Consonant Reduction in Copenhagen Danish

A study of linguistic and extra-linguistic factors in phonetic variation and change

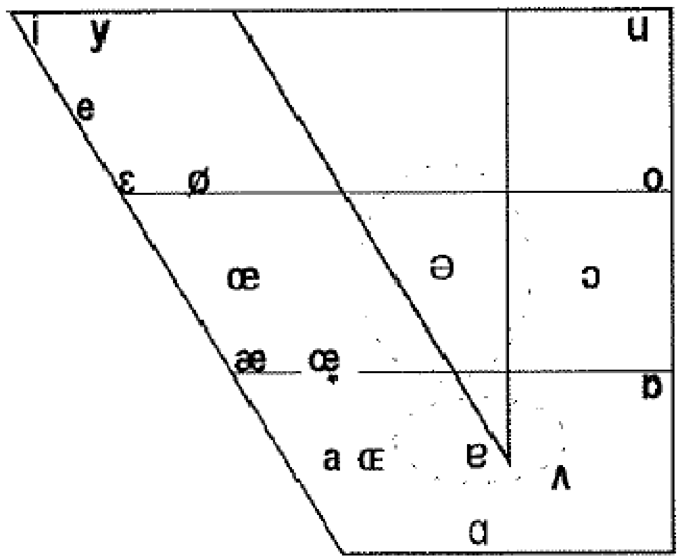
Nicolai Pharao

Department of Scandinavian Studies and Linguistics, University of Copenhagen

Contents

A note on transcriptions

The phonetic transcriptions in this dissertation use the modified version of the International Phonetic Alphabet for Danish as described in Grønnum (1998). The vowel symbols therefore have the values given in the vowel chart below.



1. Introduction to the present study

The object of the present study is to provide a quantitative analysis of a number of reduction phenomena in contemporary Copenhagen Danish spontaneous speech. In addition to providing an analysis of these phenomena and relating them to the phonetic context in which they occur, the investigation will also test some strong claims made in sociolinguistics and usage based phonology with respect to phonetic variation and change and the spread of change over time. The study will thus provide a description of selected phenomena that have not been previously described in detail for any variety of Danish, and will also contribute to the study of mechanisms of change and how studies of phonetic variation and change may contribute to our understanding of the mental representation of spoken words in memory.

1.1 Variation and change in Usage Based Phonology

The major hypotheses about the influence of word-form frequency on phonetic variation in the usage based framework that is to be investigated in the present study is the claim that reduction affects high frequency items before low frequency items as put forth in Bybee (2002), and generally supported by studies of corpora of speech from a variety of languages and has also been observed as a general tendency in reductive changes arising in Copenhagen Danish throughout the 20th century (Brink & Lund (1975), pp. 729-730). The pervasiveness of frequency effects leads Bybee (2002) to posit a conceptualization of the lexicon that is radically different from that of Generative Phonology. Rather than stripping away all redundant aspects of word form from lexical representation, and then later supplying this information by phonological rules, Bybee suggests that "[l]inguistic regularities are...schemas or organizational patterns that emerge from the way that forms are associated with one another in a vast complex network of phonological, semantic, and sequential relations." (p. 21).The phonological rules of Generative Phonology, then, are convenient abstractions over the vast and highly redundant lexicon. An example that supports this conceptualization of phonology as a part of the human faculty for Language is the variation found in the application of what would traditionally be termed the rule of schwa deletion in American English (taken from Bybee (2002) pp. 40-42[1]). In an analysis of spontaneous speech, Bybee finds that schwa deletion is obligatory in a high frequency word like 'every' which is always realized as [ˈɛvɹ̞ɪ] and never [ˈɛvəɹ̞ɪ]. Words of relatively lower frequency have a strong tendency for the schwa to be deleted, but with the following sonorant becoming syllabic, e.g. a word like 'memory' is most often realized as [ˈmɛmɹ̩ɪ] rather than as [ˈmɛmɹ̞ɪ] or [ˈmɛməɹ̞ɪ].

