MAPPING VOWELS

Variation and change in the speech of Gothenburg adolescents

Department of philosophy, linguistics and theory of science

MAPPING VOWELS

Variation and change in the speech of Gothenburg adolescents

Johan Gross

Doctoral dissertation in linguistics, University of Gothenburg
20180907

Cover: Thomas Ekholm
Typeface: Times
Printed by Repro Lorensberg, University of
Gothenburg, Gothenburg 2018
Publisher: University of Gothenburg
(Dissertations)
ISBN 978-91-7833-125-3 (print)
ISBN 978-91-7833-126-0 (pdf)

Distribution
Department of philosophy, linguistics and theory of science,
Box 200, SE-405 30 Gothenburg

ABSTRACT

Title: Mapping vowels: Variation and change in the speech of Gothenburg adolescents

Author: Johan Gross

Language: English, with a summary in Swedish

Department: Department of philosophy, linguistics and theory of science

ISBN 978-91-7833-125-3 (print)

ISBN 978-91-7833-126-0 (pdf)

Until now, sociolinguistic investigations of Gothenburg have been scarce. This compilation thesis provides some of the first steps in a quantitative investigation of language variation and change in the city. Its results suggest that sociolinguistic variation ought to be studied with a bottom-up approach investigating on the one hand relationships among variables in what might be a coherent linguistic system and on the other hand the intersection of a range of social categories, rather than relationships between one category and one variable at a time. In three of the articles, analyses of variation in long vowel productions (and in one, also perceptions) are carried out. In the remaining article, the interaction between the informants in the activity used to collect data is described and analyzed.

The data used for the articles are drawn from two corpora with young adolescents living in Gothenburg (and Stockholm): the SUF corpus, which consists of recordings of 222 informants collected for another project, and the SSG-corpus, material collected for this dissertation, which consists of recordings of 111 informants in two activities, an interview and a map-task. Acoustic and statistical analysis were carried out and analyzed on parts of both corpora. The second study in the collection describes how a map-task can be used as a sociolinguistic tool to collect a large number of tokens in a relaxed interaction between friends.

The results from the first article in the collection show that variation in production of /ɛ:/ is related to the foreign vs. Swedish background of young informants in Gothenburg, but no significant differences were found in Stockholm. The third paper in the collection investigates not only this vowel but also eight other long vowels (or allophones). The acoustic and statistical analyses show that the pattern of variation can be captured by considering how foreign background conflates with socio-economic status and educational background. These categories can be captured by considering neighborhood as a factor, due to housing segregation in Gothenburg. The final paper provides acoustic and perceptual evidence that there is an incipient merger between /i:/ and /y:/ in Gothenburg. It also suggests that the reason behind the direction of this vowel merger is that lip-rounding is a perceptually weak feature, so the merger is in the direction of rounded to unrounded and not vice versa.

Two general conclusions can be drawn from the thesis as a whole. The first is that to understand and describe sociolinguistic variation, proper attention needs to be given to how social categories intersect in a specific context before a study of the meaning attached to variation can be carried out. The second pertains to the ontological status of sociolinguistic variables. Some variables are undergoing widespread change; others are more locally bound. The adolescents in segregated suburbs can both be seen to lead widespread changes while not participating in more local variation and change. This pattern might indicate an orientation away from local dialectal norms.

FÖRORD

Det finns fyra personer som jag vill rikta ett särskilt stort tack till, det finns så klart fler, men utan er så hade inte denna avhandling blivit vad den är.

Först och främst Sally Boyd, min huvudhandledare. Du har alltid varit en tålmodig och trygg punkt, någon jag vågat vara öppen mot och fråga när jag inte förstått eller behövt stöd. Av dig har jag lärt mig alla de där sakerna som inte går att läsa sig till. Du har varit en klippa som jag alltid har kunnat lita på en stor inspirationskälla och förebild jag hade inte kunnat drömma om en bättre huvudhandledare.

Min bihandledare Therese Leinonen, du har tålmodigt stått ut med mina frågor och noggrant, korrekt och pedagogiskt svarat på ett sätt som fördjupat min kunskap och sporrat mig att lära mig mer.

James Walker min guide i den internationella forskarvärlden. Det var du som verkligen fick mig att förstå sociolingvistisk teori och metod.

Sist men inte minst, Julia Forsberg världens bästa kollega, jag fattar inte hur du orkat med mig. Du har gjort så att jag känner mig lugn och trygg i arbetsvardagen, inspirerat och stöttat, flamsat och tramsat alltid ställt upp när jag behövt hjälp. Jag hoppas att vi följs åt i vårat fortsatta forskarliv.

Göteborg, Augusti, 2018

Johan Gross

LIST OF STUDIES

STUDY 1

Gross, Johan, Boyd, Sally, Leinonen, Therese, & Walker, James A. (2016). A tale of two cities (and one vowel): Sociolinguistic variation in Swedish. *Language Variation and Change*. 28: 225–247. [Permission waived]

STUDY 2

Forsberg, Julia, & Gross, Johan. (unpublished). *"You change your speech depending on who you talk to, but I didn't change much": The map-task viewed through the lens of audience design*.

STUDY 3

Gross, Johan. (in press). Segregated vowels: language variation and dialect features among Gothenburg youth. *Language Variation and Change*. [Permission waived]

STUDY 4

Gross, Johan & Forsberg, Julia. (under review after revision) *Weak lips? A possible merger of /i:/ and /y:/ in Gothenburg*. [With permission of the copyright holder.]

PART 1: THESIS FRAME

PART 2: STUDIES

APPENDIX

CHAPTER 1

Introduction

1.1 Overall aims

The linguistic situation in today's major cities is characterized by heterogeneity, which originates from large-scale population movements from rural areas, other cities, or other parts of the world. Thus, urban centers are places where people with different dialects and languages come in contact. This process of urbanization is far from new, although it has accelerated during the 20th Century and today, cities are growing and changing rapidly. This rapid change in population inevitably leads to increased language variation and change (Trudgill 2011), something that sociolinguists have been aware of ever since sociolinguistic theory and practice started to contribute to general linguistics. As Labov described it when investigating language variation and change in New York in 1966, "Variability is an integral part of the linguistic system and no less a part of the behavior of the city." (p. 3), highlighting that variability is something as natural in the linguistic system as it is natural in urbanized areas and that the two phenomena are intertwined.

Still, theoretical constructs which rest on an assumption of homogeneity and stability from linguistic sub-disciplines such as dialectology have often been reified in sociolinguistic studies over the years. One of these is the idea of the native speaker, a theoretical construct that has attracted criticism (e.g. Boyd & Fraurud 2010, Fraurud & Boyd 2011), although the consequences of its use are still found in linguistic studies and the concept has had implications for the linguistic description of the language spoken in a city. One such case is the aforementioned New York study by Labov (1966), where the language of a rather 'homogeneous' group in a geographically small area of the city was assumed to represent the language spoken in the city as a whole (criticized by Horvath 1985). Sociolinguistic research in Sweden is by no means free from these constructs, and the description of the linguistic situation in Sweden's two largest cities (Stockholm and

Gothenburg) has been skewed as studies have been carried out assuming there to be well-defined speaker groups, such as working class or immigrants, within the city, without properly investigating if these groups are the most relevant to the question of linguistic variability. A contributing factor to these skewed descriptions is the lack of systematic descriptions of the dialects assumed to be spoken in the cities. For example, the best description of the language spoken in Gothenburg dates back to Björseth's dialectological work published in 1957, which was slightly updated in 1976 by Holmberg and the situation is only slightly better for the Stockholm dialect. These descriptions do not fulfil the methodological requirements that modern sociolinguistic theory puts on a study, leaving a gap to fill in 2018, both with regards to what is described and how it is carried out.

The overall aims of this thesis are therefore both descriptive and theoretical. The first aim is to contribute one piece of the descriptive puzzle, in the form of a description of the vowel system of Swedish as currently spoken by young people in Gothenburg. Although the description will be limited to vowels produced by speakers in the age range between 16-20 in a single activity, this description can, however, not be done without considering the nature of the city's population and what factors might actually contribute to variation in the city. Therefore, the second aim is to investigate the social factors motivating variation in the vowel system on a macro level in Gothenburg specifically, and – hopefully – also more generally. Today, terms such as *multiethnolect* are used as labels for linguistic practices connected to foreign background. But, as pointed out by Jaspers (2008) and Eckert (2008), this kind of terminology puts forward ethnicity as the strongest factor behind linguistic practice, on the one hand ignoring that it is only people with a foreign background that are considered as ethnic and on the other hand, as is often pointed out, that youths with a non-foreign background also often use these same features, or speech styles (Eckert 2008, contributions to Källström & Lindberg 2011). In addition, the other possibility – that persons with a foreign background do not use the alleged ethnolectal features – seems also to be ignored. If this is the case, are they behaving more in accordance with a local, 'native' or non-foreign norm? The second aim of the thesis is therefore to investigate if features that are *not* commonly associated with language contact are used in a similar way between speakers with and without foreign background.

Finally, the linguistic features that have thus far been associated with alleged new urban varieties have in common that they are all highly salient, often associated with second language learning (e.g. diversions from the V2 rule, grammatical gender, loan words etc.) and relatively easy for any speaker of the language to incorporate into their language use, although social pressure will delimit their use. This thesis aims to raise the question of what linguistic variables other than those that are connected with L2 language features should be used in order to ascertain if different groups of speakers in a community share the same linguistic system. The suffix *-lect* seems to imply a shared system; however, assuming some kind of -lect on the basis of a few salient features seems shaky. A more general aim of the thesis is to take a step back from the study of single variables and try to see if an examination of a larger part of the linguistic system might give us another picture of linguistic variation. As already mentioned, this thesis aims to give a description of the long vowel system of the Swedish used in Gothenburg. Vowels are nice in this way; they can be described as a system, i.e. a set of components that are linked to each other so that they form an orderly whole[1]. So, we can ask about the relations, or links, between the vowels, rather than simply describing their variation as the phonetic realizations of single phonemes. The final aim of this thesis is therefore to examine something that is more difficult for the speaker to manipulate and less salient, namely the (articulatory) *coherence* (Wiese & Rehbein, 2016:46) that can be observed between vowels. Some vowels tend to co-vary so that they appear in pairs, for example /i/ and /y/ tend to have the same articulatory position. This thesis aims to raise the question if coherence between variables could be used in order to understanding language variation and change in a city experiencing both in-migration from other parts of Sweden and immigration from other countries.

To sum up, there are three general aims of this thesis:

1. To provide a sociolinguistic description of the vowel system of Gothenburg Swedish

[1] This is my translation of the definition of *system* in the Swedish National Encyclopedia: *"*samling element som hänger samman med varandra så att de bildar en ordnad helhe*t"* (Nationalencyklopedin).

2. To find out how speakers in different groups are using features associated with more general on-going changes in the linguistic system. What is the relationship between linguistic variation and social factors such as foreign background, gender, class and neighborhood?
3. a. To investigate long vowels, which are less salient features of the alleged new varieties of urban Swedish, than variables previously studied.

   b. To investigate degrees of coherence among long vowel phonemes and allophones. To do this, it will be necessary to study the vowels as a system, rather than as individual variables.

1.2 Structure of the thesis

The thesis frame as a whole provides a more extensive discussion of the theoretical and methodological background for my work than is possible in the individual articles, and places the work in a broader research context. In chapter two of the frame, the theoretical framework for the studies is described, including a discussion of key concepts and relevant previous research within the field of variationist sociolinguistics, sociophonetics and studies of language contact in urban settings. Chapter three describes the material used for the studies, which consists of two corpora, as well as the methods selected for acoustical analyses, vowel normalization, perception experiments, transcription and interactional analysis. Chapter four summarizes the individual studies and my contribution to each study. In chapter five, I draw conclusions based on all four studies and develop some ideas which have arisen subsequent to the finalization of the articles; the chapter ends with some ideas for future research.

CHAPTER 2

Theoretical background and framework

> *"I have resisted the term sociolinguistics for many years, since it implies that there can be a successful linguistic theory or practice which is not social. "(Labov 1972, p: xiii)*

In this chapter I will outline some of the more important aspects of the branch of sociolinguistics often called variationism and to place the research in this thesis within it. A primary aim of variationist theory has been to contribute to answering one of the most fundamental questions in linguistics, namely the one about the nature of linguistic competence as an internalized part of the human mind. An important part of variationists' contribution has been to examine how the social life of humans is part of what predicts and shape linguistic structure. For example, gender and ethnicity have often been shown to be social factors correlating with linguistic variation. The question, however, is not if these and other social and contextual factors are significant in certain cases, but why. The why question will demand more in-depth analysis of the social life that creates the more abstract categories, as Eckert describes it (2016:11):

> *"(1) variation is socially meaningful and (2) variation patterns socially on both the macro- and the micro-social scale. A robust theory of variation will integrate these facts, tracing the links between local stylistic practice and macro-social patterns, examining meaning-making on the ground in view of the conditions on life represented by macro-social categories."*

In other words, we need to combine perspectives on different levels of abstraction, to be able to construct a viable linguistic theory. However, no matter the perspective, some fundamental concepts seem to be shared among linguists within the variationist paradigm while other concepts have been the object of debate,

which has pushed the theoretical development forward; some of the concepts and objects of debate will be discussed in 2.1 and in 2.2 I will discuss my use of the terms *language*, *dialect* and *variety*. In section 2.3 I will move on to discuss relevant research about language variation and language use in bilingual settings. The paradigm of sociolinguistics which provides a background for the work in this thesis has at times been practiced within the field of phonetics suing the term *sociophonetics*. The methodological implications of this practice will be discussed in 2.4. Relevant variationist research in Sweden will be presented in section 2.5. A summary of the theoretical background of the thesis is provided in 2.6.

2.1 Variationism

The quantitative study of language in large part originated with William Labov's studies in New York City during the 1960s (e.g. 1966, 1968). In these studies, the point of departure was the *speech community*, i.e. a group, rather than an individual and her/his idiolect. With this point of departure, it became possible to see that what was previously assumed to be free variation or performance errors (e.g. Saussure 1916, Chomsky 1963), in fact was *orderly differentiation* (later more aptly called *orderly heterogeneity*), which Weinreich, Labov and Herzog (1968) argued must be an integral part of the linguistic system and as such part of a speaker's competence. The approach of Labov's New York study and the concept of orderly differentiation were corroborated in similar studies in other parts of the USA and in other countries during the late 1960s and 1970s, (e.g. Trudgill 1974, Wolfram 1969, Nordberg 1975). This work created an empirical foundation showing that language variation patterned in accordance with macro sociological factors such as socioeconomic status, gender, and ethnicity, as well as contextual and internal linguistic factors.

Once the group was established as the object of linguistic study, it became necessary to define the speech community as a theoretical and scientific object. Kerswill (1993: 35-36) tracks four criteria for a speech community based on Labov's variationist work during the 1960s and 1970s: (1) the nativeness of speech community members; (2) the presence of uniform patterns of linguistic variation; (3) the shared social evaluation of linguistic parameters; and (4) the systemic identity (or very close relatedness) of the linguistic varieties on all linguistic levels.

