Chapter 3

Aims and research questions

The main aim of this thesis was to describe the geographic variation in vowel pronunciation across the Swedish language area. As described in §§ 2.2 and 2.3 both rural Swedish dialects and regional varieties of Standard Swedish vary a lot when it comes to vowel pronunciation, and vowels have been important for characterizing varieties of Swedish and classifying dialects. Still, no exhaustive acoustic description of variation in Swedish vowel pronunciation exists (Bruce, 2010, 103).

By carrying out an acoustic analysis of vowels from a large number of varieties of Swedish, I hoped to be able to answer the following questions:

1. How is the variation in Swedish vowel pronunciation distributed geographically?
2. Do different vowel features show co-variation?

The Swedish dialects have undergone big changes during the last century. General questions to answer include questions about the dialect situation around year 2000:

3. How large is the dialectal variation and in which areas are divergent rural dialects still spoken?
4. Which Swedish dialects are changing? Which are stable?

The data set analyzed in this thesis includes speakers of two generations, which made an apparent time study of language change possible. Based on the societal and linguistic changes described in § 2.1.2 the hypothesis was that the distances between dialects would be shorter in the younger generation of speakers than in the older generation. Relevant questions were:

5. How much change in vowel pronunciation can be observed between older and younger speakers?
6. Which vowels are changing and in what direction?
7. Which vowel features are stable?

Gender is a social variable that has been shown to correlate with linguistic variation in many studies (Chambers & Trudgill, 1998, 61). The gender-related variation in vowel pronunciation was not studied in as much detail as the variation across age groups in this thesis, but at an aggregate level an answer was sought to the following question:

8. Is there gender-related variation in Swedish vowel pronunciation?

Principal component analysis of Bark filtered vowel spectra was chosen as a measure of vowel quality for this thesis, since this approach is more reliably automatable than formant measurements. A representation in Bark filters gives a good perceptual representation of vowels, because the Bark scale corresponds to the critical bandwidth of human hearing. Acoustic analysis of vowels is not unproblematic, as explained in § 2.4. The main problem for dialectological and sociolinguistic studies of vowel pronunciation is how to normalize for speaker variability related to the anatomy/physiology of speakers in order to be able to analyze linguistic differences. A large number of normalization procedures have been proposed (see § 2.4.4), but most of them depend on the varieties being compared sharing some common traits that can be used as a basis for the normalization. When no common denominators, like comparable mean values and standard deviations or common point vowels, exist normalization fails.

The Swedish dialects show so much sub-phonemic and phonemic variation in vowels that the kind of common denominators mentioned above cannot be found for all dialects. The present study included a relatively small number of speakers per variety, and in addition the number of men and women was not equal in all speaker groups. Using pure group averages of the acoustic measures for reducing the influence of speaker-dependent variation would have been biased by the systematic differences in the vowel-spaces of men and women. A question related to the acoustic analysis of the vowels was:

9. To what extent can speaker-dependent variation in the acoustic measures be reduced?

Dialectometry has introduced aggregate analysis of dialectal variation as an alternative to detailed analysis of separate variables. Aggregate analysis allows the researcher to find out how dialects relate to each other when all available data is considered simultaneously, instead of looking at individual features, which is what generally has been done in traditional dialectology. Methods commonly used for aggregate analysis of dialects are cluster analysis and multidimensional scaling. The problem of how to identify linguistic structure in the aggregate has not been completely solved yet. A methodological aim for this thesis was to analyze the relationship between aggregate analysis and underlying distributions of individual features. In the paper *Factor Analysis of Vowel Pronunciation in Swedish Dialects* (Leinonen,
2008) I showed that factor analysis is an effective method for identifying linguistic features that show similar geographic distributions, and displaying these distribution patterns on maps. By comparing the results obtained by factor analysis and by multidimensional scaling in this thesis I wanted to approach the questions:

10. How can analysis on the variable level and aggregate analysis supplement each other in the study of dialectal variation?

11. Can a comparison of variation on the variable level and aggregate analysis explain what kind of variation the aggregate analysis accounts for?

As the aggregate analysis shows how varieties relate to each other when all variables are considered, it can provide a basis for a dialect classification. Questions for the aggregate analysis were:

12. How can the Swedish dialects be classified based on vowel pronunciation?

13. Does a classification of modern varieties of Swedish correspond to traditional divisions of Swedish dialects?

The data for this study comes from the SweDia database (see Chapter 4). Within the SweDia project work has been carried out that encompass the same Swedish sites and speakers as the present study but other linguistic levels (see § 2.2.3). A comparison with these studies can give an account of the association between linguistic levels:

14. Does a classification of Swedish dialects based on vowel pronunciation correspond to typologies based on other linguistic features?

Answers to questions 1–7 are given in Chapter 6, where variation on the variable level is studied. Question number 1 is more specifically related to § 6.1 and question 2 to § 6.3. Answers to questions 3–7 are found in all sections of Chapter 6.

Chapter 5 describes the acoustic analysis of the vowels. Question number 9 is being dealt with more specifically in § 5.1.5.

In Chapter 7 aggregate analyses are described and answers are given to questions 8, 11 and 12. Also questions 3–5 are partially answered in Chapter 7.

In Chapter 8, question number 10 is approached by comparing the results of Chapters 6 and 7. Questions number 13 and 14 are not subject to any quantitative analyses but are discussed in general terms in § 8.2.1.