



## UvA-DARE (Digital Academic Repository)

### On variation and change in diphthongs and long vowels of spoken Dutch

Jacobi, I.

**Publication date**

2009

**Document Version**

Final published version

[Link to publication](#)

**Citation for published version (APA):**

Jacobi, I. (2009). *On variation and change in diphthongs and long vowels of spoken Dutch*. [Thesis, fully internal, Universiteit van Amsterdam].

**General rights**

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

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

## 6. ON SPEECH VARIATION AND SOCIAL BEHAVIOUR

*Abstract* The previous chapters showed that the social background of speakers in terms of age and educational/occupational level had an effect on the sub-phonemic realization of the vowel phonemes /e:/, /o:/, /ɛi/, /œy/, and /au/. From the small perception task that we described in chapter 5 we can infer that the listener's age had a comparable impact on the perceptual categorization of vowel variants. This chapter will offer a literature overview on how and why phonetic variation is socially intertwined. With the objective of defining its structure, and to explain origin and change in variation, linguistic approaches as well as processes studied in psychology will be considered. We will show that the effects found in literature on the articulatory-auditory interaction in human beings coincide with the effects found in our acoustic and perception data.

### 6.1 *Introduction*

Chapter 4 showed that socially structured variation can be found in fine-grained phonetic detail such as sub-phonemic differences in vowel realization. Our cohort of 70 speakers differing in social background, differed also in their pronunciation patterns of /e:/, /o:/, /eɪ/, /æy/, and /au/. The speakers' level of education (or occupation) could be related to their vowel realizations. Generally, the low educated speakers showed higher onsets and less diphthongization than the high educated speakers.

Moreover, the speakers' vowel realization could be grouped according to the speakers' age group at the time of recording. Realization differences between the two educational levels varied characteristically between speaker generations: the largest difference between the two educational levels were found for the mid generation (aged 36 to 54 at the time of recording), which included the speakers who were born between 1945 and 1965. The pronunciation of the high educated changed remarkably from the old to younger speaker generations, whereas the pattern of the low educated hardly changed with the generations. Considering the role of age in listener behavior, the results of the perception experiment in chapter 5 coincide with the age effects that were found in the speakers' realizations in chapter 4. When split into the young, mid and old generations as the speakers in chapter 4, the listeners in chapter 5 differ in their perception of acoustic differences age-group dependently.

Together, the results of the acoustic speaker analysis and the perception experiment suggest that sub-phonemic vowel perception is affected by the social background as much as vowel production. Before interpreting the results of our present study, in the following we will try to explore what causes the variation in articulatory and perceptual behavior and why both dimensions seem to be connected. We will review relevant literature from linguistics and psychology on the topic.

### 6.2 *The Structure of Variation and Change*

To determine to what extent variation is perceived and can be imitated or accommodated, the role and the processing of phonetic and social features need to be considered. Though social research in the psychological-cognitive area that goes beyond the pragmatic-semantic level is comparably scarce and recent, (and often based on new techniques of brain-imaging, which might not yet be totally reliable in terms of mapping precision), the basic concept of social recognition and processing of inter-human contact will give further explanation to variation and changing speech behavior.

### 6.2.1 *Sociolinguistic Approach*

Starting with a study on the sociolinguistic structures and social tension in the community of Martha's Vineyard (Labov, 1963 [80]), foremost research on the connection of sound patterns and a speaker's social background or social awareness was carried out on American English by William Labov from the 1950's on. In his seminal study from 1966 on the pronunciation of the postvocalic /r/ in New York City, he found the production variants to correlate with socio-economic class (Labov, 1966 [81]). Labov's research resulted in a class stratification pattern, where a variant is used most frequently by the highest-status class and least frequently by the lowest-status class, and where alternatives of saying the same will have social significance.