The first criterion, nativeness, is usually met by setting some kind of age of acquisition and stability of residence criteria on the sample population. In an urban setting this runs the risk of leading to a rather drastic decrease in target population. However, without some kind of age of acquisition criterion the definition runs the risk of being too inclusive. The second criterion is based on Labov's (1966) findings that the frequency of certain features will increase or decrease predictably with social status and the formality of the speech style. This criterion seems to assume that variation is unidirectional and that the vernacular is a base that the speaker orientates away from, depending on the formality of the context (more on the vernacular later in this section). The third criterion is based on Lambert's et al. (1960) matched guise findings, that speakers in the same community (even belonging to different first language groups) share a set of norms and evaluations toward linguistic forms and variants. And, the final criterion is according to Kerswill (1993: 36) never postulated by Labov, but, something that seems to be presupposed by the analytical procedure he uses. It assumes that phonological variation will be on a phonetic continuum or as discreet variants within a phoneme. Kerswill (1993) tested these criteria on the speech of people living in Bergen, Norway, and showed that the other criteria were fulfilled by the 'native' informants. However, in-migrated informants from neighboring dialect areas were shown to share the native informants' norms and conceptions about their own as well as about the Bergen dialect, which explained patterns in the in-migrated informants' speech. Kerswill argues on the basis of these results that the city as a social structure will connect smaller speech communities into a larger one, where the smaller ones fulfill the Labovian definition and the larger one is based on shared norms of varieties as a whole.

A third and central concept in the theoretical construct of variationist theory is the notion of the vernacular. Labov (1972:112) defines this as follows: "*the style which is most regular in its structure and in its relation to the evolution of the language is the vernacular, in which the minimum attention is paid to speech"*[2]. In contrast to Labov, who situates the vernacular as an unreflected (Eckert (2012) calls

[2] However, later Labov (2001a) adjusts this definition arguing that the 'attention paid to speech' part was a definition intended to capture the stylistic variation in the data collection methods (interviews and reading tasks) he used for his dissertation, i.e. not necessarily the vernacular but something close to it.

it "passive") speech style belonging to the individual speaker, Eckert (2001) situates the vernacular as a clearly differentiated style associated with a group of speakers. With Eckert's definition, the individual speaker is disconnected from the vernacular and it is seen as something created and updated by groups of speakers connected to some kind of social structure i.e. a speech community or a community of practice. There is, however, one thing that the two definitions have in common: they presuppose that the vernacular is but one style among many that members of a speech community master. As such, the vernacular seems to be some kind of baseline, but far from the only object of study if we want to understand language variation and change. However, the possibility of ascribing social agency to the use of a vernacular speech style separates them. In the early work of Labov (1966, 1972) and Trudgill (1972), the vernacular was connected with local values, covert prestige and the working class, while the shift away from the vernacular was explained by attention paid to speech in more formal contexts. However, the use of attention paid to speech also made the vernacular a passive speech style that could not be explained as something that the speaker used with social agency. This notion of the vernacular was difficult to connect with the findings of, for example, Milroy (1986), who showed that vernacular speech was connected to more multiplex and dense networks. She argued that the use of a vernacular style was a way of engaging and strengthening local norms, established in the networks. More recently, Eckert (2016) has argued that since the concept of the vernacular does not capture the stylistic range that we as linguists can observe, it might be better to talk of a linguistic habitus, e.g. a cognitive embedding of the entirety of the individual's experience of social interaction, so that linguistic habitus should be understood as the sum of all styles and their associated roles in the social world. In this way, Eckert removes the vernacular from its special status within the variationist theoretical construct. This actually follows from Labov's claims that there is no such thing as a single style speaker and his belief and precept that the object of linguistic study should be the group, rather than the individual.

The discussion of variationism above leaves many later assumptions or principles of variationist theory uncommented. My intention in this section has not been to give a complete description of what variationism is but to highlight those concepts which I have identified as most central in connection to aims of the paradigm and the uses I have made of it in this thesis. These concepts are not always

used directly in the papers, but with the understanding that these concepts are relevant in the theoretical construct that underlie them, the studies can rise above just being descriptive and connect their results to a more abstract variationist theoretical construct, thus aiming to understand linguistic variation and change and to connect variation and change to human linguistic competence. The most robust concept out of the three discussed above, seems to be ordered heterogeneity. With all of the variationist research corroborating this concept, a description of an understudied speech community in its widest sense needs to begin, as Eckert (2016) describes it, with the macro sociological patterns, in order to provide a base to connect more micro patterns. In my own studies, starting on the macro level, I have applied a rather loosely defined concept of the speech community as a point of departure, and the only criterion for inclusion of an informant has been their age of acquisition and continued residence. However, the question as to when an individual started to learn language X or came in contact with variety Y is a nearly impossible question to answer if it is not in recent time, or the learning has taken place as an adult. There is today ample research (e.g. Milroy 1986, Eckert 1989, 2016, Sharma 2011) claiming that Labov's concept of the vernacular is outdated, which leads me to argue that it is not a good construct to use. And my aim in this work has been to study the vowel system in a sample of relaxed, casual speech – carrying out a complex map task with a friend. What this type of speech is similar to is discussed in Forsberg and Gross (unpublished), and hopefully more research will shed light on it. However, some features seem to be more robust than others across styles and harder to manipulate; I believe one such feature is coherence in the vowel system. If this is the case is however hard to say and further research on the stability of different features across styles is needed.

2. 2 Language, dialect and standard language

The terms, *language*, *dialect* and *standard language* are central to the theoretical construct of my argumentation in the papers. These concepts are difficult to define, particularly *language* and *dialect*, but all of them are so commonly used in linguistics that they are often taken for granted. In the first paper, 'language contact' and 'dialect contact' are central concepts, and the picky reader might ask; what do you mean by *language* and how do you separate this from a *dialect*?

Hyltenstam (2010 p. 116-117, own transl.), discussing whether Meänkieli should be considered a language in its own right or as a dialect of (Sweden) Finnish, tries to separate the two concepts by postulating seven criteria for distinguishing them:

1. For a variety to be considered as a dialect of a language it has to have a historical kinship [with other varieties of the language], so that there are similarities in pronunciation, grammar and lexicon and that the parallels are clear [between the varieties].
2. Varieties that are mutually intelligible are in most cases (but not all) considered dialects of a language, while varieties that are not mutually intelligible are considered languages.
3. Varieties that are standardized and have some kind of orthography are usually considered a language rather than dialect.
4. Standardized varieties that have official functions in a state are usually (but not always) considered a language rather than a dialect.
5. Standardized varieties that don't have any official functions in a state are usually (but not always) considered a language rather than a dialect.
6. Varieties that don't have orthography and that are associated with a standardized language are usually considered dialects rather than languages.
7. Varieties that the speakers themselves consider languages are usually considered as such.

None of the criteria are waterproof but they provide an (unusually detailed) background to my own thinking about dialect and language in the papers that constitute the thesis. With these criteria in mind, *dialect contact* is, for example, contact between geographically delimited, historically related and mutually intelligible varieties of a generally accepted language e.g. in the case of Sweden, contact between Dalecarlian (*dalmål*) and Scanian (*skånska*). Language contact on the other hand is contact between varieties that have standardized varieties and are generally accepted as languages e.g. contact[3] between Swedish and German (not to

[3] In the thesis dialect and language are compounded with *contact*, and as such we might also ask: what constitutes "contact"? However, the important aspect here is not to go in to a discussion of

forget that language contact in Sweden today can include languages that don't fulfill all of them e.g. Kurdish).

Just as languages can be divided into smaller entities e.g. dialects, the type of dialects I exemplify can be divided into smaller entities, like a Matrjosjka (the Russian wooden doll that hides a wooden doll within itself and so on). However, there is no need, as a point of departure in this thesis, to divide the dialects of Stockholm and Gothenburg into any smaller geographical entities. The city has brought together people from different geographical areas of the same country that all see themselves as speakers of, in this case, Swedish, but of different varieties of Swedish associated with different geographical parts of the country. Worth noting here is that there is a tendency in the Swedish dialectological literature to also write about and discuss the varieties in the cities as dialects e.g. the Gothenburg dialect or the Stockholm dialect. I will also use the term *dialect* when I write about what previous research has described as the varieties spoken in Stockholm or Gothenburg, as all descriptions are generalizations, some level of abstraction must be assumed. Therefore, the Gothenburg dialect is not assumed to be something you will find anyone speaking all the time. There is always variability and variation at all linguistic levels.

Connected to these two concepts is the concept of standard language. Milroy (2001: 531) provides a definition of standardization: "standardization consists of the imposition of uniformity upon a class of objects" i.e. the removal of variation. With this definition, a standard language would be a language free of variation. So, if you take variationist theory seriously, such a language is impossible since language is inherently variable. Although a person can use forms that are in accord with the standard description, a standard language will not be anything that anyone produces when speaking, or for that matter writing, as there will always be variation and choices to be made. Instead, I view the standard language as a highly idealized description that can be used as a linguistic tool or maybe more accurately as some kind of linguistic navigation point in a linguistic analysis, but never produced by a speaker in spontaneous speech. The standard language is thus an idea codified in a

what should be counted as contact but rather to make it clear that languages and dialects come in contact because people come in contact. In other words, languages and dialects do not have contact with each other independent of speakers.

description and not something possible to acquire. In this sense the linguistic habitus stands in opposition to the standard language, as something actually produced and acquired, situating the object of scrutiny in the actual world and as something that can be recorded and researched.

2.3 Ethnolects

The investigation of and conclusions regarding the labeling and thereby categorization of language use in multilingual settings in Europe has been an integral part of sociolinguistic study ever since Clyne (2000) coined the term *multiethnolect,* based primarily on reports from Kotsinas (1992) and Rampton (1995). Clyne argued that new ethnolect varieties had surfaced in Europe similar to those described in North America, although the social circumstances differed somewhat compared to e.g. Afro American Vernacular English. Clyne saw the new types of ethnolects as a result of speakers using the national language as a lingua franca in communities where native speakers of no language, not even the national language, were dominant. While the lingua franca view of ethnolect has not gained that much attention, the term multiethnolect has become widely used to describe alleged new varieties resembling what in North America have been called ethnolects.

Early on Kotsinas (1988a, 1988b/2014) found that youths in the suburb of Stockholm called Rinkeby, had started to create a radically new way of speaking; she called it *Rinkebysvenska* inspired by the informants' own naming practices. In these studies of youth language in Stockholm, Kotsinas (1988a, 1988b/2014, 1992, 1994, 1998, 2001) focused on new innovative features which, she argued, were connected to typologically marked subsystems of Swedish. Her assumption was that language contact was the underlying factor behind the variation, because of the multilingual character of Rinkeby. Widening the scope of her research on youth language, Kotsinas (1994) investigated language variation among adolescents in western, northern and southern parts of Stockholm. In this study she looked at variables from all levels of the linguistic system: both those associated with the traditional dialect and the innovations associated with language contact. However, she did not investigate if the variables associated with the traditional Stockholm dialect were used by the youths in the western multilingual area and vice versa,

whether the innovative features found in the west were used by youths outside the multilingual area. Furthermore, the social situation of Rinkeby was not thoroughly investigated and situated with in the social context of Stockholm, rather, foreign background was used as an explanation without investigating other potentially intersecting social categories.

In Denmark, Quist (2008) and in Norway Svendsen & Røyneland (2008) employ a qualitative perspective investigating the adolescents' language use as strategically deployed resources within communities of practice. However, they argue for the use of the term *multiethnolect* because of the systematic use of variables similar to what is found in dialects or ethnolects with the difference that the language use is not delimited by geographic factors or by a single ethnicity. Wiese (2009) also approaches *Kiezdeutsch* (the German counterpart to *Rinkebysvenska*) as a multiethnolect, comparing new syntactic constructions in the variety to standard German.

A somewhat more restrictive approach when categorizing the language use among adolescents in multilingual settings is made by Schooen and Appel (2005). They investigate the reported use of what in the Netherlands is called Street language and found that all young people (including adolescents with a non-multilingual background) were using it though the adolescents with Surinamese background used the features associated with Street language most frequently. The use of Street language was contextually sensitive so that the young people were more inclined to use Street language in peer talk and in informal settings, leading Schooen and Appel to describe it as a multilingual youth register.

Their results also show that there are problems in naming a linguistic practice as some kind of ethnolect, as discussed by Jaspers (2008). He argues that the naming practice works to homogenize groups that are not necessarily homogeneous, or maybe should not even be considered to be well-defined groups. These naming practices also create unnecessary abstractions concealing stylistic practices that speakers deploy to create their persona. In his argumentation Jaspers sees ethnicity as something that the individual creates using linguistic resources, reminiscent of Rampton's notion of crossing (1995), and how feminists and gender researchers regard gender, as something one does rather than something one has. The use of the term *ethnolect* and what this usage conceals has also been discussed by Eckert (2008) when studying the complex nature of day-to-day interaction in

two Northern California elementary schools. She showed that even if there are distinctive ways of speaking for different ethnic groups, the motivating factor for sound changes and variation might not be ethnicity but socially indexed values that might (or might not) originate within a certain ethnic group. However, rather than marking ethnicity, use of ethnolectal variants reflects the peer-based social order at the school i.e. the features are free to use for all the kids in the school in their orientation towards a social identity not just an ethnic one. In this way variation and change are not motivated by ethnicity but by identity work transcending ethnicity. Jaspers' (2008) and Eckert's (2008) approach together with those of others makes it possible to describe how a person is navigating towards or away from an ethnic stereotype (cf. LePage & Tabouret-Keller 1985).

Hoffman and Walker (2010) try to capture this "navigation" using a variationist macro-perspective when examining the social stratification of the Canadian vowel shift and t/d-deletion in Toronto, by introducing the notion of ethnic orientation (EO) as an independent variable. This variable is described as an emic, subjective variable based on the informant's self-reported degree of affiliation with an ethnic group and is tested together with the classic way of categorizing informants in ethnic groups solely by their ethnic heritage. Hoffman and Walker found that the phonological and morphological factors conditioning final consonant cluster simplification, known as t/d deletion, in the Anglophone population as a whole did not condition t/d deletion in the group of Chinese first-generation immigrants. This difference in morphological and phonological rules was, however, not passed on to the second or third generation Chinese, who all deleted t/d according to the general morphological and phonological rules conditioning this variable in Toronto as a whole. However, the degree of use of the variables differed in the second and third generations, with regards to EO although ethnicity by itself was not significant. Hoffman and Walker interpreted this as an argument against the ethnolect hypothesis, or at least a strong version of this hypothesis i.e. that the varieties in monoethnic enclaves should be different than the standard because of the speaker's ethnic background.

On the other hand, Cheshire et al. (2011) have found a connection in London between second-generation immigrants and innovations in the vowel and morphosyntactic system of English. They call the variety they study Multicultural London English to avoid unnecessary labelling of it as a lect or variety. Their

explanation for the emergence of this variety is that a lack of adult native speaker models gives rise to group second-language acquisition. At the same time, as innovations emerge due to group second language acquisition, the second-generation speakers can be shown to have adopted more widespread changes in British English such as /u:/ fronting. The adoption of this widespread ongoing change, together with the fact that other innovations do not have a similar origin together contribute to the impossibility of a single source explanation for the emergence of Multicultural London English (MLE). The language variation is instead seen as constituting a heterogenous *feature pool*, a concept introduced by Mufwene (2002) as a way to provide an explanation for how different features are selected by individual speakers to create a linguistic system, idiolect or larger aggregates, such as languages. Cheshire et al. (2011: 176) argues that the choice of linguistic variants from the heterogenous 'feature pool' is guided by frequency, salience and communicative strategies. According to them, linguistic variants undergoing global or widespread changes are one type of linguistic variant that is highly available for all speakers to select and use.

