined by high positive loadings on PC1.

F2 of back vowels has an average frequency in the same frequency area as F1 of the most open vowels. F2 of back vowels is, thus, defined by high positive loadings on PC1. The frequency area of F2 of front vowels has high positive loadings on PC2. At the frequency area of F2 of central vowels the loadings of PC1 and PC2 cross each other (compare also to Figure 5.6, p. 65), which means that F2 of central vowels is defined by moderately high positive loadings on both PCs.

Different weightings of the Bark filters on the two extracted components thus account for the varying frequencies of F1 and F2. The overlapping frequency areas of F1 of open vowels and F2 of back vowels explain the fact that the formants cannot be completely separated by the PCA, but PC1 correlates not only with F1 but also, to some extent, with F2 (see Table 5.9, p. 75).

F3 of all vowels is defined by high positive loadings on PC2, with somewhat lower loadings for the highest F3 values. Variation in the intensity at the frequency area of F3, hence, influences PC2. But because the whole frequency area has high positive loadings, differences in the frequency of F3 is not likely to be well distinguished by PC2.

Figure 5.13 (p. 77) shows that the relationship between PC2 and F2 is weaker than the one between PC1 and F1. As can be seen in Table 5.11, PC2 is strongly influenced by a frequency area higher than where F2 is found. PC2 is, hence, a combination of the effect of F2 and of higher frequency regions. The varying intensities in the higher frequency area most likely causes the variation in PC2 which cannot be explained by F2.

Since Bark filters correspond to the critical bandwidth of human hearing and include information from the whole vowel spectrum, they model perception very well. Formant frequencies have on the other hand been shown by numerous studies to be very important cues for the perception of vowel quality. The PCs that result from reducing Bark filter data to the most important underlying components are influenced to a great deal by formants, as has been shown above.

Results from perception experiments that specifically compare the role of formants respectively PCs of Bark-filtered spectra in perception would be interesting in order to understand the relationships even better.

Table 5.11. Mean formant frequencies of Swedish long vowels produced by male speakers according to Eklund & Traunmüller (1997), and the loadings of PC1 and PC2 (two components extracted with varimax rotation; loadings >0.6 and < -0.6 in boldface). Multiple vowels within the same Bark filter are given in ascending order of formant frequency.

	loadings		formant frequencies		
bf	PC1	PC2	F1	F2	F3
2	**−0.80**	−0.04			
3	**−0.85**	−0.13	[yː] [iː] [uː] [ʉː]		
4	−0.46	−0.31	[eː] [oː] [øː]		
5	**0.75**	−0.14	[ɑː]		
6	**0.93**	−0.02	[æː]	[uː] [oː]	
7	**0.96**	−0.06			
8	**0.94**	−0.09		[ɑː]	
9	**0.93**	0.10			
10	**0.76**	0.54			
11	0.49	**0.77**		[æː] [øː]	
12	0.25	**0.91**		[ʉː]	
13	−0.02	**0.96**		[yː] [iː] [eː]	
14	0.03	**0.94**			[øː] [æː] [uː] [ʉː]
15	0.03	**0.90**			[ɑː] [oː] [eː] [yː]
16	−0.12	**0.83**			[iː]
17	−0.11	**0.79**			