With the findings on correlations of social factors and linguistic variants, the linguistic tradition to focus on competence (or internal language) in distinction to performance (or external language) became problematic (Milroy & Gordon, 2003 [100]). Rather than treating language structure as invariant and variation as asocial, the variationist approach is based on the assumption that language variation is intrinsic and structured. Behavior variables and their social embeddings are thereby seen as essential in understanding the dynamics of language change. By comparing the existence of variants and their relative frequency at different points in time, the quantitative research paradigm enabled linguists to propose social explanations for changing frequencies and emergence of varieties in time, space, and social space (Milroy & Milroy, 1997 [99]). Without data on usage and attitudes, and without interlocking the collected linguistic forms with ordinary verbal interaction, linguistic changes can hardly be explained (Labov, 1989 [83]).

According to Labov, research on change should focus on the following points: Firstly, find the continuous matrix of social and linguistic behaviour in which the linguistic change is embedded (embedding problem). Secondly, find the trigger of the linguistic change (actuation problem). And thirdly, find out if the change from below (below the level of conscious awareness) is dependent upon high status and will become a prestige model, or if it is dependent upon low status and will be stigmatized. There has been some debate about the role of prestige and other tacit generalizations, and by now it is assumed that the crucial indicators of language change are rather locally determined social categories (Milroy & Gordon, 2003 [100]). Since these have different meanings in different communities, various interactions need to be considered, which complicates any proposal of generalization.

### 6.2.2 *On the Origin of Sound Change*

The cultural and psychological forces that were found to structure variation and change are only accounts of the spread of a variant, the 'maxi-sound change' (Ohala, 1993 [110]). The origin of a sound change, i.e. the fine-grained phonetic detail which selectively becomes

spread or not, referred to as 'mini-sound change', cannot be explained by these factors.

Often, a principle of least effort has been put forward as motivator or mechanism for sound change. While ease of effort might fit other levels of speech (e.g. grammar), least effort in terms of articulation is hardly defensible when it comes to phonetic structures. The structures of the existing languages differ too much and are too diverse to support articulatory ease as the leading mechanism. Also, languages with similar structures fail to show similar sound changes under comparable conditions. From a sociolinguistic point of view, the principle of least effort cannot explain the underlying process of sound change, since it regularly appeared that speakers of advanced social positions within a local community who use language effectively and vigorously, are the innovators of sound changes (Labov, 1980 [82]).

Where previously the speaker and his striving for ease or intelligibility was seen as the origin of a sound change, to Ohala, it is the listener rather than the speaker who is assigned the leading role in the emergence of a new phonetic variant. He assumes that for the sake of communication, the interlocutors will pronounce and use words the way they (think they have) heard them (Ohala, 1981 [109]). Assimilatory and dissimilatory sound changes are due to misperception, and listeners develop new forms as a cause of failure in normalizing or correcting perceived speech variations. Inherent to his approach is also that sound changes are phonetically abrupt, however, he assumes that the phonetic changes might be easier to detect by outsiders than by speakers of the affected speaker community (Ohala, 1993 [110]).

If indeed based on misperception, the (abrupt) 'mini-sound change' should happen within phoneme classes, since the sound actually produced is misperceived in such a way that the listener is not aware of the misperception and his following 'misproduction' (in terms of deviating from the norm or mean): With multiple sources of information, including phonetic examples of various speakers, as well as knowledge of spelling or grammar, perception errors would be discovered. Thus, an 'accepted' misperception can only appear regularly, if the misperception and the new production do not break phonological rules.

A model of sound change based on inappropriate normalization or correction, however, is hardly able to explain more complex sound variation or change, i.e. the phenomenon of speech convergence during conversations, or chain shifts. The assumption that "speaker and hearer are interested in communicating and will pronounce words only as they have heard them (or think they have heard them) pronounced by others" (Ohala, 1981, p.197 [109]), could also hold for arising variants without misperception: As the perception experiences (input) differ for each listener, pronunciations following the individual input information will do, too. Instead of 'misperception', the speaker variation in pronunciation could simply be due to the individual auditory input. With strong connections between hearing and articulation no abnormal processes need to be included to explain variation.