The fact that immigration influences society and thereby the language is well established. The problem seems to come when claiming that what arises out of the contact is always new varieties. In some cases, there appears to be a fairly firm ground for the claim (Cheshire et al. 2011) while in other cases there seem to be less firm ground for it (Kotsinas 1988-2014). What we need is some kind of definition of what a variety is, and theoretical tools to handle the "difficult" cases. Hudson (1980) defines a variety as *"a set of linguistic items with a similar social distribution"* (1980:24). This is of course not the only definition but it is the one that Boyd (2010) used as a point of departure when trying to determine if research carried out in the SUF-project (cf. 3.1.1 below) pointed in the direction of one or more new varieties of Swedish, or whether the observed differences could be regarded as "normal" social variation within a linguistic system. Her conclusion was that the adolescents' use of the alleged variants was not as frequent or as homogeneous as one would expect when claiming that it constitutes the base for a new variety. However, orderly heterogeneity predicts that there will be conditioning linguistic, contextual or social factors in the use of linguistic variants.

What the conditioning social factors might be, where linguistic ones seem to be missing, differs depending on what quantitative level of abstraction the

investigation chooses to use – more macro or more micro. Efforts to circumvent and not use ethnicity as an established natural category include Hoffman and Walker's (2010) introduction of ethnic orientation (EO), further elaborated in Boyd, Hoffman and Walker (2015) as ethnolinguistic orientation, while more qualitative approaches (Jaspers 2008, Eckert 2008) investigate the variation by providing descriptions of intricate identity work among children, adolescents and young adults. Also, theoretical tools have been suggested to better explain the heterogeneous features used in alleged new varieties, in the form of a "stylistic practice" (Quist 2008), "ethnolinguistic repertoire" (Benor 2010) or "feature pool" (Cheshire et al. 2011). All these are similar ways of coming to terms with the heterogeneous nature of the linguistic practices among young people in multilingual communities. Wolfram (2007) argues in connection to Afro-American English vernacular (AAEV) that the metonymic practice to name an alleged variety has led to a belief that the new ways of speaking are more homogeneous and less varying than actually is the case. In the case of AAEV this has had the effect that regional and social variation has been disregarded in favor of variation contingent on the ethnic category. We can see that a similar process has taken place with the introduction of *multiethnolect,* where focus has been directed toward the non-majority background (such as ethnicity or foreign background) of the speaker, which has masked a more complex situation where a number of social factors interact with foreign background on the macro level.

2.4 Sociophonetics

This thesis is grounded in variationist sociolinguistic theory, but it also makes use of phonetic methods when trying to answer its research questions. The two areas have been conflated under the term sociophonetics, as studies such as this one need to combine phonetic and variationist methods of data collection and analysis (Thomas 2011). As an example, within acoustic phonetics the quality of the recordings is of the highest priority and recordings of reading tasks in sound-proofed rooms are often used for analysis, as this kind of recording has an extremely good sound quality and enables detailed analysis of the acoustic signal. Also, control of the material ensures that the desired number of tokens will be obtained, and that the tokens are the same for every speaker studied. However, within

sociolinguistics, this kind of setting is considered highly artificial and as tending to affect the informant's speech style to be more controlled and careful, i.e. further from the more informal or relaxed speech styles which ought to be the object of study for sociolinguists. At the same time, even the sociolinguist recognizes that if the recording conditions are not controlled in any way, the acoustic signal will be of inferior quality and acoustic analysis will be more difficult, so they often complement more spontaneous speech material with reading passages and word lists. Likewise, for both the phonetician and the quantitative sociolinguist, token production needs to be controlled for so that the data fulfills statistical requirements on sample size. Even though the vernacular might not always be the style the sociolinguist aims to study, the sociolinguist usually aims to create or to record speech in a setting where social and contextual factors that we know from sociolinguistic research will affect the informant's speech style tend to be more relaxed. These issues are discussed further in paper 2 (Forsberg & Gross submitted).

2.5 Variationist research in Sweden

Studies within variationism, sociolinguistic studies of language contact and bilingualism and sociophonetic studies have been carried out in Sweden, as elsewhere in Europe and North America. As one of my goals is to contribute to a description of variation in Swedish, this section will summarize some of the work that has been carried out here which is most relevant to my research questions, the work of Bengt Nordberg, Eva Sundgren and Mats Thelander.

Variationist sociolinguistics in Sweden can be said to have started in the late 1960's with Bengt Nordberg's studies of variation in the spoken language of Eskilstuna (Nordberg 1972, 1975). In a number of studies, he investigated morphological (1972) and phonological (1985) variation and showed how both internal and social factors contributed to language variation and potential change in the town. The material that Nordberg collected during the 1960's has been utilized in a number of subsequent sociolinguistic studies.

Most notable among the follow-up studies is Eva Sundgren's (2002) real time study, where she combined Nordberg's data with a new sample collected 29 years later, sampling data from 17 of the original informants and 72 new informants. In her study Sundgren (2002) investigated the same morphological and morpho-

phonological variables as Nordberg, showing that the assumed change toward a more leveled standard-like speech was slower than expected, and that the variation associated with certain social factors had decreased. In addition, she showed that integration in the local community and social mobility were important factors to consider when analyzing the linguistic behavior of the informants, a finding that Grönberg (2004) also found when investigating the linguistic behavior of adolescents in Alingsås. However, Sundgren (2002) points out that there might be differences regarding how rapidly variables on different linguistic levels (phonological, morphological, syntactic etc.) are changing. Apart from contributing with empirical data to the theoretical understanding of the differences between apparent vs. real time studies and panel vs trend studies, Sundgren (2002) also investigated if there was coherence in the use of the morphological variables. By doing so she could (with some caution) classify the informants into three groups: one for whom use of local forms dominated, one in-between group, who used both local and standard forms and one where the standard forms dominated. She showed thereby that variants were used systematically by different groups of speakers so that they could be described as creating a linguistically coherent system in the city.

The investigation of coherence and division that Sundgren used was also inspired by Thelander (1979a, 1979b) who investigated the use of dialect and standard variants in Bjurträsk. In his study he showed how linguistic variants co-occurred in a systematic way so that three varieties could be distinguished: dialect, regional standard and standard. Switching among the three different speech varieties was dependent on social and situational factors, but the informants tended to use variants in a coherent way so that they switched primarily between two varieties: either dialect and regional standard or standard and regional standard, depending on context. Furthermore, Thelander (1979a:116) showed that the stability in use of a variant was connected to the size of the geographical area that the variable was used in: the larger the area that a dialect variant was found in, the more stable it was. Thelander's Burträsk studies thus contributed significantly to our understanding of dialect leveling, coherence and quantitative approaches to switching among varieties in a limited geographical area.

2.6 Summary

In this chapter, I have sketched the framework of variationist sociolinguistics and introduced some of the key concepts of greatest relevance to my study: speech community, ordered heterogeneity and the vernacular. I have also discussed the concepts of language, dialect and standard language in order to clarify how I use these terms in the rest of the dissertation. Previous work on ethnolects is discussed critically, particularly how the term *multiethnolect* has emerged in sociolinguistic reserach. Relevant sociophonetic research is discussed in section 2.4, including its methodological implications for my work. The chapter is concluded with an account of some Swedish variationist research which is most relevant to my studies.

CHAPTER 3

Data collection and analytical methods

This chapter describes the two corpora that were used, the data selected and the analytical approaches taken to make sense of the data. I will briefly describe the SUF corpus (*Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer* 'Language and language use among adolescents in multilingual urban settings' [Bijvoet et al. 2001]), from which interview data was used for the first paper (Gross et al. 2016) and then focus on the new data that Julia Forsberg and I collected during 5 intense months in the autumn of 2014. The result of the latter data collection is a corpus that we named the SSG-corpus (*Språkbruk i Stockholm och Göteborg* 'Language use in Stockholm and Gothenburg'), which consists of 74.5 hours of audio recordings (both interviews and map-tasks) from 111 informants who have grown up in the two cities (more in section 3.2). However, the SSG-corpus is not the only new data that was collected. During the initial work with segmenting the data I stumbled upon something peculiar: some of the informants seemed to be merging their /i/ and /y/ to [ɨ], which is why Forsberg and I performed a perceptual experiment to gain an understanding of what I had noticed. For the perceptual experiment, we collected classifications from 203 listeners through a web form. To sum up, the material for this thesis consists of; interview recordings from the SUF project, interview and map-task recordings from the SSG corpus, and perceptual data.

When collecting sociolinguistic data, ethnographic observations are often of great value, which has been shown in numerous studies ever since Labov's (1963) groundbreaking work in Martha's Vineyard. However, there are always decisions to make when it comes to what data to gather: more qualitative data which allows us to gain knowledge in the microcosmos of a small group, or more quantitative data, allowing us to gain access to the more abstract language structures of larger groups. It is my firm belief that in order to gain a 'full' understanding of a variety, a combination of both approaches is needed, but that the more abstract quantitative

work should precede the more qualitative work. Since there is a shortage of research on variation and change on vowels in Gothenburg – and on the Gothenburg dialect in general – I decided to start to draw a picture of the more general aspects gained from quantitative analysis, hoping to provide a brick in the foundation for later more in-depth ethnographic studies in the future.

This thesis is not just a description of the Swedish vowel system used among youths in Gothenburg, the initial study (Gross et al. 2016) is a comparison of Stockholm and Gothenburg. The approaches taken in that study and a desire to collect comparable data to the SUF data in both cities were strong contributing factors and a point of departure when we set out to collect the SSG data. However, the time that it would take to segment and analyze the amount of data that was collected for the entire SSG-corpus proved to be too long and I decided to focus on the Gothenburg recordings, which is why the Stockholm recordings in SSG have not been analyzed in this thesis.

3.1 Data

In this section I will briefly introduce the SUF corpus (3.1.1), then move on to the SSG corpus (3.1.2) and finally discuss some relevant differences between the two corpora (3.1.3).

3.1.1 The SUF corpus

This corpus was collected in 2002 in Sweden's three largest cities, Stockholm, Gothenburg and Malmö for the SUF-project. The focus of this project was to examine the situation that Kotsinas in a number of books and papers (Kotsinas 1988a, 1988b/2014, 1992, 1994, 1998, 2001) had described in Rinkeby, Stockholm, but with a larger sample and in similar settings in additional cities. In total, 222 informants from eight different upper secondary schools were interviewed and recorded in a number of situations, such as retelling tasks and group discussions. The sampling was based on a judgement sample where the relevant speakers were identified and schools where they could be found contacted (Ganuza 2008). Judgment samples were chosen in favor of the more statistically kosher random

sample. This type of sampling technique is usually preferred in quantitative sociolinguistic studies, since a random sample requires a larger sample to ensure that all background factors are represented. A large sample tends to result in a data-handling problem where the amount of linguistic data collected will be too large to handle (Milroy and Gordon 2003).

The recordings of interviews from Stockholm and Gothenburg in the SUF corpus were used in study 1 of the thesis. The length of these ranged between 20-40 minutes (depending on the informant's willingness to speak) and the content included the informant's social situation, daily life, family and friends, language abilities, reading habits, music tastes, hobbies, plans for the future and so on. I listened to and annotated all of the 119 recordings from Gothenburg and Stockholm in search of /ɛ:/ in non-pre-rhotic as well as pre-rhotic position. Out of these in total 119 informants, only 57 informants met the number of tokens criterion (see Gross et al. 2016 p. 234, for this and other criteria). At the beginning of the study, the intention was also to investigate /ø:/ as well /ɛ:/, since other studies has shown that a similar pattern of variation occurs for the /ø:/ vowel in other parts of the country (Nordberg 1975, Leinonen 2010). However, the /ø:/ vowel in non-pre-rhotic as well as pre-rhotic position were not produced often enough to be included in the study.

3.1.2 The SSG corpus

This corpus was collected at four different schools, two in Stockholm and two in Gothenburg. The sampling method was inspired by the method employed in the SUF project, i.e. it was a judgment sample. However, one of the upper secondary schools in Gothenburg used in the SUF project had closed since 2002 and we therefore opted for an inner-city school which resembled the school used in SUF but was somewhat more similar to the inner-city school in Stockholm, in order to enable a more straight-forward comparison between the two cities. In Stockholm, it turned out that none of the three schools that were used in the SUF project wanted to participate again, and we had to find two completely new schools, one in the suburbs and one in central Stockholm. To be included in the new SSG corpus, informants had to have received schooling in the relevant city since before the third grade (at the latest 8 years of age). At the time of the recording the informants' ages

ranged from 16-20 years with a mean of 17.5 years. This choice was made to control at least to some degree for the dialectal experience assumed to be acquired within a critical period (Werker & Hensch 2015). Other than this criterion, any pupil from any educational program was welcome to participate, and the informants came from a wide range of different programs. The removal of the program and the age criteria are two aspects where the SUF sample differs from the SSG one, as the former project had no strict age criterion, but all the informants came from the social science program. Furthermore, by removing the program criterion the informants were able more freely to choose their partner in the map-task part (more about this in 4.2).

Two ways of collecting data were used: an interview and a map-task. The interview was inspired by the interviews in the SUF-project, which can be described as short sociolinguistic interviews in the sense that the interview was structured into conversational modules focusing on the informant's social life and language use. The limited length of the shortened interview was probably one of the reasons why the informants did not produce enough tokens in the SUF interviews. However, the general time length that is opted for in a sociolinguistic interview, about 1-2 hours (Labov 1984), was shortened to about 30 minutes since it would be almost impossible to find timeslots of 2 hours (or more) for each of the pupils during school hours. The small token count in the SUF-interviews was a problem that Forsberg and I had to deal with when planning the SSG project. Our solution was to create a map-task where the words depicted on the maps had long vowels (or consonants in the relevant environments for Forsberg's research). The map-task was the main tool for our linguistic data collection, and the interviews were used as a way of gaining demographic and metalinguistic data, as well as information about the informants' daily lives and educational plans.

Before contacting schools, the SSG project was approved by the Regional ethical review board of Gothenburg (registration number 465-14). After this, schools were contacted by e-mail to the principal who declined or gave the initial permission for us to contact teachers and be present in the school. Our first contact with the informants was usually in class where we gave a short introduction about the project, both verbally and in written form. After this, the informants who were interested in participating and met the school attendance criterion were invited to sign consent forms and provide their name and phone number in given timeslots.

All informants who participated were given a cinema voucher as thanks for their participation. Informants were also recruited in the hallways and by informants asking their friends if they wanted to participate. These pupils also received written information and signed consent forms. In total, 111 interviews (a total of 36.5 hours) and 55 map-tasks – each with two participants – (a total of 38 hours) were recorded. One informant dropped out after the first part of the map-task recording and one informant was used as a conversational partner who did not meet the school attendance criterion that we had set up for the informants. Both were excluded from the analysis. This is why the number of interviews is not exactly twice the number of map tasks. The distribution of the informant gender and mother's[4] birthplaces are shown in Table. 1 Kripke (inner city) and Frye (suburb) are the pseudonyms for the two schools in Stockholm, and Descartes (inner city) and Hume (suburb) are the pseudonyms for the two schools in Gothenburg. Studying the distribution between the inner-city schools and the suburban schools, a first thing to notice is that the distribution between those with a mother born in Sweden and those with a mother born abroad is more even in the inner-city schools than in the suburban ones. This imbalance is most striking in Hume, the suburban school in Gothenburg where we had to abandon the goal of obtaining an even distribution between informants with Swedish and foreign-born parents, as so few such pupils attended the school.