[1] Phonetic representations are mine; Bybee (2002) gives orthographic forms – NP.

Words of very low frequency are never affected by schwa deletion in the corpus, e.g. 'mammary' is always realized as [ˈmæməɹɪ] and never as [ˈmæmɹ̩ɪ] or [ˈmæmɹ̥ɪ]. This example shows that a variable phonological process, schwa deletion before sonorants in post-tonic syllables, is guided by frequency of use of the words in which the structure is embedded: the more frequent the word, the more likely the application of the process. While this process can be abstracted and formulated as a rule with variable application, Bybee (2002) maintains that it is more properly modeled as a property of individual words or word-forms. This notion has been taken up in recent psycholinguistic approaches to the modelling of the mental lexicon couched in Exemplar Theory. In these models (e.g. Johnson (1997) and Pierrehumbert (2002)) the mental lexicon *is* an immense memory storage of all tokens of words encountered and produced by the individual language user. Each word-form is represented by a cloud of tokens of that word-form with all of their phonetic details retained. Thus variability which is closely associated with the word-form's prosodic and segmental context is not abstracted away, nor is information about the gender, age, geographical origin or even mood of the speaker producing the specific token. There are more aspects to the Exemplar Theoretic account of lexical storage which I will return to later, but an important aspect of these models in relation to sound change is that the frequency effect which motivates the exemplar based model must be observable at the level of the individual language user, in order for word-form frequency to be explanatory and relevant in the description of mental phonology (Pierrehumbert (2001) p. 148).

The claims to be tested in the present investigation are that

1) reduction affects high frequency word-forms more often than low frequency word-forms,

and that

2) high frequency word-forms are affected before low frequency word forms.

Furthermore,

3) this effect of frequency should be observable at the level of individual speakers

This will be done by including word-form frequency as a factor in the statistical modelling of consonantal variation in two corpora of spontaneous Copenhagen speech.

1.2 Studying on-going sound change – sociolinguistic approaches

The study of sound changes in progress has been conducted mainly within the tradition of quantitative sociolinguistics since Labov (1966(2006)). Within sociolinguistics, the so-called "free" variation of mainstream phonology is treated as gradient but regular properties of the speech of sub-groups of a speech community. In this field, identification of an on-going change was until recently based exclusively upon the comparison of the pronunciation of different generations of speakers. Such studies are known as apparent time studies and rest upon the Apparent Time Hypothesis (Labov (1994)). The last decade has seen an increase in longitudinal studies of phonetic variation in adults. In studies like these, the pronunciation of groups of adults is studied at (at least) two different points in time typically spaced two or more decades apart. In sociolinguistics, these studies are referred to as real time studies. The subsequent analysis of the phenomena studied at the time of the first recording can either be done using the exact same people, a so called panel study, or people from similar social backgrounds, a so called trend study.

The study of language change lies at the heart of classical variationist sociolinguistics. From the very beginning of the study of social influences on phonetic variation in U.S. English, the objective was also to provide an empirical study of linguistic change and its correlation with social characteristics of speakers (Labov (1972)) and (1966(2006)). Due to the lack of recorded speech material from the areas studied (Martha's Vineyard and the lower east side of Manhattan in New York City), these initial investigations of on-going sound change relied on evidence from apparent time investigations. Labov (1966(2006)) states that "[b]y studying the differences between the linguistic behavior of successive age levels in our sample, we can make inferences about linguistic change." (p. 200). The limitations of this methodology were of course also recognized immediately, and the validity of the inferences about sound change was acknowledged to be restricted to circumstances where "1)...there are no differences between older and younger speakers which are repeated in each generation; and 2)...the older speakers remain isolated from the effects of the language used by younger speakers." (pp. 200-201) In other words, if we are to say that the difference in pronunciation of the vowel in words like 'ret, frem' /rɛd frɛmˀ/ *straight, forwards* that we observe between 20 year old speakers and 60 year old speakers in 2008 reflects a sound change that has progressed within the past 40 years, we must assume that the speech of the 60 year olds today is, at least with respect to their pronunciation of vowels, identical to their speech in 1968.