The strong connections between the auditory and articulatory system have been sup-

ported by various findings, ranging from speech experiments with delayed feedback, and pseudo-word repetition tasks to neuro-anatomy. With respect to the latter, Hickok & Poeppel (2000 [50]) describe an interfacing network between auditory and articulatory representations of speech, where a sound-based representation is linked to a motor-articulatory system, as well as to an auditory-motor interface, and the auditory-conceptual interface. The motor-articulatory system is furthermore directly connected with the auditory-motor interface.<sup>1</sup> The network might establish when the child tunes his articulatory productions to the sounds of the target language. Next to playing a key-role during this critical period, it forms the basis of the phonological working memory in adults, providing access to sub-lexical speech segments (Hickok & Poeppel, 2000 [50]). The network's continuing activity, beyond the so-called 'critical period' of language development, can account for the phonetic tuning and therewith for changes in the productions of adults. Examples of tuning activity in adults are various, reflected by e.g. the phenomenon of (temporary) speech convergence during dialogs. In Pardo (2006 [112]), phonetic change and vowel variation could be linked to social interaction patterns of the interlocutors. A less temporary example are the gestural drifts that were found for the productions of bilingual speakers after a long stay in either of the two countries where one of their native languages is spoken (Sancier & Fowler, 1997 [130]). Furthermore, a longitudinal study over 50 years on the British Queen's realizations of vowels during her broadcast annual Christmas messages revealed a considerable shift over time (Harrington, 2006 [44]).

As mentioned at the beginning of this chapter, many changes in the phonetic repertoire are found to have social significance. Moreover, sound changes usually spread from groups of speakers with advanced social positions (Labov, 1980 [82]). Since there is no plausible physical explanation (i.e. misperception or individual perception) why some communities show salient changes over time, whereas others hardly do (excluding the unequal existence of hearing impairments), the reason for the different sound dynamics will be due to the unequal characteristics of the speaker groups. Previously it was concluded that the 'mini-sound change' will hardly be perceptible to speakers inside the affected group, and so a conscious adaptation of a pronunciation variant within a social group is rather unlikely. In the following section, we will consider to what extent the social structure and connections of a group can carry, adapt, or slow down innovations, and how findings from social psychology explain the impact of social relations on speech realizations.

### 6.2.3 *Social Relations, Identity and Social Cognition*

A speaker's social network is often described as a web of strong and weak ties, which interpersonal relations are defined in terms of strength, structure and density (Milroy &

---

<sup>1</sup> This dual-route sensory-motor interface is supported by neural research on the auditory switching from non-speech to speech modes (Dehaene-Lambertz et al., 2005 [25]).

Gordon, 2003 [100]). Weak interconnections within a network are seen as being more sensitive to external influences, and hence favourable to changes. Conversely, a linguistic system might be successfully supported within a network that consists of dense and multiplex interconnections.

According to theories from social psychology, intra-group differences are minimized within the social network, whereas inter-group differences are maximized (Tjafel, 1982 [146]). This accentuation of differences protects the group's value system and helps maintaining or enhancing it. Next to the value function, the accentuation of differences and the forming of stereotypes has a cognitive function: With the utilization of category membership, the complex network of social groups the individual has to deal with can be simplified.

Similar to the 'community of practice' within a social theory of learning (see Lave & Wenger, 1991 [87]), Eckert (1999) related linguistic variation to social practice and defined such communities of practice within the structure of school, where collections of people meet through common endeavours as "common goals, dreams, desires, jobs, necessities, and/ or problems, finding joint responses and strategies for dealing" (Eckert, 1999, p.40 [32]). Entering these multiple communities of practice, each scholar finds a personal path in juggling the benefits of the various communities. This principle of personal development or social conformation can easily be mapped on situations outside school.