	Male		Female		
Mother's birthplace	Sweden	Abroad	Sweden	Abroad	Total
Kripke Sthlm	10	5	11	7	33
Frye Sthlm	2	10	10	4	26
Descartes Gbg	8	7	6	7	28
Hume Gbg	0	9	1	14	24
Total	20	31	28	32	**111**

Table 1. Distribution of the informants in the SSG corpus.

[4] Mothers birthplaces is chosen instead of parents since this was the factor used to operationalize foreign background in Gross et al. (2016). However as is apparent in Gross (in press) using both parents place of birth gives rise to a third category "both" e.g. on parent with a Swedish background and one with a foreign.

3.1.3 Similarities and differences between SUF and SSG

The two corpora share a number of similarities but they are also different. Both are based on a judgment sample, a sampling method that has been favored in quantitative sociolinguistics since it solves the methodological problems arising with handling large samples of linguistic data. All sampling methods must consider three decisions when planning data collection: the sample universe, the relevant dimension of social variation and sample size (Sankoff 1980). As mentioned above, the SUF corpus based its sample universe on Kotsinas' (1988a, 1988b/2014, 1992, 1994, 1998, 2001) findings and identified multilingual urban settings (usually suburbs) as the sample universe. Informants were then asked to participate by contacting schools in the designated areas. The SSG corpus was to a large degree inspired by SUF but also by the findings and argumentation in Gross et.al (2016), which explains why one inner-city school and one suburban school were chosen in Gothenburg instead of two suburban schools as in SUF. The dimensions of variation in SUF were also determined on the basis of Kotsinas' findings: residential area, immigrant vs. native population and age (youth), but a number of different situational contexts were included e.g. interviews, language tasks, written school assignments, spontaneous speech (self-recordings) and group discussions. In SSG, residential area, immigrant population and age were still used as points of departure. The situational contexts, though, were limited to interviews (similar to SUF) and map-tasks (not used in SUF). The differences in situational data stem from the differences in focus and the number of people collecting data.

3.2 Methods

A number of different methods have been used to collect, prepare and analyze the data in the four papers. Most of the data analyzed consists of phonetic acoustic measurements. Two different techniques were used to elicit this data, as discussed in 3.2.1. Before the acoustic data from different individuals can be compared it needs to be *normalized*. What it means to normalize acoustic data from vowels and different algorithms used for this is discussed in 3.2.2. For the fourth paper, investigating a possible merger of /i:/ and /y:/, both acoustic and perceptual data were used. The collection of perceptual data is discussed in 3.3 and finally in 3.4

the method used for analyzing the interaction in the map-task recordings is discussed.

3.2.1 *Acoustic analyses*

Acoustic analysis was performed on the recorded data from both of the aforementioned corpora. However, different methods were used to elicit the relevant acoustic parameters correlated with the position of the tongue body and the degree of opening of the mouth. In the first study, where the data was taken from the SUF corpus, the measurements were performed manually using the linear predictive coding (LPC) algorithm in Praat (Boersma & Weenink, 2016) while the data from the second study was analyzed with a whole spectrum method.

The first method mentioned is often referred to as a formant-based method and it identifies broad peaks in the acoustic vowel spectrum by calculating a set of predictors which predicts the amplitude of the peaks in the waveform. These predicted peaks correspond to the resonance frequencies of the vocal tract (Johnson 2011). The method has been shown to have quite large error rates when automated: Eklund and Traunmüller (1997) found that about a quarter of their measurements had to be manually corrected, and Adank, et al. (2004) obtained similar results analyzing a different dataset. This large error rate stems from the fact that the number of peaks, or coefficients, needs to be specified in advance for the algorithm to be able to calculate the predictors (Harrison 2013). In Praat, the number of coefficients is set by defining the number of formants in formant settings. Harrison (2013) showed that this specification of coefficients made the method sensitive to measurement errors if the specifications were not checked before each new speaker and before measuring each new token, so that the predictions made by the algorithm corresponded with visual observations made by inspecting the spectrogram and the LPC-predicts which are 'printed' with red dots on top of the spectrogram (Figure 1). Similar results were found by Derdemezis, et al. (2016) who, when examining the default parameter in a number of phonetic software programs, ended up recommending that the parameters should be checked and manipulated if necessary before every new informant and depending on the vowels and formants of interest.

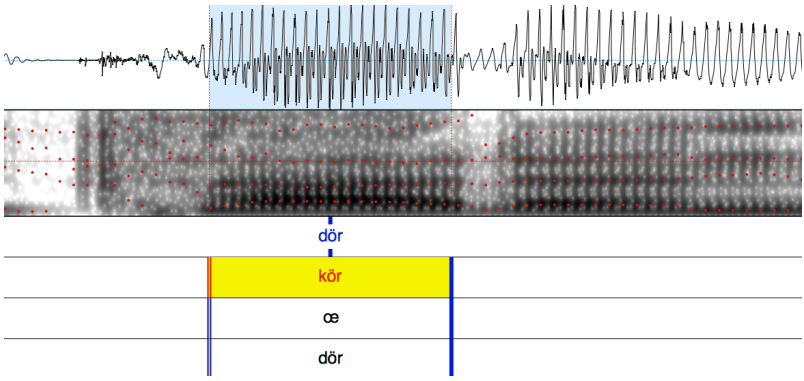


Figure 1. An example of the coding and how the vowels were segmented in Praat prior to acoustic analysis. The red dots are the LPC-predictions of the formants. The horizontal black/ dark streaks are amplitude peaks in the spectrogram which indicate the formants; the first formant is the lowest. The coding in the tiers indicates from top to bottom: center of the segment indexed with type word, the word, the impressionistic judgment of the vowel quality and on the lowest tier phoneme/ allophone indicated with type words.

Also, the algorithm has been shown to have difficulties finding the second formant in high back vowels such as [u] since F1 and F2 are closer to each other, which is why extra attention needs to be paid when measuring high back vowels (Di Paolo et. al 2011). Nevertheless, the method is attractive in that the values are given in Hz that can be easily understood. The method is also easy to use and incorporated in a number of software packages and as long as the analyst has a knowledge of the algorithm and is careful in correcting the number of coefficients before doing each measurement the values can be trusted and fine differences can be detected. Manual measurements are, however, time-consuming and errors might be introduced due to the repetitious work better suited for computers than humans.

All of the acoustic measurements in Gross et. al (2016) were done manually in Praat, the LPC estimates were controlled against the spectrogram, and the number of formants was then adjusted if necessary so that the LPC estimates plotted over the spectrogram matched the dark horizontal parts in the spectrogram indicating increased amplitude in the sound spectrum (see figure 1). In cases where these LPC estimates were hard to control through this kind of spectrogram inspection, a Fast Fourier Transform (FFT) spectral slice was produced and the formants were measured this way. This way of analyzing the formants was, however, extremely time-consuming and would have been impossible to use when

measuring and analyzing all of the vowels in Gross (in press) and Gross & Forsberg (submitted). Instead, I opted for a whole-spectrum method which has been shown to be more reliable when automated (Jacobi 2009).

The whole spectrum method used in Gross (in press) and in Gross & Forsberg (submitted) is to a large extent the same as in Leinonen (2010) and can be described as a two-step procedure: first, acoustic analysis with band-pass filters, and then reducing the output with statistical methods. This way of analyzing acoustic data was introduced by Plomp et al. (1967) who showed that band pass filtered data could be analyzed to meaningful components with the aid of principal component analysis (PCA). A band-pass filter is an acoustic filter that allows acoustic energy within a certain frequency range to pass through. The acoustic energy that remains after filtering is called a pass-band. PCA is a statistical method of reducing a large set of data to a set of components that explain the variance in the data set.

The band pass filter method was later refined for acoustic vowel analysis by Jacobi (2009). She showed that by applying Bark filters (a type of band-pass filters where the defining scale is in Bark), where each pass-band has a bandwidth of one Bark overlapping with -3dB and a cutoff point at 18 Bark, will capture the frequency range where the three first formants will be found. Leinonen (2010) improved the second step by showing that the best correlation between formants and PC's was found when two PC's were calculated and then rotated with a varimax rotation. However, Leinonen (2010: 76) points out that although the correlation is high, there is no one-to-one match between formants and components i.e. F1, F2 and PC1, PC2. PC1, which should explain most of the variance, also captures some variation caused by F2 and F3, whereas PC2 only captures variation connected to F2. In other words, the whole spectrum method can be reliably automatized but captures rather large frequency regions that characterize the vowel quality. On the other hand, LPC cannot be reliably automatized (without careful hand corrections) but manages to capture more fine-grained differences in the vowel quality.

To summarize, in the first study (Gross et al. 2016) LPC was applied manually since the number of vowels was possible to segment and measure manually. However, in Gross (in press) and Gross & Forsberg (submitted) an automatized method was used, where the first step was an acoustic analysis performed in Praat with Bark filters, and the second step was to reduce the data

from the acoustic analysis to articulatory meaningful components with principal component analysis performed in R (R Core Team, 2017 version 3.4.2).

3.2.2 Normalization

No matter the statistical approach taken to analyze the acoustic data, the problem of making it comparable between individuals remains. This difference between speakers arises from the fact that the size of the vocal tract differs between people, and especially so between men and women. This is solved by normalizing the data i.e. by mathematical means eliminating the physiological differences while preserving phonological distinctions and social/dialectal variation. Normalization can also be a way of modeling the cognitive processes that humans employ when associating vowel productions by different speakers with the same vowel category (Disner 1980, Thomas 2002). A number of algorithms have been suggested to accommodate the criteria. These normalization algorithms can be typologically classified as to whether the information they take as input is *vowel intrinsic* or *extrinsic* and *formant intrinsic* or *extrinsic*; this division gives us four types of algorithm (Watt et al. 2011). A vowel intrinsic method only considers a single vowel whereas an extrinsic one uses the information from many vowels assumed to be in some kind of linguistic system; a formant intrinsic one uses the information from one formant *x* (e.g. all the measurements of F1) at a time whereas the formant extrinsic considers two or more formants at a time.

The vowel and formant intrinsic algorithms are scale transformations e.g. from Hz to Mel or Bark. These transformations have been shown to be quite bad at reducing anatomical differences but can be used as a step in a normalization process. An example of a vowel intrinsic formant extrinsic method is that of Syrdal and Gopal (1986) who transformed their data by comparing the distance between adjacent formants transformed to Bark. The vowel intrinsic algorithms are usually used to model some aspect of the human perception of vowels, as in Syrdal and Gopal (1986). The vowel extrinsic methods on the other hand were developed for classifying vowel tokens with automated speech recognition (Adank et al. 2004). An example of a vowel and formant extrinsic method is Nordström and Lindblom (1975). They calculated the mean from each individual's F3 for low vowels and

used this mean value as a scaling parameter. Lastly, two examples of vowel extrinsic formant intrinsic methods are Nearey (1978) and Lobanov (1971). These two methods were shown to perform best when normalizing acoustic vowel measurements for the purpose of variationist studies, when 12 different methods were tested (Adank et al. 2004). In (Gross et.al 2016) the Neary algorithm was used and in Gross and Forsberg (submitted) and Gross (in press) the Lobanov procedure was used. Below I explain more about these two algorithms.

The Lobanov (1971) algorithm is a z-score transformation i.e. the measurements from a group of informants are all transformed into the same scale by, for each speaker, subtracting formant (*n*) from vowel (*v*) with the mean of formant *n (*$MEAN_n$*)* and then dividing it with the standard deviation of formant *n* (S_n) as in equation (1)[5].

(1) $F_{n[V]}^{N} = (F_{n[V]} - MEAN_n)/S_n$

The Lobanov method was marginally better or tied with the Nearey algorithm in the different tests that were performed by Adank et al. (2004) to compare the methods. However, the Nearey algorithm was shown by Labov (2001b) to do an excellent job preserving sociolinguistic data and Disner (1980) showed that the Neary method performed better than Lobanov in reducing scatter when normalizing for cross-language comparison. The Nearey algorithm is also sometimes referred to as the log-mean method since it normalizes by using the logarithmic mean, as in (2).

(2) $F^{*}_{n[V]} = \text{anti-log}(\log(F_{n[V]}) - \text{mean}(\log(F_n))$

Here, the log-transformed value of formant *n* from vowel *v* for each speaker is subtracted with the log transformed mean from all the formants from that single speaker *mean(log(*F_n*))*. Both methods have similar disadvantages and work best when all of the vowels in a vowel system are incorporated; otherwise the vowel space can get skewed. Also, if data from different linguistic systems are used, the

[5] equation (1) and (2) are taken from the NORM webpage created by Thomas, Erik R. and Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite. [Online Resource: http://lingtools.uoregon.edu/norm/index.php]

normalization might skew the output since they are both making use of the mean of all vowels, a procedure which might make the typological differences smaller than they really are (Disner 1980). Both of the methods transform the formant values to a new scale and the new values cannot be related back to Hz or Bark values without a new scaling transformation. However, the values can easily be plotted and the shape of the vowel space is preserved rather well with both algorithms (Watt et al. 2011). In the end, what is most important when choosing a normalization procedure is to understand the limitations of the data and to decide on the method which suits it best and distorts the data the least (Thomas 2002: 174).

3.2.3 Perception experiment

In Gross and Forsberg (submitted) perception tests were made as a means to gain an understanding of observations during the initial work with segmenting and annotating the vowels in the Gothenburg part of the SSG corpus. During this work, I noticed that the informants in the inner-city school, both with foreign and Swedish born parents, that did not live in the northeastern suburbs seemed to merge their /i:/ and /y:/ vowels into /i:/. This group is a rather homogeneous group with a high Swedish proficiency. In addition, the merger has never been reported on in Gothenburg which led us to wonder whether this might be a change in its early stages and well below the level of conscious manipulation of the speakers (Labov 1994: 78). Two conditions are often given as as backgrounds to the merger of two phonemes: 1) contact between varieties (languages [Herold 1990] or dialects [Trudgill & Foxcroft 1978]), and 2) language internal factors such as crowding or weakness of some articulatory feature. Thus, the perception experiment was designed to investigate the language internal factors that might be in play.

The experiment was designed as a segmental identification task. Normally in such tasks the categories or labels are provided. However, we were afraid that phonological categories might interfere in the identification and did not want to draw attention to the phonemes under investigation so we decided to opt for an open-set format where the informants had to write the grapheme that represented the sound that they heard. Productions from the map-tasks in SSG of /i:/ and /y:/ and some fillers were then selected in a two-step procedure: in the first step we picked out all the productions that we could find that had a reasonably good quality;

in the second step the selected tokens were transcribed by both of us independently and the tokens that were judged (by both of us) as good enough quality to be played in isolation to naïve speakers were chosen. Each production was then put in a web form where they occurred in a randomly scrambled order two times, once in the whole word where they occurred and once in isolation. The informants could listen to the word or segment as many times as they wished but as soon as they answered and had pressed "continue" they could not go back and correct their answers. In this way, the listeners carried out a discrimination task without being aware of it and we could be fairly sure that the listeners' attention was not drawn to the merged vowels as such.