One of the arguments cited to support the apparent time hypothesis is the notion of the critical period, borrowed from research in language acquisition and first proposed by Lenneberg (1967). The critical period denotes the time in a speaker's life during which he/she must be exposed to a given linguistic variety in order to acquire native competence of it. This period ends some time during adolescence and presumably most speakers do not change their linguistic systems significantly after this. Speakers may adopt novel lexical items or may change their realization of speech sounds as a function of aging effects on vocal tract physiology, but they should not participate in processes of sound change that they have not already engaged in once they reach early adulthood.

The validity of such a claim, that the difference between generations reflects a change in progress, can only be firmly established through the use of real time studies, either by re-recording the panel of speakers used as informants before, or by recording a socially identical group of speakers in order to examine the trend of the development. The assumption that speakers do not change their phonological systems, including here the phonetic implementation rules of the contrastive phonemes, is essentially a hypothesis about the plasticity of the language faculty of the individual speaker. This assumption is being called into question by recent evidence from a small number of real time studies of changes in the phonetic implementation of phonological systems. While there is no evidence to date of speakers participating in changes that lead to mergers, i.e. abolition of contrasts, there is a gathering amount of evidence that speakers may change their realization of one or more phonemes rather drastically throughout the lifespan. For example, Blondeau & Sankoff (2008) show that a minority of adult speakers of Montreal French shift their productions of /r/ from an apical to a uvular pronunciation. While they do not become categorical uvular-/r/ users, they shift significantly in the sense that the uvular variant becomes the predominant one in their speech. A similar finding was presented for lowering of /ɛː/ by Yaeger-Dror (1994) in an analysis of Quebec French. Yaeger-Dror (1994) describes the development for 13 male speakers (out of a sample of 40) and shows that these 13 individuals continue to change their realizations of original long /ɛː/to [æː] i.e. they participate in an ongoing process of lowering well into adulthood.

Sankoff (2006) summarizes the evidence from real time trend and panel studies conducted until the early 2000s, and highlights the following characteristics on the basis of 17 studies:

1) The majority of speakers do not change

2) The speakers that do change, do not do it to as large an extent as the overall community

3) The speakers that change, change in the direction of the community

4) The changers are most often young adults – although that term covers the age range of speakers in their late twenties up to fifty years of age.

Real time studies seem to have mainly supported the validity of the apparent time hypothesis, or at least that is the conclusion reached by the majority of researchers working with real time data. But it is interesting to note that a subset of speakers do change their speech. The tendencies outlined above can be tested by conducting a real time investigation of the changes in consonant articulation in Copenhagen speech. Thus, the claim that speakers remain stable will be tested by comparing the speech of individuals as attested in recordings made in the 1980s to recordings of the same speakers made in 2005 and 2006 by analysis of the speech of middle class speakers from Copenhagen contained in the

LANCHART corpus. Some of the tendencies outlined above will be tested, but the present study will combine the claims made in quantitative sociolinguistics with the claims from usage based phonology that the spread of change is governed by word-form frequency. As mentioned above, the frequency effects should be observable at the level of individual language users. This means that for the speakers who do participate in on-going change throughout adulthood, the distributional patterns of variants should either

1) show an increased use of innovative variants in words of high frequency in the new recordings compared to the original recordings

or

2) show a tendency for innovative variants to also occur, in the new recordings, in words of lower frequency than in the original recordings.

1.3 Major goals of the investigation

By combining the approaches from usage based phonology and quantitative sociolinguistics the following overall goals of the investigation can be set up as follows:

1) Describe the relationship between phonetic and extra-linguistic factors in processes of consonant reduction in contemporary Copenhagen Danish

2) Investigate on-going change in the consonantal system of contemporary Copenhagen Danish

3) Investigate the role of language use in synchronic variation and in the spread of an on-going sound change

4) Investigate the degree to which individual language users change during the life span

5) Determine the role of language use in the progress of a change in the speech of individuals

If changes are found in the behavior of individual speakers and if such change is shown to be regulated by language use, as reflected by the correlation of word-form frequency with the spread of change, the findings will lend support to models of the lexicon couched in the framework of Exemplar Theory.