Another theory originating from the area of social learning that explains the impact of social relations on speech is the 'Social Cognitive Theory'. It holds that environment, behavior and cognitive factors are interacting in a reciprocal relationship, and thereby are causing each other (Bandura, 1989 [5]). In being selective in their environments, people can get control over the happenings, and the social support helps in managing daily life. In anthropology studies the possibility is discussed that language evolved primarily to subserve social behavior (Adolphs, 2003 [3]). These theories recall the work of Vygotsky, who was one of the first researchers from the language field to emphasize the role of social learning. Earlier, speech was seen as the expression of thoughts, the latter being an inner process. Vygotsky argued that social interaction precedes development: Using tools such as speech to mediate the social environment, consciousness and cognition are formed through this socialization process and social behavior (Vygotsky, 1986 [162]). In a recent article, early speech learning is tied to social factors (Kuhl, 2007 [78]): In natural linguistic settings meaningful social cues like referential information (e.g. objects of reference or eye gazes) cause significantly higher attention and arousal, as well as overall increases in remembered and coded speech quality and quantity, both in perception and production. Social interaction, or even the simple presence of a human being (as opposed to a virtual human being on the tv-screen) significantly affects early speech learning. As inherent features of natural social settings, contingency and interactivity seem to be key components of speech learning. Furthermore, findings from studies with children with

autism spectrum disorder couple social deficits with early language disabilities (Kuhl et al., 2005 [79]).

Social interaction is not only a major factor during the acquisition of speech. Investigations on the ‘Chameleon’ effect, the non-conscious mimicry of various aspects of one’s interaction partner, show, that mere perception triggers mimicry, also in adults (Chartrand & Bargh, 1999 [14]). Mimicry seems to smoothen and increase the linking between the interaction partners. For interaction partners who are well-disposed towards each other, the effect is found to be even greater. A similar effect is found in speech communication, where interlocutors converge their speech patterns during conversations (Vallabha & Tuller, 2004 [149], Pardo, 2006 [112], Magnus & Nusbaum, 2007 [93], Delvaux & Soquet, 2007 [27]). Pickering and Garrod (2004 [116]) assume that a largely automatic process causes the interlocutors’ linguistic representations to become aligned at many levels in dialogue. In their model of ‘interactive alignment account’, the channels are bidirectional, and are assumed to be similar to the perception-behavior link that plays a central role in imitation according to Chartrand & Bargh (1999 [14]).

Considering the reciprocal relationship of social interaction, today, two ways of research are pursued: the representation of other minds, and the experiencing of other states of mind. The ‘Theory of Mind’ is used as a general term for research that investigates how we reason about others’ mental states, mediated by our own social rules and norms, hence on the basis of our own theories of minds (Lieberman, 2007 [90]). The Theory of Mind as domain-specific format in terms of its claimed independence from general intellectual capabilities, or as a purely theoretical model, has been challenged by an alternative approach: Experiencing others’ states of mind is associated with empathy and internally-focused processes, rather than interpretation based on theoretical concepts. The suggestion of a reciprocal relationship of environment, behaviour and cognition is supported by results from social and cognitive psychology. Most observations suggest that the performance of social actions and the processing of social stimuli are cognitively not different from the performance of other actions, or the processing of other stimuli (Hommel, 2006 [53]). Human behavior in general seems to be constantly affected and even conditioned by social interaction. Neural processes have been revealed, which mediate perception and the planning of action: Studies on the neuron system using brain-imaging techniques<sup>2</sup> suggest a neural mechanism that mediates own self-experienced multilevel knowledge and the implicit certainties we hold about others. Research on neural links between oneself and others, the *mirror*-neurons, display their role in social understanding and imitation (cf. Kuhl, 2007 [78]): During the third-person experience of action or emotion, the same structures are active as in self-experience. Besides a cognitive interpretation of what is perceived when for example recognizing emotions, experiential knowledge is generated

<sup>2</sup> such as C(A)Tscan (Computed (Axial) Tomography), PET (Positron Emission Tomography), fMRI (functional Magnetic Resonance Imaging), EEG (Electroencephalogram), and ERPs (Event-related Potentials).

by a functional mechanism (Gallese et al., 2004 [38]): Observed social stimuli are mapped directly onto motor neural structures, that were generated by the experiential knowledge of the concerning social stimulus. Thus, in dichotomy with the sensory description of the observed stimuli, also the associated internal representation of the state which is evoked during self-experience of similar states is activated. In other words, the structures that are normally involved in personal experience take a part in how we perceive and understand the behaviour or states of third persons. The impact of social relations on speech is reflected by findings concerning Broca's area, a brain area involved in both speech and the adult's mirror system (Rizzolatti & Craighero, 2004 [128], Pulvermüller, 2005 [126]).