3.2.4 Analysis of the interaction in the map-tasks

Gross (in press) and Gross and Forsberg (submitted) made use of data collected with the help of a map-task. The interaction between the informants (and researcher) when solving this task was analyzed in Forsberg and Gross (unpublished), in order to gain an understanding of the nature of the data this method generated. The initial stage in this analysis was data-driven and Forsberg and I tried to find patterns and things that captured our interest. Systematically listening to all of the recordings, taking notes and discussing different aspects of the participants' interaction we asked questions such as: How do the informants start the interaction? How do we as researchers interact and act? What comments do they make in between maps? and How do they solve obstacles in the maps?

The first goal, at this point, was to find and transcribe short sequences that could be used during a data session where people from different departments at the Faculty of Arts participated and provided their perspectives and thoughts about the interaction. The process of choosing what part to transcribe and how to transcribe them are two crucial points. Cibulka (2016: 18) discusses these two aspects of 'the activity of transcribing' – the analysis and the presentation – and argues that the two could be seen as two separate aspects or layers of transcription which can very well be manifested in different kinds of transcripts, since the transcript best suited for analysis might be less good when presenting the data for an audience.

*H28: hittar du den
*H27: nä men a: jag vet vicken d e
(1.2)
*H28: yani har du ingen bil >ens elle män<
*H27: nä: °d° asså ja sa ju till dig förut hon sa d e två olika nästan
helt olika e:: nästan helt samma men (0.6)
[eller hur
*H28: [aha: aha oj [ja: märkte inte de
*H27: [a
*H27: a
*H28: okej alla fall (0.7) e::# (0.9) du vet den här musikgrejen

Example 1. The transcription used in the data session, transcribed in in CLAN.

The first point of choosing the parts to transcribe is crucial to how the data will be presented to an audience, in our case first at the data session and then later in publications. However, the first audience is part of the analytical process and part of producing results and the second audience will 'only' receive the results. The orientation towards two (slightly different) audiences highlights the fact that the transcripts themselves are tools for analyzing and presenting the data, and not data per se. For the data sessions, more detailed transcripts were used (Example 1), transcribed using CLAN (MacWhinney 2000) This software is useful when the analyst wants to capture and analyze in detail the sequentiality and turn-taking in conversation and annotate the length of pauses and prosodic features. In the paper, less detailed transcripts were used (Example 2). At the point when we were writing up the results, the analysis had led us in the direction of Bell's theory of audience design (Bell 1984) and we realized that the more detailed transcripts would only obfuscate our purposes and conclusions. Although the transcript still captured the turn-taking, other aspects such as prosodic features (e.g. speech rate, intonation, prolonging of segments and so on), overlap in turn taking and exact length of pauses were removed in favor of other aspects that we wanted to highlight.

1. *H28: hittar du den
2. *H27: nä men a jag vet vicken d e
3. *H28: yani har du ingen bil (ens elle män)
4. *H27: nä: asså ja sa ju till dig förut hon sa d e två olika nästan helt olika e nästan helt samma men eller hur
5. *H28: aha aha oj ja märkte inte de
6. *H27: a
7. *H27: a
8. *H28: okej alla fall e du vet den här musikgrejen

Example 2. The same transcription as in Example 1. but adjusted for the purpose of the map-task paper. See example 5 in (Forsberg & Gross unpublished) for translation.

To sum up, the initial analysis relied on a more detailed transcription where we tried to capture as much of the intricate nature of the interaction as possible since we at that point did not want to impose too much of our own assumptions about what kind of features might be of relevance. When we had a clearer picture of what in the interaction was of relevance for the paper we adapted the transcript to capture and highlight these aspects. We are well aware that a transcript is not the data but only one of many possible representations of the interaction.

3.3 summary of chapter 3

In this chapter I have discussed the two corpora from where the acoustic material is taken. The first corpus is the SUF corpus which was collected in 2002 with the general aim to investigate the claims of Kotsinas (1988a, 1988b/2014, 1992, 1994, 1998, 2001). The second corpus was collected by myself and Julia Forsberg, in order to get a similar material as the SUF corpus though adjusted for our purposes. There are similarities between the two corpora such as one of the schools in Gothenburg, the sampling method, the interviewees, the recording location within the school to name the more prominent. However, there are also differences such as the fact that we did not limit the sample to pupils from one educational program, an age of arrival criterion for participator was deployed and the use of a map-task to ensure that enough high-quality tokens were produced.

The tokens from the two corpora where then analyzed acoustically and the two first formants where measured. The material from the SUF corpus was analyzed by employing a formant-based method (LPC) where all the tokens where measured manually and the LPC coefficients where adjusted to ensure that each vowel was correctly analyzed. When analyzing the material from the SSG corpus a whole spectrum method was used since this has been shown to be more reliable when automated. This type of acoustic analysis is done in two steps: first an acoustic filter is applied so that certain frequencies are filtered out; second the data that is left after the filtration is analyzed by the means of a PCA, where PC1 and PC2 have been shown to capture the same variation as F1 and F2, though PC1 also captures some of the variation of F2. The output from the acoustic analysis was then normalized to enable comparison between speakers. Two different vowel extrinsic formant intrinsic methods were used: the SUF data was normalized with the Neary

1 algorithm and the SSG data was normalized with the Lobanov algorithm. Though no normalization algorithm is perfect these two have been shown in a number of studies to be among the most reliable.

Apart from the acoustic analysis, perception experiments were used to gain an understanding of why some informants seemed to merged /i:/ and /y:/ to /i:/. This perception experiment was a segmental identification task distributed through a webform. The segments were taken from the map task recordings. Each segment was classified two times: one time in the produced word and one time in isolation without the listener knowing that this was the case. In this way we managed to see if the if a possible top down process obscured the merger and to obtain classifications of the merged and unmerged forms.

Lastly, the analysis of the interaction between the informants when solving the MT was done through careful study of the interaction with the help of transcriptions and data sessions. The transcriptions used during the analytical process differ from those presented in the final paper since some aspects in the original transcriptions obfuscated aspects of the interaction relevant for our analysis.

CHAPTER 4

Summarizing the studies

This chapter provides short descriptions of the papers which are included in this thesis, the main result from each one and some of the more important motivating factors behind them. The papers follow their own internal argumentation and as such they stand on their own. However, they have in part inspired each other and in this way are part of a single research process. Because of relationships among the papers I will show how they connect to each other and what in the earlier studies gave rise to the more recent ones. By doing so I hope to clarify the thinking behind the papers and assist in the reading of them. Three of the papers are coauthored and one is single-authored, and the summaries of the co-authored papers will end with a short section where I clarify my contribution to each paper.

4.1 The first study, Gross et al. (2016): "A tale of two cities (and one vowel): Sociolinguistic variation in Swedish"

This study approaches the language variation and change described by Kotsinas (1988a, 1988b/2014, 1992, 1994, 1998, 2001). But instead of looking at the innovative features, such as loan words and diversions from the V2 rule that Kotsinas described, this study looks at variation which can be related to a traditional dialect variable in Stockholm, often called *Stockholm-e* but which has also been described as a more general change in the Swedish vowel system. The older dialect variable is a reported merger between /e:/ and /ɛ:/ as [e:]. However, the merger was never completed since the contrast was kept in phonological environments where the vowel preceded a rhotic consonant. The variable became a *stereotype* associated with the traditional dialect of Stockholm. When Kotsinas studied the variable (1994) she found that the merged forms were losing ground in favor of the more standard-like [ɛ:], and that a more open allophone [æ:] was making an entry. Similar evidence of a more open variant of /ɛ:/ emerged in Nordberg (1975) and Leinonen (2010), i.e. in other parts of Sweden as well.

Inspired by findings in London (Cheshire et al. 2011) and Toronto (Hoffman 2010, Hoffman & Walker 2010), where widespread changes were found to diffuse to multilingual speech groups, we, in this study, set out to investigate if the change of /ɛ:/ also diffused in a similar fashion in Stockholm and Gothenburg. Gothenburg and Stockholm are both located in the same dialect area (Elert 1994) but have two quite distinct dialects and the vowel-change could be in different stages or take different directions in the two cities. The first research question of this study was therefore formulated as "[a]re there differences between speakers in Stockholm and those in Gothenburg as regards this variable?". This question was necessary to be able to disentangle the results in the two cities and uncover the variation within them. When this question was answered we could turn to the second question, the one that the whole paper started with: "[d]o young people with foreign-born parents behave differently from those with Swedish-born parents regarding this variable?". As we know that the change is widespread in the central Swedish area it can be seen as a change in the phonology of the Swedish system and as such the diffusion ought to be connected with dialect contact. However; the youths with foreign backgrounds might not participate in the change to the same degree if language contact is interfering. If this is the case we could expect that the youths with Swedish backgrounds behave differently from the youths with foreign backgrounds. The last question asked "[a]re the effects of foreign background, sex, and linguistic context on this variable the same or different in the two cities?", tries to address the question of the sociolinguistic patterning, e.g. even if the change has advanced to different degrees in the two cities (question 1) it might be the case that there are similarities in how the change is distributed socially in the two cities.

The results show that there are interesting differences between the pattern of variation in Stockholm and Gothenburg. The difference that we found was that the informants in Gothenburg with a Swedish-born mother are clearly keeping the allophones apart, whereas in Stockholm all of the informants have a rather small distance, if any, between the allophones and are behaving more uniformly. Based on this difference we decided to separate the two cities and answer the second question by testing the persons with foreign born mothers against the persons with Swedish born mothers one city at the time. Here, we found that there was no significant difference in the variation between the groups in Stockholm and that the

change was moving in the direction of a merger of the two allophones ending in a more open vowel, though a small contrast still exists between the vowels. In Gothenburg, on the other hand, the sociolinguistic patterning was somewhat more complicated. The informants with a foreign background tended to behave similarly to the Stockholm informants (both with and without foreign background) and showed a high degree of open merged allophones. However, in the Swedish-background group the informants could be separated into two groups based on sex, whereby the girls showed a tendency to behave more in line with the standard description, i.e. separate allophones pre-rhotically as opposed to elsewhere. However, the variation in this group is large; some of the girls had more open merged forms. For their part, the boys were more dialectal in the sense that they tended to have a more closed realization of both of the allophones though still making a difference between them. Since the dialect has more closed forms of /ɛ:/ and /ø:/ and are preserving the pre- rhotic contrast the influence from dialectal speakers might be slowing down the change, though not stopping it.

As previous research investigating multilingual urban areas has focused on describing new and innovative features occurring among groups of speakers in these areas, the picture has been skewed in favor of language contact explanations of variation in multilingual urban settings and groups. However, in modern day cities the picture is more complicated and needs to be balanced. This study shows how this can be accomplished by considering how the complex nature of urbanization and ethnic diversity interact and, as a consequence, how processes of language and dialect contact must be viewed together, rather than as separate isolated processes.

4.1.1 My contribution to the paper

This paper was co-authored with Sally Boyd, Therese Leinonen and James A. Walker. The structure of the paper was outlined together during our initial meetings. However, as first author of the paper I was the one responsible for directing and organizing the paper when writing the results. I wrote most of the sections and delegated smaller sections where I felt that the other authors could contribute. However, no part has managed to escape my touch or vice versa. All of the measurements, the statistical testing (and figures) were carried out by me under the

close supervision of Leinonen. The text got its final structure when I visited Walker in Toronto and he helped give the text a consistent stylistic voice. I was responsible for rewriting and editing the text after getting the paper accepted and receiving the editors' and anonymous reviewers' comments, all the while discussing how to treat the comments with all of the co-authors.

4.2 The second study, Forsberg & Gross (unpublished): "You change your speech depending on who you talk to, but I didn't change much": The map-task viewed through the lens of audience design

One indirect result of the first study was that the interviews from the SUF data turned out not to provide a guarantee that the informants produced enough tokens of the vowels of interest. Therefore, when planning the new project and its corpus, different ideas (longer interviews, word lists etc.) were discussed. However, we did not want to rely on written prompts and wanted the language production to be produced in conversation, ideally between friends. One of the ways used in other linguistic disciplines, often with a more experimental focus, is the use of map-tasks, which we ultimately decided on (the maps used can be found in the appendix). However, the method is rarely used in sociolinguistic data collection, although it has the potential to provide a way to stimulate discussion between peers and at the same time control and increase token production, without any written prompts. After finishing the data collection and observing the informants in solving the task we felt like the task was successful in creating a situation where informants sat down and talked to each other in a relaxed though friendly manner.

Data collection is one of the core issues in sociolinguistics and methods of collecting speech data are often discussed in textbooks on sociolinguistics, though rarely analyzed in journal articles. Since this method is uncommon in sociolinguistics we felt the need to analyze and make sense of the method as a way of understanding how it can be part of the tool box of data collection methods employed in sociolinguistics, as well as what kind of data that can be expected from using the method. For these purposes we analyzed the different participant roles from an audience design perspective (Bell 1984, 2001) and the four *non-audience* components associated with a speech event perspective: genre; purpose/function; topic; and setting (Saville-Troike 2003). In addition to these analyses, we made

short phone interviews with seven of the informants to obtain their views of the activity.

The results from both the phone interviews and the analyses of the interaction showed that the best way of characterizing the *genre* is as a game where direction-giving is a central part. This result led us to split the *purpose* of the speech event between the aims of the informants and the aims of the researchers, since the informants primarily tried to solve a game-like task whereas the researchers wanted to obtain recordings with a large number of high quality tokens of long vowels and certain consonants for acoustic analysis. The analysis of *topic* showed that it can be divided into three sub-components: direction-giving and receiving; establishing, maintaining and updating common ground; and meta-communication around the task. The participant roles in the map task can be categorized as 1st person speakers, 2nd person addressees and auditors. 1st person speakers and 2nd person addressees are generally the two informants and the roles were switched during the speech event depending on the participants' requirements for solving the task. The researcher, on the other hand, is best characterized as an *auditor*. However, it is important for the researcher to understand that during the map-task, the 1st person speakers and 2nd person addressees are the ones shaping the rules of interaction and thus controlling the researcher's role as either an active speaker, auditor or overhearer. The phone interviews showed that the informants had quite good intuitions regarding the purpose of the project, that they were aware of changing how they speak depending on interlocutor and that the map-task more often (but not always) was considered to be less formal than the interview they also participated in prior to the map task.

To sum up, the map-task is a situation where interaction is triggered by the researchers' introduction of a task that sets the basis for the informant's common goal. To accomplish this common goal the informants will have to shape their own rules of interaction and decide on the roles of the different participants. From the researchers' perspective, a number of components need to be considered and manipulated beforehand for the data collection to be successful. The map task, if carefully and successfully planned and carried out, can be a useful component of the tool box used to collect sociolinguistic data. As the method provides a good way of controlling more or less spontaneous peer interaction so that a large number of

high quality tokens can be collected, it makes the use of word list or other 'written-prompt' methods superfluous.

4.2.1 My contribution to the paper

This paper was co-authored with Julia Forsberg, who is the first author, and both of us were responsible for creating the maps and recording the map task interactions. The work with transcription and analysis has been split more or less equally between the two of us and most of the writing has been done together. However, Forsberg as first author was responsible for structuring the text and writing the main part of the section about previous use of map-tasks in linguistic research.