2. METHODOLOGY

This chapter describes the methodological aspects of the present investigation. First, a description of the corpora used in the investigation is given, next the nature of the phonetic data analysed is described, and last a brief introduction to the statistical methods that are used in the investigation.

2.1 DESCRIPTION OF THE CORPORA USED

2.1.1 The DanPASS Dialogues

The corpus of Danish phonetically annotated spontaneous speech, the DanPASS corpus, was designed and collected by Nina Grønnum at the Linguistics Laboratory in the Department of General and Applied Linguistics at the University of Copenhagen. The corpus consists of a set of 18 monologues and 11 dialogues, comprising more than 10 hours of running speech and more than 70.000 word tokens. The monologues were recorded in 1996, and the dialogues were recorded in 2004.

The DanPASS dialogues, like the entire DanPASS corpus, contain annotation of phenomena at several levels of representation, stored in Praat TextGrids. The basic level is the orthographic representation of the speech contained in the associated sound recording. This representation is segmented and time aligned with the sound signal at the word level. In addition to the standard orthographic representation, this tier contains information about pauses, hesitations and word stress, with stressed tokens containing a comma before the letter representing the stressed vowel. For example, an unstressed token of the first person singular pronoun is transcribed 'jeg', whereas a stressed token is transcribed 'j,eg'. Associated with this orthographic representation are tiers containing a semi-narrow phonetic transcription at the word and syllable level, a relatively abstract (morphologically informed) phonological representation, annotation of the pitch patterns of stress groups, and an annotation of perceived pitch course for individual prosodic phrases (Grønnum (2009)).

The transcription and prosodic annotation of the recordings were done by a group of phonetically trained transcribers. Each recording was transcribed by two transcribers, who subsequently compared their transcriptions, and disagreements were resolved in conference with the project leader. In addition both the segmental transcription and the prosodic annotations were proof read by the project leader (Grønnum (2009)). Only the proof read transcriptions and annotations are contained in the corpus. Therefore, it is not possible to examine inter-transcriber agreement or intra-transcriber consistency in relation to the variability in the speech signal. Studies of inter- and intra-transcriber agreement show that even experienced phoneticians do not agree completely with each other or with themselves (cf. Shriberg & Lof (1991)), it is a pre-requisite for using the DanPASS corpus that this source of error be accepted. However, the many stages that the transcriptions have gone through clearly reduce the risk of error in the transcriptions, and the corpus certainly constitutes the most richly and carefully annotated corpus of spontaneous spoken Danish available.

The entire corpus has also been tagged with part-of-speech labels (henceforth: PoS-tags). The PoS-tagging was done automatically using a tagger developed by Peter Juel Henrichsen (2002). The automatically generated PoS-tagging was subsequently hand corrected for the *entire* corpus. The PoS-tagging is given in two versions: a detailed representation conforming to the conventions of the Parole part-of-speech tag set, developed for the analysis and parsing of several European languages, and a reduced version containing information about word classes only. A few distinctions have been added to the set of PoS-tags, notably the personal pronouns. For a full list of additions see Grønnum (2006) and for further information about the changes made to the initial automatic PoS-tagging see Grønnum (2009), p. 598.

Each TextGrid file in the DanPASS dialogue corpus represents the speech of *one* participant only. Furthermore, each dialogue is split into 5 sections: in two of them, the speaker is guiding his or her partner through a map, and in another two, the speaker is following the directions given by his or her interlocutor. In the fifth section, the speaker reads a list of words pertaining to the map task.