The implications of socially constrained variation and its effect on speech perception and language acquisition led to adjustments of several (phonological) models. Featural models of speech processing that directly map phonetic features to lexical representations (with the possibility of a post-lexical phonemic level) account for the representation and processing of speech variation. Not only socially constrained variation seems to be part of the mental representations. Following e.g. McMurray et al. (2002 [95]), fine-grained subphonemic differences are preserved and of use for higher levels of speech processing. The information might be important for the perceptual system, where the fine-grained acoustic/phonetic information is correlated with information of its phonetic environment. Other studies support abstract underspecified representations of phonological features in the mental lexicon (Eulitz, 2007 [35]). Based on both findings, the newest models of speech processing are hybrid, assuming a coexistence of abstract and episodic representations, with one type of representation being dominant, dependent on individual experience: During the perception of speech, at the same time that knowledge of linguistic meaning is added, detailed episodic memory traces are created of the words that were spoken. Suggesting that accumulated episodic traces represent the mental lexicon, not all phonemic features are stored in the mental lexicon, and top-down processes influence language-specifically the perception of phonetic contrast. Following this, a complementary system is suggested, consisting of both episodic and abstract perceptions and memories that work combined (Goldinger, 2007 [42]), and, depending on factors, with one of the representations being dominant.

By stating that detailed episodic memory traces are created of perceived words, which in accumulation form the mental lexicon, the hybrid model would account for an automatic and unconsciously acquired pronunciation pattern, and for a certain degree of flexibility and changes in phonetic realizations. The latter is needed to explain speech accommodation over time. It would also explain the increasing influence of phonological knowledge on the categorization of speech sounds during language development. The different weighing possibilities of the complementary system can account for the experience-based differences found in L1 and L2 learners (see Cutler & Wagner, 2007 [23]). Furthermore, the assumption that the patterns used in production are more or less dependent on perceived

patterns, goes together very well with the previous results from neuroanatomy considering a dual-route sensory-motor interface (see Hickok & Poeppel, 2000 [50]). However, though this hybrid model seems promising, more detail and challenges have yet to be investigated.

#### 6.2.4 *Summary*

Theories holding that the speakers' representations are based or conditioned on the (frequencies of occurrence of) input, are supported by neuro-anatomical findings: With linked auditory- and articulatory systems, third-person experiences being mapped on the same areas as own experience, and brain areas that are involved in both speech and the mirror system, there is a clear interdependence of the speech input, social relations, and one's own behavior.

With speakers usually being unaware of using a particular pronunciation, it is not surprising that the acquisition process of pronunciation patterns is largely automatic. In view of code switching, as well as the stigmatizing of interlocutors, which leads to less convergence in speech, a volitional process seems to have an influence on the pattern adaptation process. Considering the processes that mediate perception and one's own behavior, stigmatization could as well be partly automatic; due to no or little contact with the rejected group, there will be no or little ability to mirror its behavior. Research must show, whether for example speakers who dislike their interlocutors, converge speech segments anyway when communicating over a long period of time. Nonetheless, the processes included in selective pattern accommodation are heavily based on social relations: Convergence goes with social understanding and the will to communicate in an optimal way, whereas non-convergence goes with social distance or reluctance to communicate. For an explanation of speech variation and change, an analysis of the embedding of the speaker within each community and the attributes of these social networks, as well as with whom the speaker identifies the most, is indispensable.