4.3 The third study, Gross (in press): "Segregated vowels: language variation and dialect features among Gothenburg youth"

With the insights from the first study, that language and dialect contact should be considered together, I show in this single-authored paper that ethnic background and socioeconomic status interact and manifest themselves through housing segregation; this is something that affects the linguistic situation in Gothenburg. Another insight from the first study was that there was no clear picture of the entire vowel system in the Gothenburg dialect, which made it harder to understand how different vowels might affect one another. A possibility to address this problem was provided by using the map-task, which gave us more vowel tokens and types than the interviews, and a picture of the vowel system as a whole could be painted. So instead of just describing the social patterning of a single variable, this paper uses *coherence* between two vowel pairs, i.e. the systematic covariation between /i:/-/y:/ and /ɛ:/-/ø:/, as a point of departure. Furthermore, the study shows, through a data driven process where the vowel measurements are analyzed by means of K-mean clustering, that the speakers with the least within-group variation are associated with the neighborhoods where they live; in the northeastern suburbs (NE) or not (non-NE). This result shows that the 'strongest' independent variable is associated with housing or neighborhood rather than parent's birthplace or socioeconomic status alone, since we in both 'neighborhood' groups find informants with a foreign background. Moving on from this finding, I use the groups NE and non-NE as independent variables when testing 9 out of 12 vowels. Among these, two vowels;

/i:/ and /y:/, are significantly different in both openness and advancement and four are significantly different in advancement; /ɛ:/, /ʉ:/ and the pre- rhotic allophones of /ɛ:/ and /ø:/. The difference found for /i:/ and /y:/ can be explained by the fact that the two vowels both have a coherent centralized position in the traditional Gothenburg dialect. However, the NE informants do not show coherence between the two vowels and have a fronted /i:/ but a centralized /y:/. For two of the other vowel pairs, /ɛ:/ and /ø:/, a similar pattern of coherence is present in many of the central Swedish dialects, such that the vowels have overlapping places of articulation and both are becoming more open and retracted before rhotic consonants. As shown in the first study (which was based on data collected in 2002), the /ɛ:/ vowel is undergoing a change and becoming more open, and the allophonic rule is disappearing among youths with foreign-born mothers in Gothenburg. In that study, it was assumed that the informants with foreign-born mothers were leading the change and that the other groups were lagging. This assumption is corroborated in the third study, by the fact that there is no distinction in openness of /ɛ:/ between NE and non-NE informants. However, the NE group is still making a distinction in advancement, indicating that all of the informants are affected by the change but conforming to it at different paces.

The study shows that it is not ethnicity per se that affects language variation, but rather a complex interaction of socioeconomic status, inequality in the opportunity to choose where to live, and foreign-background. These three parameters result in segregated suburbs where people with a low socio-economic status are grouped together due to not having the same opportunities to move away as those in a higher socio-economic status group do. In other words, part of the linguistic situation can be explained by the density of communication between people who speak the traditional Gothenburg dialect and the people in the northeastern suburbs who have less exposure to this dialect. This is not to say that NE informants do not use linguistic features associated with the dialect. They do, but rather that the coherence between /i:/ and /y:/ is not present to the same degree among the informants in the NE group. Just as with /i:/ and /y:/, the lag in the widespread change of /ɛ:/ and /ø:/ in the non-NE group can be assumed to connect to more contact with the traditional dialect.

By no means is this study an attempt to argue against the fact that identity, and other more contextual factors such as interlocutor, play a role in language variation and change; that would be absurd with the abundance of studies showing that this is the case. However, in many studies focusing on language variation and change among people with foreign backgrounds, the aspect of ethnicity is highlighted and by doing so we miss that the variation might stem from a more complex combination of social factors which in this study are captured by the variable of neighborhood and the concept of communication density.

4.4 The fourth study, Gross & Forsberg (submitted): "Weak lips? A possible merger of /i:/ and /y:/ in Gothenburg"

The final study focuses on the finding from the third study that /i:/ and /y:/ exhibit coherence and have a centralized place of articulation for most of the Gothenburg speakers. During the initial segmentation of the SSG material I observed that some informants seemed to merge their /y:/ and /i:/ into [ɨ:]. However, it can be hard to know if individual observations are observations of a slip of the tongue from a single informant or a more consistent pattern in a group, especially so when working with materials the size of SSG, which is why I started to make impressionistic annotations of the actual productions of /i:/ and /y:/.

The results from the third study showed that the NE group separated the two vowels using more than one feature, and the impressionistic coding showed that the apparent merger never seemed to occur among informants from NE neighborhoods. In other words, if there is a merger it is only present among the informants who do not live in the northeastern suburbs, and explanations based on language contact cannot be used to explain the observations. As was evident from the third study, the non-NE informants have a coherent pattern for place of articulation for /i:/ and /y:/ i.e. there is no difference in advancement and openness, and the only articulatory feature that separates the two vowels appears to be lip rounding.

In the present study, which is co-authored with Julia Forsberg, we analyzed both acoustic and perceptual data to try to ascertain the underlying reason for the merger, while looking at the individual variation in the group. The results from the acoustic analysis showed no difference in PC1 and PC2. The analysis of the individual variation showed that a majority of the informants had merged forms,

though most also produce some unmerged forms. As a consequence of the acoustic results, we designed a perception experiment to test if listeners could distinguish productions of /i:/ and /y:/ when segments were played in isolation and if they were able to hear the merged forms when they occurred in a word; i.e. could the listeners hear that /y:/ was produced as [ɨ:] when the production occurred in a word. The results show that the listeners classified the [y:] productions as /i:/ when played in isolation, and that the merged forms were perceived as /y:/ when the listeners heard the vowel within a word context. Since the acoustic analysis confirmed that only lip rounding was separating the two phonemes, and that listeners more often classified [y:] as /i:/, but not the other way around i.e. [i:] as /y:/, when listeners hear the productions in isolation, we can conclude that lip rounding is a weak perceptual feature. However, the phonemic knowledge that speakers of Swedish have of /y:/ seems to hide the merged productions, and merged forms are very hard to detect, at least to speakers of the variety in question.

The results from this study confirm Eckert's and Labov's (2017) argumentation that mergers are "invisible" processes and therefore cannot be motivated by conscious manipulation. Instead, other factors such as language or dialect contact, or internal factors will motivate the change. And in a case where the opposition is based on only one distinctive feature, the direction of the change will be in favor of the perceptually stronger feature and the weaker one will be lost.

4.4.1 My contribution to the paper

This paper was coauthored with Julia Forsberg; I am the first author. The acoustic analysis, segmentation and statistics were done by me. The perception experiment was designed by both of us, but Forsberg put it together and created the web form. Most of the writing was done together. However, due to the division detailed above, the background section, sections on acoustic analysis and statistics were mainly written by me, and Forsberg was responsible for the part describing the perception test. However, as first author I was responsible for finalizing the text.

CHAPTER 5

General conclusions

One of the aims of this thesis has been to provide a sociolinguistic description of the structuring and variation of vowels in Gothenburg. However, this descriptive aim is not possible to fulfil without considering the social context of the linguistic variation. The social and internal motivation behind this variation has been investigated in three studies: the first by contrasting variation and change of /ɛ:/ in the two largest cities of Sweden (Stockholm and Gothenburg); the second by focusing on Gothenburg and relating the variation found in individual vowels to a system; and the third by drawing special attention to the group which grew up outside the northeastern suburbs of Gothenburg (the non-NE group), described in the second study, and the two vowels that often are merged by informants in this group. I would like to draw two general conclusions from the research process: first we have to accept the fact that social reality is always more complicated than we imagine when we design a sociolinguistic study, so we have to adjust our questions and conclusions accordingly. Second the ontological status of the variable under investigation must be considered when we draw conclusions about the observed variation.

What I aim for with the first conclusion is that predefined categories such as ethnicity, sex, neighborhood etc. should not govern how we analyze the sample population. As I have already stated in previous chapters of the thesis frame, the focus in this thesis is on the macro level. However, this does not imply that macro sociological categories should be used in an inflexible way as often has been the case in quantitative variationist work (Drummond 2018). Instead, we need to observe the heterogeneous environment that the city constitutes as a speech community, in the sense that Kerswill (1993) describes it. This means that we have to understand how "simple" categories intersect in ways that we cannot predict without a proper investigation of them. For example, what does it mean that a neighborhood is multiethnic? Usually it seems to be a way to describe a

neighborhood where the majority has an ethnicity other than the one that is assumed to be the country's "natural" or inherent ethnicity, whatever that is. Still, this way of speaking about a neighborhood masks the fact that most neighborhoods are more or less diverse and people with all kinds of background live their lives in them. To frame it in another way, there are people with different ethnicities in most neighborhoods just like there are people with different genders in most neighborhoods. If we sample a population from different parts of a city this fact should be reflected and simple categories ought not to be imposed on the population when we proceed in our linguistic research. A population sampled from only one area will inevitably have implications for what conclusions we can draw from the results. If we only sample speakers from a neighborhood usually characterized as multiethnic we cannot conclude that their foreign background is the reason behind their way of speaking. Instead we have to understand how a neighborhood is situated as a part of a larger whole, namely an urban area, and the reasons behind the demographic composition of the neighborhood. This reasoning leads us in the direction of bottom-up approaches where we first investigate how social categories intersect and create more complex social structures that we then can apply in a second step to investigate linguistic variation. A combination of careful reflection of the demographic data together with a non-hierarchical clustering method helped me to approach this goal and to show that housing segregation had an important effect on how young people produced their vowels.

The second conclusion is connected to the concept of *widespread change*. This concept was introduced in the first study to explain the patterns of variation observed in the two cities. With this concept I have tried to capture the fact that the same process where /ɛ:/ becomes more open and retracted is something that affects more than one dialect in a rather larger part of central Sweden. Thelander (1979a) already noted that widespread variables tend to be more stable than more geographically limited ones. As mentioned in Chapter 2, the concept of widespread change has been linked to the concept of *feature pool* by Cheshire et al. (2011: 176). This provides a partial explanation as to how a feature undergoing widespread change will be highly accessible for all speakers to use. This use will, according to Cheshire et al. (2011), not necessarily be dependent on participation in close-knit networks or indexed with local social values (Cheshire et al. 2011). The ontological

status of this variable is therefore an important aspect to consider with regard to how the variation of /ɛ:/ is patterning in Gothenburg.

The /ɛ:/ variable can be put in contrast with the variable /i:/, which in the Gothenburg dialect has a centralized variant [ɨ:]. Judging from previous reports on /i:/, the local centralized variant has undergone a change from being a variant used by the less dialectal speakers to be the variant used by persons who speak the Gothenburg dialect (Björseth 1957, Andersson 2006). Thus it is something that could be characterized as a *local* change. This contrast between local and widespread is probably a reason why we find differences in the patterning of the two long vowels, /i:/ and /ɛ:/. Using a widespread change in one's repertoire could be a safe way to mark social boundaries without taking a strong stance. A local change will however be indexed with local values and using local variants will be a way of taking a stance and positioning oneself as part of the local community in question. A first condition for a speaker to choose the variant is therefore that she or he feels a sense of belonging to the community to which the variant has its indexical association. If a speaker does not feel like they belong to the community, she or he will be less likely to use the variant. There is actually evidence that adolescents in the northeastern suburbs do not feel like they belong in the parts of the city where the Gothenburg dialect is more often spoken, and the centralized variant is the norm. In the radio documentary *Återvändarna* 'The returners' (Collmar 2014-08-29) the issue is discussed in the context of pupils and their parents choosing an upper secondary school. The schools in central Gothenburg are valued as high prestige while the northeast suburban school has low prestige. Still, young people from the northeastern suburbs who have managed to get a place in an inner-city school sometimes return to the suburban ones after a short time in the more prestigious central school, describing a sense of not feeling at home in the inner-city. This sense of not belonging could help explain both why the speakers growing up in the northeastern suburbs (the NE-group) seem to lead the widespread change of /ɛ:/ toward a more open and retracted /ä:/-vowel and do not use the centralized [ɨ:] and centralized /y/. The open /ɛ:/ variant is the default variant to choose if you do not want to use the more dialectal variants and the same goes for the "cardinal" /i:/ variant. However, an interesting phenomenon in connection with this is that /i:/ and /y:/ do not have a coherent place of articulation in the NE-group;

why this is the case is more difficult to connect to this reasoning. However, the coherence between /i:/ and /y:/ among the non-NE informants is a prerequisite for a merger, as described in the third paper, a change that will lead to accentuating the differences in the use of the /i:/ variable in Gothenburg between neighborhoods.

	Front	Central	Back
Close	[i:]	[ɨ:] [y:]	[ʉ:]
Close-mid	[ɐ:] [ʉ:]	[y:] [ʉ:]	[o:]
Open-mid			
Open	[ɛ:]~[æ:] [ø:]~[œ:]		[ɑ:]
	[æ:] [œ:]		

Figure 2. A schematic description of the vowel variation in Gothenburg youth language, red= NE & light blue= non-NE, "~"indicates variation consistent with the pre-rhotic rule.

To conclude, figure 2 gives a schematic description of the vowel variation among adolescents in Gothenburg. Black indicates that I have not found any variation for the vowel, light blue indicates the variants in the non-NE variety and the red indicates the variants found in the northeastern suburbs. All variation is connected to the front or central vowels and the back vowels are relatively stable. The /i:/ vowel has two variants in the city: [i:] and [ɨ:]. In the dialect [ɨ:] and [y:] have the same place of articulation and are only separated by the shape of the lips [ɨ:] unrounded and [y:] out-rounded; the two variants are also somewhat lower than the fronted counterpart [i:]. There is furthermore a tendency for these two vowels to be merged to [ɨ:], however, if this is a change in progress or not, can only be answered after further studies situating the merger diachronically are carried out. The /ʉ:/ vowel has not been investigated thoroughly, but two variants are found, separated by the level of fronting. However, the two variants are both lower and could be described as close-mid rather than close. Finally, the /ɛ:/ and /ø:/ vowels are both better characterized as open than open-mid in both groups, though internal factors seem to affect the variation in the non-NE varieties. Although, judging from Gross et.al (2016) and Gross (in press) the allophonic rule seems to disappear in the dialect as well.

5.1 *Future research*

In this thesis I have focused on quite broadly defined groups of speakers, and it should be kept in mind that the within these groups more fine-grained differences are hidden. Also, only one age group has been investigated, using language in a single activity, albeit a rather informal one. There is, in other words, plenty of further work to be done before we can actually say anything definite about variation in Gothenburg vowels. More work is needed to uncover if on the one hand there are variables that are stable across styles both within and between smaller groups, such as communities of practice, and on the other hand what variables are subject to stylistic variation. Sharma (2011) showed that speakers varied considerably depending on context between more ethnically colored variants and more standard ones. The studies in this thesis provide a possible starting point for investigating how (or if) the variables are used in stylistic identity work.

The concept of coherence between linguistic variants should also be further investigated. So far only phonetic and phonological features have been discussed. We do not know if there are there other linguistic features that might covary in a systematic way, for example, the fronted [i:] and non-inversion of subject and verb. There are many occurrences of XSV word order in some of the map task data and my impression is that this only occurs among some speakers in the NE-group; if this is the case, and if so what its relation to vowel production might be are however open questions.