The temporary adjustment found in the speech of interlocutors in the cause of their dialogue, and the long-term speech adjustment such as for example in the longitudinal study of Queen Elizabeth's vowels are connected but have to be considered separately. (Phonetic) speech convergence during dialogues in general can be explained with the temporary storage of (acoustic) information, which, in being activated, will be the information that is primarily accessed. Following the general assumptions on memory, the temporarily converged speech patterns will have an effect on long-term memory if they are regularly reinforced. To change a speaker's speech pattern in the long run, first, the speaker's input will have to undergo a long-term change. These long-term effects of (temporary) speech accommodation have been reported with reference to the tuning of bilinguals, before and after they spent a considerable time in either of their two mother-tongue countries (Sancier & Fowler, 1997 [130]).

The fact that more or less stable interacting communities share a certain pronunciation pattern, and that convergence that arises during a conversation disappears later on, indicates that frequency of occurrence (probably in various communication partners) plays a major role in the longer adaptation and maintenance of pronunciation patterns. As speakers will choose to spend most of their time in communities they identify with, their speech pattern is very likely to have lineaments of the very community. Since there is no clear ending or critical period in the flexibility or adaptation of phonetic speech patterns, a stable pronunciation pattern will be more or less the result of a stable environment. To what extent these findings help interpreting the variation in the present data will be discussed in the following section.

### 6.3 *Interpretation of the Results of the Present Study*

In chapter 4, except for the oldest generation of speakers, the variation found in the acoustic analysis of the vowel realizations of 70 speakers was most significantly affected by the speakers' either high or low educational and occupational level (compare figures 4.22 and 4.23, p. 79).

Following the first sections of the present chapter, the two educational groups (high and low) can be seen as two different speech communities. Since the acoustic categories have been built individually to facilitate conversation, the social communities in the given data seem to be based on interactions and strong ties between more highly educated/occupied speakers, and on the other hand between the low educated/occupied speakers, with less strong contact between speakers of different educational levels. In the oldest generation, however, educational groups were not apparent in the acoustics, and hence, the contact between speakers of both levels might have been stronger and speech communities less differentiable than within the age generations thereafter.

In general, we assume that the results of our acoustic analysis, based on data all taken at the same given moment in time, reflect long-term speech patterns. However, considering the results of chapter 4, we assume that the speakers who belong to the middle and old age group are more settled in terms of their social communities than the speakers of the youngest age group. Thus, were all speakers measured again for a longitudinal research, we would suggest that our youngest speakers are the most probable age group to show a change in pronunciation.

The effect of age was of different relevance when the higher and low educated groups were surveyed separately. Unlike the high educated speaker group of chapter 4, the low educated speaker group hardly showed changes in its pronunciation pattern over time (compare the plots on the right-hand side of figure 4.25, page 82, with the plots on the left-hand side). For the low educated there was only a small linear effect of age and changes in the pronunciation of /æy/ and /au/, which was significant only for the diphthongization

of /au/ (section 4.4.4, p. 75).

Various suggestions are imaginable to explain the structure and spread of the pronunciation patterns within our limited data. Sociological research will have to show to what extent politics and economy or other factors affected the social structures in the society during the last decades, and thereby the social groupings and their pronunciation patterns. Then we could conclude if the pronunciation patterns of our sample of speakers indeed reflect the networks within Dutch society, and their changes over time.

The stable pronunciation pattern within the low educated group would for example suggest dense interpersonal relations, little sensitivity towards external influences, and hence the repression of changes over time. In view of the shared pronunciation pattern, there should have been pronounced ties between speakers of different age groups. Contrarily, within the high educated, the age differences in pronunciation might point to weak ties across age groups, suggesting less steady contact over time between speakers of consecutive generations. This would suggest the allowance of new contacts, thereby external influences, and hence changes in pronunciation.

If this is the case, we could hypothesize that the population of the high educated fluctuated much more than the population of the low educated. One could speculate that students who enroll for higher education might for example have come from various parts of the country, as well as from abroad, bringing various pronunciation patterns into the more highly educated/occupied community. With the origin of the more highly educated speakers being a rather unstable factor, the proportions of various vowel variants, and thus the pronunciation pattern of the speech community, would have been in a permanent state of flux. By contrast, lower education (and occupation) might have been available regionally, and the probability that students of various parts of the country mixed during lower education would have been comparatively small. Then, the low educated would have been much less affected by external influences and new pronunciation patterns than the high educated.

The finding of some almost significant regional traces in the pronunciation patterns of the low educated, as described in section 4.4.6, p. 84, would also suggest that the low educated speakers are a less firm speech community as a whole, but consist of several regional sub-communities. When this argumentation is transferred to the high educated speech group, where no significant regional traces were found but significant differences between the age groups, it would implicate sub-communities according to age. So, within the low educated group social sub-groups would have been structured by the factor 'region', whereas for the high educated, the factor 'age' would have been liable. However, more data and sociological research is needed to show which structures (and which changes of structures) of the Dutch society are indeed mirrored in our pronunciation data.

Within our high educated speakers (see table 4.9, p. 83), the pronunciations of the mid age group differed considerably from the pronunciations of the old age group. The pronunciation behavior of the youngest generation approached the pattern of the oldest age

group again though not differing significantly from the mid age group's pattern. Whereas the young age group might socialize with the old and the mid age group, in view of the outstanding of the high educated mid age group (36-54 years of age), thus the high educated speakers born between 1945 and 1965, we could hypothesize that these speakers were more dissociated from the previous generation than the young age group. Then, given the pronunciation differences in our data, similarly, they would have been dissociated from speakers of lower educational or occupational level. According to Stroop (1998 [140]), the 'poldermodel' in the Netherlands (see section 1.1, p. 3) triggered this behavior of separation, especially in women who profited by the growing equality. Whether the distinctiveness of this high educated speaker group has its roots in the 'poldermodel', or other societal movements, such as for example the well-known '68ers'-movement, is difficult to disentangle, and more complex investigations are needed. If we follow Stroop's argumentation of the 'poldermodel' as a supporter of women's emancipation, though in our data there were no gender effects or female precursors apparent, the women's new social strength in the early seventies might as well be reflected by the fact that they show the same behavior as male speakers. Nonetheless, in the generation before (the old generation) we found no gender effects either, nor in the generation thereafter (the young age group), and so we would suggest that the females' social strength could already have grown earlier, presumably within the war-generation, reflected by those raised before 1945. However, it could as well be the case that the females' vowel behavior has never differed significantly from the males'.

When considering our listening experiment, 90% of the listeners turned out to be high educated/occupied, and the effect of level of education thus could not be analyzed. The factor 'age group' that had been significant for the high educated speakers in the acoustic analysis had a (significant) influence on the listeners' categorization of some of the data as well. Despite the inclusion of rather few listeners per age group, the mid aged listeners judged some stimulus pairs significantly different from speakers of other age groups. The mid generation thus differed remarkably from the other generations, in acoustic terms as well as in the perception of these acoustics. The findings of the preceding sections on the auditory-acoustic linkage offer an explanation for our age-dependent speaker- and (presumably) listener behavior.

The acoustic distinctiveness of the mid age speakers' productions and the perceptual distinctiveness of the mid age listeners' behavior suggests that this age group was shaped by an acoustic input that differed considerably from the input that shaped the old or young age group. It could have been the case that social information got attached to acoustic variants in the time the mid aged group's speech was tuned. In view of the other age groups, we would then hypothesize that this social information had been irrelevant in the time before, and, given the smaller distinctiveness of the young and the old age group, and given as well the smaller distinctiveness of the young and mid age group, this social information

presumably became less important again in the tuning of the young generation.