Furthermore, the intention in the beginning of this project was to compare adolescents in Stockholm and Gothenburg. However, time did not allow the analysis of the vowel variation in Stockholm, so this is left for the future, aside from some minor comments on the lowering of the /ɛ:/ vowel in paper 1. Furthermore, investigation of how social categories reflect and are reflected in variation in the two cities could help us understand how linguistic variation associated with multilingual neighborhoods is situated in a national context where both language contact and dialect contact are in play.

SAMMANFATTNING

Sociolingvistiska studier om språklig variation och förändring i Göteborg lyser med sin frånvaro. Därför har fokus i de artiklar som utgör kärnan i denna avhandling legat på att beskriva, fördjupa och kontrastera den språkliga variationen i staden, genom att undersöka språkbruk bland ungdomar i staden. Varje artikel har skrivits utifrån mer specifika frågeställningar och syften, men tre mer generella syften kan ändå skiljas ut:

1. Bidra med en sociolingvistisk beskrivning av vokalsystemet i göteborgskan.
2. Utröna hur talare i olika grupper använder språkdrag som har visat sig vara del i mer allmänna pågående förändringar i det svenska språket. Vad är förhållandet mellan språklig variation och sociala faktorer såsom utländsk bakgrund, kön, klass och bostadsområde?
3. a. Undersöka vokaler, vilka är mindre framträdande språkdrag i de nya varieteter som påstås växa fram i urbana storstadsmiljöer, än de variabler som tidigare undersökts.
   b. Undersöka graden av samvariation mellan vokalfonem och allofoner. För att göra detta studeras vokalerna som ett system, snarare än som individuella variabler.

Två talspråkskorpusar har legat till grund för artiklarna. Den första korpusen samlades in 2002 inom ramen för projektet S*pråk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer* (SUF, [Bijvoet et al. 2001]) och består av inspelningar av 222 ungdomar i olika samtalssituationer från 8 skolor i Stockholm, Göteborg och Malmö. Åldern på informanterna var mellan 16–19 och alla informanter gick en samhällsvetenskaplig linje på gymnasiet. För denna avhandling har samtliga 119 intervjuer från Stockholm och Göteborg lyssnats igenom, varav

57 informanter producerade tillräckligt många belägg av /ɛ:/ samt den öppnare allofonen [æ:] för att kunna räknas in i studien.

Den andra korpusen, *Språkbruk i Stockholm och Göteborg (SSG),* samlades in av mig och Julia Forsberg hösten 2014. Korpusen är inspirerad av SUF-materialet och intervjuerna är snarlika, en lärdom från arbetet med SUF-materialet var dock att intervjuerna inte kunde garantera att tillräckligt många vokalförekomster producerades. Detta resulterade i att jag och Forsberg utvecklade en map-task där informanterna fick samarbeta verbalt i självvalda kompis-par för att rita ut vägen på en karta (se appendix). Totalt samlades talspråksmaterial in från 111 informanter distribuerade över fyra skolor, två i Stockholm och två i Göteborg.

De fyra artiklarna

Tre av artiklarna fokuserar på fonetisk och fonologisk variation i vokaluttalet bland informanterna i de två korpusarna. Den första artikeln undersöker socialt betingad variation i användningen av /ɛ:/ allofonerna genom att kontrastera språkvariationen i Stockholm och Göteborg, och materialet är taget från SUF-korpusen. De utomspråkliga faktorer som undersöks är kön och huruvida talarna har svensk eller utländsk bakgrund, samt om det är någon skillnad mellan talarna i de två städerna. Resultaten från analyserna visar att de två städerna skiljer sig åt med hänsyn till om det finns någon socialt betingad variation eller inte. I Stockholm finns det inga signifikanta skillnader mellan ungdomar med utländsk bakgrund och svensk bakgrund, alla använder mer öppna varianter av /ɛ:/ och den allofoniska variationen är så pass liten att den kan antas ha neutraliserats. I Göteborg är situationen annorlunda och tre grupper kan identifieras: ungdomar med utländsk bakgrund, tjejer med svensk bakgrund och killar med svensk bakgrund. De olika grupperna verkar orientera sig mot olika normer: ungdomarna med utländsk bakgrund beter sig på ett likartat sätt som ungdomarna i Stockholm och deltar i förändringsprocessen mot ett öppnare uttal av /ɛ:/ och en neutralisering av allofonvariationen; tjejerna med svensk bakgrund är den grupp som uppvisar störst variation i uttalet trots att de upprätthåller en allofonisk distribution; medan killarna med svensk bakgrund beter sig mer dialektalt. Resultaten visar hur dialekt- och

språkkontakt bör beaktas gemensamt när man undersöker variation och förändring i städer.

Den andra artikeln problematiserar den sociolingvistiska variationen i Göteborg genom att visa hur socioekonomisk bakgrund, utbildningsnivå och utländsk bakgrund är överlappande bakgrundsfaktorer vilka påverkar den lingvistiska situationen i staden, något som kan fångas upp genom att använda en binär uppdelning mellan talarna som bor i de nordöstra stadsdelarna (NE) å ena sidan och resten av Göteborg (notNE) å andra som en övergripande oberoende variabel. I artikeln studerar jag vokalerna som ett system och fokuserar på hur vokaler samvarierar på ett systematiskt sätt. Resultaten visar att det finns signifikanta skillnader i uttalet av /i:/ och /y:/ vilka har tydligt göteborgsdialektala varianter som uppvisar samvariation, där notNE-ungdomarna använder dialektvarianterna på ett koherent vis medan NE-ungdomarna inte använder dialektvarianterna, och de två vokalerna uppvisar ingen samvariation. Ett annat vokalpar som uppvisar samvariation är /ɛ:/ och /ø:/ och även här finns det en signifikant skillnad mellan de två grupperna. Uttalet av de två vokalerna är i förändring mot ett öppnare uttal i en stor del av det centralsvenska dialektområdet och NE-ungdomarna leder förändringen medan notNE är på efterkälken. Utifrån dessa resultat drar jag slutsatsen att det inte finns något enkelt förhållande mellan utländsk bakgrund och den lingvistiska situationen i Göteborg. Istället är det bostadssegregationen som påverkar och begränsar möjligheterna till använda och överföring av dialektdragen och, liksom att NE-ungdomarna leder förändringen av /ɛ:/ och /ø:/ då de orienterar sig bort från den lokala normen.

Den tredje studien undersöker ett möjligt sammanfall mellan /i:/ och /y:/ till /i:/, genom ett perceptionsexperiment med vokalproduktioner från SSG-korpusen och akustisk analys av 19 talare som alla har notNE bakgrund från samma korpus. Perceptionsexperimentet visar att lyssnarna klassificerar de orundade /y:/ produktionerna som /i:/ både i isolering och när lyssnarna hör ordet där produktionen förekommer. Dock visar det sig att lyssnarna även klassificerar de rundade /y:/ produktionerna som /i:/ när de lyssnar på dem i isolation, men att de klassificeras som /y:/ när de hör produktionen i ett ord. Resultaten från perceptionsexperimentet visar att /y:/ rör sig mot /i:/ perceptuellt och att lyssnarna använder sig av kategorisk perception när de hör produktionen i ett ord. Den

akustiska analysen visar att /i:/ produceras som en central orundad sluten vokal [ɨ:] av talarna i SSG-korpusen och att det inte finns någon skillnad mellan /i:/ och /y:/ produktionerna kopplat till de två första formanterna, vilket betyder att läpprundning är den enda särdraget som skiljer de två vokalerna åt. Slutsatsen är att /i:/ och /y:/ håller på att sammanfalla till /i:/ bland göteborgsungdomar och att detta sammanfall följer merger-by-approximationmodellen. Orsaken till att /y:/ blir /i:/ och inte tvärtom förklaras genom att läpprundning är ett perceptuellt svagare drag.

Den fjärde artikeln är en metodartikel som beskriver map-task som sociolingvistisk datainsamlingsmetod och lyfter fram de olika komponenterna i dialogen som kan utnyttjas för att påverka och stimulera informanternas interaktion. I artikeln presenteras också hur informanterna upplevde deltagandet i studien och hur de upplevde de olika inspelningstillfällena, intervju och map-task. Som redskap för att beskriva interaktionen under inspelningarna används terminologi hämtad från Bells beskrivning av audience design (1984) och Hymes beskrivning av speech event (1972) för att beskriva de delar som inte är direkt kopplade till deltagarnas roller. Analysen visar att deltagarna bäst kan beskrivas i termer av talare (1st person speakers), tilltalad (+ 2nd person addressees) samt lyssnare (auditors). Talare och tilltalad är informanter och de är fria att kliva i och ur de olika rollerna och forma reglerna för interaktionen, medan lyssnaren är forskaren som gör inspelningen. Denne har efter att uppgiften startat inte samma rättighet att forma interaktionen, istället är det upp till de två informanterna att nominera lyssnaren till tilltalad. När det kommer till de delar som inte direkt har med deltagarnas roller att göra identifierar vi fyra komponenter: *genre, syfte/orsak, miljö* och *samtalsämne*. **Genren** kan beskrivas som vägvisning men under intervjuerna med informanterna förstod vi att de upplevde genren som en typ av spel. **Syftet** skiljer sig åt med hänsyn till deltagarrollerna. Informanterna upplever huvudsyftet som att de skulle lösa en klurig uppgift, men beskrev också ett antal sekundära syften så som att lösa uppgiften så snabbt och/eller bra som möjligt eller att lösa uppgiften bättre än några andra kompisar som gjort uppgiften före dem. För forskaren är istället syftet att få så många typer av alla relevanta tokens som möjligt och att spela in dessa med så bra ljudkvalitet som möjligt. **Miljön** manipuleras genom att placera deltagarna i rummet så att de inte kan se hur samtalspartnerns karta ser ut, för att stimulera

interaktionen. Förutom detta är **samtalsämne** viktigt att beakta för att stimulera samtalet. Detta kan bäst beskrivas som uppbyggt av tre subkomponenter: ge beskrivning av vägen, behålla och uppdatera en common ground, och metakommunikation runt uppgiften. Den subkomponent som framför allt ska beaktas vid utformandet av en map-task är behålla och uppdatera en common ground då det är under dessa episoder som de två informanterna blir mer jämlika i turtagningen.

Allmänna slutsatser

Två generella slutsatser kan lyftas fram från de fyra studierna. Den första är att vi inte kan använda förbestämda sociala kategorier så som kön, etnicitet, socialgrupp etc. för att undersöka socialt betingad språkvariation. Istället måste vi först undersöka hur dessa faktorer flätas samman och skapar sociala hierarkier och gränser. I den andra vokalstudien visade jag hur man kan undersöka detta med hjälp av en icke-hierarkisk klusteranalys, samt hur ett antal sociala faktorer sammanföll och kunde fångas upp genom att beakta hur bostadssegregationen i Göteborg är strukturerad.

Den andra generella slutsatsen är kopplad till vad jag kallar den ontologiska statusen av en variabel. De fyra variabler som kan sägas utgöra skiljelinjen mellan olika talare i Göteborg kan kopplas till geografiskt lokala och geografiskt spridda förändringar. Det två variablerna /i:/ och /y:/ som är kopplade till en lokal förändring kan antas vara indexerade med lokala värden, och att använda dem innebär att markera sin tillhörighet till den lokala göteborgska dialektgemenskapen. Om en individ inte känner att den kan identifiera sig med den gemenskapen kommer individen vara tvungen att välja en variant som inte är indexerad med ett sådant värde. Likaså kan man tänka sig att de geografiskt spridda förändringarna kommer att plockas upp av den grupp talare som inte i lika stor utsträckning påverkas av lokala dialekt-normer, vilket också är fallet i de två största studierna.

REFERENCES

Adank, P., Van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical society of America,* 116(3), 1729-1738.

Andersson, L-G. (2006). "Hur Göteborgskan växte fram" [How the Gothenburg dialect was formed] in K. Nelson, A. Larsson & L.O. Lööf (eds.). *Göteborg förr och nu*. [Gothenburg then and now] Göteborg: Rundqvists.

Bell, A. (1984). Language style as audience design. *Language in Society,* 13, 145-204.

Bell, A. (2001). Back in style: reworking audience design. In Penelope Eckert & John R. Rickford (eds.), *Style and Sociolinguistic Variation*. Cambridge, UK: Cambridge University, 139-169.

Benor, S. B. (2010). Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics*, *14*(2), 159-183.

Bijvoet, E., Boyd, S., Ekberg, L., Fraurud, K., Källström, R., & Lindberg, I. (2001). *Språk och språkbruk bland ungdomar i flerspråkiga storstadsmiljöer*. Grant proposal to the Bank of Sweden Tercentenary Foundation. June 2001, Unpublished research program.

Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer. Version 6.0.13, retrieved January 31, 2016 from http://www.praat.org/.

Boyd, S. (2010). Vems svenska ska beskrivas? Problem med att använda termen native speaker omungdomar i flerspråkiga storstadsmiljöer och termen varietet om deras språkbruk. [Whose Swedish will be described? The problem with using the term 'native speaker' about youth in multilingual large urban centers and the term 'variety' about their language use.] In C. Falk, A. Nord, & R. Palm (eds.), *Svenskans beskrivning 30. Förhandlingar vid Trettionde sammankomsten för svenskans beskrivning. [Description of Swedish 30. Proceedings of the 30th Congress for the Description of Swedish.]* Stockholm: Stockholm University, 13–36.

Boyd, S & Fraurud, K. (2010). Challenging the homogeneity assumption in language variation analysis: Findings from a study of multilingual urban spaces. In: *Language and Space. An International Handbook of Linguistic Variation*. : *Volume 1: Theories and Methods*. [ed] Auer, Peter & Schmidt, Jürgen Erich, Berlin/New York: Mouton de Gruyter, 2010, 1, 686-706.

Boyd, S., Hoffman, M. F., & Walker, J. A. (2015). 14 Sociolinguistic variation among multilingual youth: comparing Swedish cities and Toronto. In: *Language, youth and identity in the 21st century: Linguistic practices across urban spaces*. Cambridge University Press Cambridge, 290-305.

Björseth, B. (1957). *Göteborgsspråket. [The language of Gothenburg.]* Stockholm: Norstedts Förlag AB.

Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics* 15 (2), 151–196.

Chomsky, N. (1963). *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.

Cibulka, P. (2016). *A Universal of Human Interaction? – Manual Movement as Interactional Practice in Spoken and Signed Conversation*. Ph.D. dissertation. University of Gothenburg.

Clyne, M. (2000). Lingua franca and ethnolects in Europe and beyond. *Sociolinguistica*, 14, 83–89.

Collmar, L. (Sveriges radio P1). (20140829). *Återvändarna*. [podcast] downloaded 20180815.

Derdemezis, E., Vorperian, H., Kent, R., Fourakis, M., Reinicke, E., & Bolt, D. (2016). Optimizing vowel formant measurements in four acoustic analysis systems for diverse speaker groups. *American Journal of Speech-Language Pathology, 25*(3), 335.

Di Paolo, M., Yaeger-Dror, M., & Beckford Wassink, A. (2011). Analyzing vowels. In M. Di Paolo & M. Yaeger-Dror (eds.), *Sociophonetics: A student's guide*. New York: Routledge, 87–106.

Disner, S. F. (1980). Evaluation of vowel normalization procedures. *The Journal of the Acoustical Society of America*, *67*(1), 253-261.