Experiments on the perception of non-native speech sounds already showed that phonetic boundaries differ according to the listener's language background (see e.g. Ingram & Park, 2002 [55]). The perception of sub-phonemic boundaries of listeners with the same language background has not yet been a popular study objective. An investigation by van Bezooijen (2001 [151]) on what attributes younger females (aged 18-29) versus older females (aged 38-58) associate with speech samples of speakers of 'ABN', 'Polder Dutch' (see chapter 1), and two dialects, has some interesting results considering the effect of the listeners' age. To the younger, the difference between the two categories is less clear than to the older females. The younger judged ABN and Polder Dutch both as being 'normal', but ABN as less modern. The older females judged Polder Dutch as less normal and less cultivated. So, to the younger, the difference between both variants is less salient than to the older females. This suggests that the younger females are used to both variants equally, and that the only difference they perceive is the period in time that they associate with the variants. The older females, however, associate an additional factor, namely 'less cultivated' with the new variant, and thus a social attribute. In another study of van Bezooijen et al. younger and older males were included in the perception task, and the results showed that the listener generations differed in their answers towards the normalcy of the Polder Dutch variant independent of sex; contrary to the young, the old females and males judged it less normal than the Standard Dutch (ABN) variant (van Bezooijen et al., 2001 [153]). The categories of van Bezooijen's 'younger' versus 'older' listeners roughly correspond to the categories of the young versus the mid age group of our present studies. In our perception experiment, accordingly to the results of van Bezooijen, the mid generation showed categorization patterns that differed from the younger generation, suggesting that the generations are not equally sensitive to sub-phonemic differences. Though van Bezooijen's task was very different from our perception task, both results indicate the same: attributes attached to sub-phonemic variants and their perception vary at different points in time.

An interesting experiment underlines the finding that the listeners' ability to judge and categorize sub-phonemic speech sounds is influenced by the phoneme categories they experienced themselves. Recently, the effect of social information on the speech processing of merging diphthongs in New Zealand was investigated (Hay et al., 2006 [47]). In New Zealand English, the diphthongs of <near> and <square> words are merging, with the diphthong [eə] moving towards [iə]. Within the realizations of the younger, for example <air> and <ear> became homophonic [iə], and <chair> and <cheer> became [tʃiə]. Participants in this perception experiment were presented speech stimuli with and without varying (visual) social information on the speakers. When social speaker attributes were available to the listeners, the perceptually favored word of two words with merging pronunciations turned out to be biased. The results showed that – next to the influence of context- and word-specific characteristics – the accuracy in the perception task was not

only influenced by (visually) perceived speaker characteristics, but also by participant-specific factors. Though the authors note that the results are quite complex, apparently, the speaker's speech, including the merging vowel realizations, was processed not only dependent on the perceived social speaker characteristics, and participants with distinct representations of the merging vowels evoked speaker-age specific vowel distributions. To these 'unmerged' participants, perceiving a younger speaker activated less distinct vowel distributions, whereas perceiving an older speaker evoked more distinct distributions. Contrary, participants with merged representations appeared to activate less age-dependent distributions (Hay et al., 2006 [47]). Additionally, no effect of listener education on the perceptual categorization were found in contrast to the effect of listener age. Following the previous argumentation, this would imply that the different age groups were surrounded by differently equipped social-acoustic communities.

In general, where a certain sub-phonemic range of variation is socially unimportant to one generation, variation within this range can be crucial and socially structured in another generation. Given our acoustic and perception analysis and the findings in literature, we can conclude that in the same way the acoustic structure of sub-phonemic vowel realizations in Dutch differs during the course of time due to varying social constellations and influences, the sensitivity towards sub-phonemic variation differs between age groups due to own experienced social attributes.

However, interpreting the measured outcome of our speaker data remains difficult. Given the various speaker groupings, a generalization of the effects found in our sample of speakers is delimited, and many more (social) factors in addition to the ones we considered might have played a role in the distinctiveness of sub-phonemic pronunciation patterns. More research and still broader analysis is needed to identify all factors that affect the sub-phonemic tuning of production and perception. The strength of our corpus was seen in its objectivity towards the appearance of the Polder Dutch phenomenon. Now it would be interesting to see whether the same background effects that were related to the acoustic behavior of our speakers will be found in larger corpora that are specifically designed to investigate these or other social background effects in sub-phonemic speech production (with e.g. equal spreads in speaker background data and recording situation). Correspondingly, a perception experiment should be performed on the distinctiveness of the sub-phonemic variation with larger and more diverse listener groups.