Drummond, R. (2018). Maybe it's a grime [t] ing: th-stopping among urban British youth. *Language in Society*, 47(2), 171-196.

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Eckert, P. (2001). Style and social meaning. In Eckert, Penelope & John R. Rickford (eds.) *Style and sociolinguistic Variation*, 119-126. Cambridge, UK: Cambridge University.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism* 12(1–2), 25–42.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology*, *41*, 87-100.

Eckert, P. (2016). Variation, meaning and social change. *Sociolinguistics: theoretical debates*, 68-85.

Eckert, P., & Labov, W. (2017). Phonetics, phonology and social meaning. *Journal of Sociolinguistics, 21*(4), 467-496.

Eklund, I., & Traunmüller, H. (1997). Comparative study of male and female whispered and phonated versions of the long vowels of Swedish. *Phonetica*, *54*(1), 1-21.

Elert, C-C. (1994). Indelning och gränser inom området för den nu talade svenskan: En aktuell dialektografi. [Divisions and borders in the area of modern spoken Swedish: A current dialectography.] In L-E. Edlund (ed.), *Kulturgränser—myt eller verklighet? Vol. 4 of DIABAS* [Cultural borders – myth or reality?]. Umeå: Department of Language Studies, Umeå University, 215–228.

Forsberg, J., & Gross, J. (Unpublished). *"You change your speech depending on who you talk to, but I didn't change much": The map-task viewed through the lens of audience design.*

Fraurud, K., & Boyd, S. (2011). The native–non-native speaker distinction and the diversity of linguistic profiles of young people in multilingual urban contexts in Sweden. *Young Urban Swedish*, 67.

Ganuza, N. (2008). *Syntactic variation in the Swedish of adolescents in multilingual urban settings: Subject-verb order in declaratives, questions and subordinate clauses*. Ph.D. dissertation, University of Stockholm.

Gross, J. (in press). Segregated vowels: language variation and dialect features among Gothenburg youth. *Language Variation and Change*.

Gross, Johan, Boyd, S., Leinonen, T. & Walker, J. A. (2016). A tale of two cities (and one vowel): Sociolinguistic variation in Swedish. *Language Variation and Change*. 28, 225–247.

Gross, J., & Forsberg, J. (under review after revision) *Weak lips? A possible merger of /i:/ and /y:/ in Gothenburg*.

Grönberg, A. (2004) *Ungdomar och dialekt i Alingsås*. PhD thesis in Nordic languages. Gothenburg: Dept of Swedish language, University of Gothenburg.

Harrison, P. (2013). *Making accurate formant measurements: an empirical investigation of the influence of the measurement tool, analysis settings and speaker on formant measurements*. PhD thesis, University of York.

Herold, R. (1990). *Mechanisms of merger: The implementation and distribution of the low back merger in Eastern Pennsylvania*, PhD thesis, University of Pennsylvania.

Hoffman, M.F. (2010). The role of social factors in the Canadian Vowel Shift: Evidence from Toronto. *American Speech* 85(2),121–140.

Hoffman, M.F., & Walker, J.A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22(1), 37–67.

Holmberg, B. (1976). *Språket i Göteborg*. [The language of Gothenburg.] Stockholm: Esselte Studium.

Horvath, B. M. (1985). Variation in Australian English. The Sociolects of Sydney. *Cambridge Studies in Linguistics London*, (45), 1-200.

Hudson, R. A. (1980). *Sociolinguistics*. Cambridge: Cambridge University Press.

Hyltenstam, K. (2010). Begreppen språk och dialekt – om meänkielis utveckling till ett språk, In: *Sveriges sju inhemska språk – ett minnoritetsspråksperspektiv*. Ed. Kenneth Hyltenstam. Lund: Studentlieratur.

Jacobi, I. (2009). *On variation and change in diphthongs and long vowels of spoken Dutch*. Universiteit van Amsterdam. Ph.D. thesis.

Jaspers, J. (2008). Problematizing ethnolects: Naming linguistic practices in an Antwerp secondary school. *International Journal of Bilingualism* 12(1), 85–103.

Johnson, K. (2011). *Acoustic and Auditory Phonetics* (3rd ed.). Wiley-Blackwell.

Källström, R., & Lindberg, I. (2011). *Young urban Swedish: Variation and change in Multilingual settings*. Gothenburg: Department of Swedish Language, University of Gothenburg.

Kerswill, P. (1993). Rural dialect speakers in an urban speech community: the role of dialect contact in defining a sociolinguistic concept. *International Journal of Applied Linguistics*, *3*(1), 33-56.

Kotsinas, U-B. (1988a). Immigrant children's Swedish—A new variety? *Journal of Multilingual and Multicultural Development* 9(1–2),129–140.

Kotsinas, U-B. (1988b). Rinkebysvenska – en dialekt? [Rinkeby Swedish—a dialect?] In P. Linell, V. Adelswärd, T. Nilsson, & P. A. Petersson (eds.), *Svenskans beskrivning 16. SIC 21a*. Linköping: Linköping University, 264–278.

Kotsinas, U-B. (1992). Immigrant adolescents' Swedish in multicultural areas. In C. Palmgren, K. Löfgren, & G. Bolin (eds.) *Ethnicity in youth culture*. Stockholm: Youth Research Center, Stockholm University, 43–62.

Kotsinas, U-B. (1994). *Ungdomspråk*. [The language of youth.] Uppsala: Hallgren and Fallgren.

Kotsinas, U-B. (1998). Language contact in Rinkeby, an immigrant suburb. In J. Androutsopoulos & A. Scholz (eds.), *Jugendsprache/langue de jeunes/youth language*. Frankfurt am Main: Peter Lang, 125–148.

Kotsinas, U-B. (2001). Pidginization, creolization and creoloids. In N. Smith & T. Veenstra (eds.), *Creolization and contact*. Amsterdam: John Benjamins, 125–156.

Kotsinas, U-B. (2014) Rinkebysvenskan—en dialekt? [Rinkeby Swedish—a dialect?] In S. Hellberg (ed.), *Nordiska språk då och nu: Artiklar av stockholmsforskare från skilda tider*. [Nordic languages then and now: Articles by Stockholm researchers from different times.] Stockholm: Department of Swedish Language and Multilingualism, Stockholm University.

Labov, W. (1963). The social motivation of a sound change. Word, 19(3), 273-309.

Labov, W. (1966). *The social stratification of english in new york city*. Washington: Center for Applied Linguistics.

Labov, W. (1968). A Study of the Non-Standard English of Negro and Puerto Rican Speakers in New York City. *Volume II: The Use of Language in the Speech Community*.

Labov, W. (1972). *Sociolinguistic patterns* (No. 4). University of Pennsylvania Press.

Labov, W. (1984). Field methods of the project on linguistic change and variation. *Language in use: Readings in sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28-53.

Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors*. Oxford: Blackwell.

Labov, W. (2001a). The anatomy of style-shifting. In Penelope Eckert & John R. Rickford (eds.) *Style and social variation*. Cambridge: Cambridge University Press.

Labov, W. (2001b). *Principles of linguistic change. Vol. 2, Social factors*. Oxford: Blackwell.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology*, *60*(1), 44.

Le Page, R.B., & Tabouret-Keller, A. (1985). *Acts of identity: Creole-based approaches to language and ethnicity*. CUP Archive.

Leinonen, Therese. (2010). *An acoustic analysis of vowel pronunciation in Swedish dialects*. Ph.D. dissertation, University of Groningen.

Lobanov, B.M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, *49*(2B), 606-608.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk. 3rd Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of sociolinguistics*, *5*(4), 530-555.

Milroy, L. (1986). Social network and linguistic focusing. In *Dialect and language variation*, 367-380.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: Method and interpretation* (Language in society (Oxford), 34). Malden, MA: Blackwell Pub.

Mufwene, S.S. (2002). Competition and selection in language evolution. *Selection*, 3(1), 45-56.

Nationalencyklopedin. (2018). *System*. Retrieved: 2018-03-22. https://www.ne.se/uppslags-verk/encyklopedi/lång/system.

Nordberg, B. (1972). Böjningen av neutrala substantiv i Eskilstunaspråket: Utomspråkligt och inomspråkligt betingad variation. In: *Nysvensaka studier* 51 (1971), 117-227.

Nordberg, B. (1975). Contemporary social variation as a stage in a long-term phonological change. In K-H. Dahlstedt (ed.), *The Nordic languages and modern linguistics*. Stockholm: Almqvist & Wiksell, 587–608.

Nordberg, B. (1985). Det mångskiftande språket: Om variation i nusvenskan. *Ord och stil* 14. Malmö.

Nordström, P.E., & Lindblom, B. (1975). A normalization procedure for vowel formant data. In: *Proceedings of the eighth international congress of phonetic sciences*. Leeds.

Nearey, T. (1978). *Phonetic feature systems for vowels*. University of Connecticut. Ph.D. thesis

Plomp, R., Pols, L. C. W., & Van de Geer, J. P. (1967). Dimensional analysis of vowel spectra. *The Journal of the Acoustical Society of America*, 41(3), 707-712.

Quist, P. (2008). Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice. *International Journal of Bilingualism* 12(1–2), 43–61.

Rampton, B. (1995). Language crossing and the problematisation of ethnicity and socialisation. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, *5*(4), 485-513.

R Core Team. (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/.

Sankoff, G. (1980) A quantitative paradigm for the study of communicative competence. In G. Sankoff (ed.), *The Social Life of Language*. Philadelphia: University of Pennsylvania Press, 47–79.

Saussure, F. D. (1916). *Kurs i allmän lingvistik*. Lund: Arkiv.

Saville-Troike, M. (2003). *The ethnography of communication: An introduction*. (3rd edition). Oxford: Blackwell.

Sharma, D. (2011). Style repertoire and social change in British Asian English. *Journal of Sociolinguistics*, 15(4), 464-492.

Schoonen, R., & Appel, R. (2005). Street language: A multilingual youth register in the Netherlands. *Journal of multilingual and multicultural development*, *26*(2), 85-117.

Sundgren, E. (2002). *Återbesök i Eskilstuna: en undersökning av morfologisk variation och förändring i nutida talspråk*. Ph.D. dissertation, Uppsala universitet.

Svendsen, B. A., & Røyneland, U. (2008). Multiethnolectal facts and functions in Oslo, Norway. *International Journal of Bilingualism* 12(1), 63–83.

Syrdal, A. K., & Gopal, H.S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, *79*(4), 1086-1100.

Thelander, M. (1979a). *Språkliga variationsmodeller tillämpade på nutida burträsktal. Del 1, Materialbeskrivning och analys på språkdragsnivå*. Studia philologiae Scandinavicae Upsaliensia, 14. Uppsala; Stockholm Universitet: Almqvist & Wiksell international (distr.).

Thelander, M. (1979b). *Språkliga variationsmodeller tillämpade på nutida burträsktal. Del 2, Teoretiska utgångspunkter och analys på språkartsnivå*. Studia philologiae Scandinavicae Upsaliensia, 14. Uppsala; Stockholm: Universitet: Almqvist & Wiksell international (distr.).

Thomas, E. R. (2002). Instrumental phonetics. *The handbook of language variation and change*, 168-200.

Thomas, E. R. (2011). *Sociophonetics: an introduction*. Palgrave Macmillan.

Trudgill, P. (1972). Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in society*, *1*(2), 179-195.

Trudgill, P. (1974). The social differentiation of English in Norwich. *Cambridge studies in linguistics,* 13. Cambridge: Cambridge U.P.

Trudgill, P. (2011). *Sociolinguistic typology: Social determinants of linguistic complexity* (Oxford linguistics). Oxford: Oxford University Press.

Trudgill, P., & Foxcroft, T. (1978). On the sociolinguistics of vocalic mergers: Transfers and approximations in East Anglia, in P. Trudgill (ed.) *Sociolinguistic patterns in British English*, London: Edwin Arnold, 69-79.

Watt, D., Fabricius, A., & Kendall, T. (2011). More on vowels: plotting and normalization. In M. Di Paolo & M. Yaeger-Dror (eds.), *Sociophonetics: A student's guide*. New York: Routledge, 107–118.

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: new directions. *Annual Review of Psychology*, *66*.

Weinreich, U, William, L & Herzog, M. I. (1968). *Empirical foundations for a theory of language change*. Directions for Historical Linguistics: A symposium. University of Texas press: Austin.

Wiese, H. (2009). Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119(5):782–806.

Wiese, H. & Rehbein, I. (2016). Coherence in new urban dialects: A case study. *Lingua*. 172-173, 45-61.

Wolfram, W. (1969). *A sociolinguistic description of Detroit negro speech* (Urban language series, 5). Washington, D.C.: Center for applied linguistics.

Wolfram, W. (2007). Sociolinguistic folklore in the study of African American English. *Language and Linguistics Compass*, *1*(4), 292-313.

STUDY I

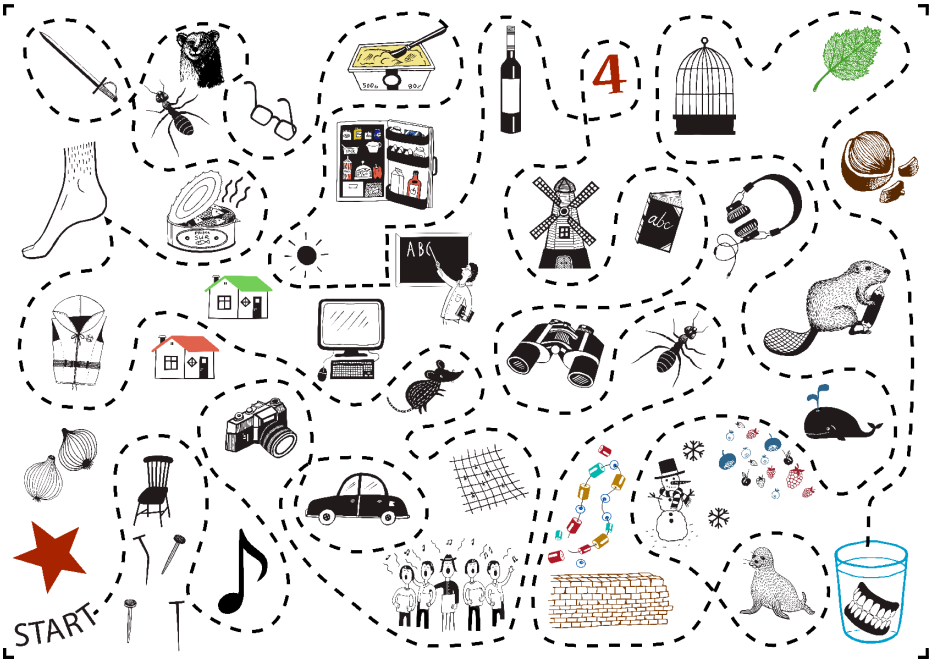
STUDY II

STUDY III

STUDY IV

APPENDIX

This appendix includes the maps that Forsberg and I developed for the data collection in the SSG project. The maps are organised so that the two first are the ones that were used to collect the Swedish vowels and the third was designed to collect the English plosives. The top map is the one that the nominated direction-giver used and the one below is the one that was used to draw the path. The pictures where drawn by Emelie Höcks.


4
START
ABC
abc


4
4
abc


START
NYHETER
TV
Väder


START