Participants were paired with a person that they already knew well in the hope that this would make the conversation between them more natural given the restrictions imposed by the experimental setup. The task itself was a replication of the HCRC Map Task (Anderson et al. (1991)) in which two speakers take turns guiding each other around a map of an imaginary "landscape". The maps differ from each other by not containing the exact same objects and places, although the majority of items are identical. The purpose of the discrepancies was to elicit discussion between the guide and the follower. While the task itself arguably is a fairly close approximation to an everyday conversational situation, the circumstances of the recording were not. Participants were placed in separate rooms, and communicated over headsets with built-in microphones (for technical specifications, see Grønnum (2006)). In this respect the DanPASS dialogues resemble telephone conversations, albeit with vastly superior sound quality. While the recording circumstances are highly unusual compared to everyday interaction, the corpus contains a sample of speech that is perhaps more readily comparable across speakers than those contained in corpora of sociolinguistic interviews like the LANCHART corpus described below. The situation is exactly the same for every participant; the only situational factor that varies is the identity of the interlocutor, but not the relation between the interlocutors in each pair. This makes the likelihood of any significant stylistic variation very small, regardless of whether this is conceptualized as a degree of formality, or the amount of attention paid to speech (like Labov (1966[2006], pp. 58-65), or even categories like genre or stance-taking like those employed in LANCHART (Gregersen, Nielsen & Thøgersen (fthc.)). The genre in the DanPASS dialogues is the same throughout, the exchange of directions around a map, which in the LANCHART corpus would be classified as exchange of information. The individual speaker alternately is the giver and recipient of the information, but the dialogues very rarely diverge from this topic. So while the restrictions imposed by having a specific task to perform as well as being recorded in rather unfamiliar circumstances may lead to the classification of the speech contained in the DanPASS dialogue corpus as non-scripted rather than casual, it still provides the best possible starting point for the investigation of the phonetic characteristics of Danish spontaneous speech. The span of variability may be, and quite probably is, limited compared to that which takes place in completely unrestricted interaction between the same (types of or even actual) speakers, and even when compared to the relatively free setting of the sociolinguistic interviews of the LANCHART corpus.

The DanPASS dialogues, then, readily provide the opportunity to investigate the variability in the phonetic realization of (morpho)phonemes as registered by the transcribers of the recorded dialogues. They also allow for investigation of the contribution from 1) stress, 2) segmental environment, 3) morphological categories, and 4) position in the prosodic phrase to the patterning of variation for individual phonemes. Below is an overview of the informants participating in the DanPASS dialogues giving their year of birth and gender.

Table 2.1 – Informants in the DanPASS Dialogues

Informant number	**year of birth**	**Gender**	**Informant number**	**year of birth**	**gender**
001	1976	M	015	1977	M
002	1945	F	016	1974	F
003	1976	M	017	1973	M
004	1957	F	018	1974	F
006	1966	F	020	1946	F
007	1963	M	021	1950	M
008	1957	F	023	1942	M
009	1955	M	025	1977	M
010	1974	F	027	1962	M
012	1982	F	031	1968	M
013	1928	M	033	1974	M

Since all 22 participants in the corpus are academics, either students or lecturers, and since there is no information about the socioeconomic background of the participants and their families, the dialogues do not provide an opportunity to study social factors in the patterns of variation that may be found across speakers, except for gender. The sample is, however, stratified for age, allowing for a small scale apparent time investigation of the tendencies for change in the production of consonants. This apparent time study may serve as input to a real time trend study by comparing the original recordings from the LANCHART Cph Corpus (cf. next section) with the DanPASS dialogues. This is also why the investigation is restricted to the dialogues in the corpus: all speech in the recordings that will be used to provide real time data for the variation of consonants comes from dialogues.

2.1.2 The LANCHART Cph Corpus

The LANCHART Corpus consists of a large collection of sociolinguistic interviews. The corpus contains both sound recordings and orthographic transcriptions of the interviews stored as Praat TextGrids, one for each interview. The orthographic transcription is time aligned with the sound recording at the utterance level, i.e. beginnings and ends of utterances in the transcriptions coincide with their beginnings and ends in the sound recording, but individual words are not aligned with the sound recording. The entire corpus contains interviews from 7 different regions of Denmark, but since the present investigation is concerned only with the phonetic variation found in the speech of Copenhagen, the remainder of this description is confined to the details of the Copenhagen sub-corpus, henceforth the LANCHART Cph Corpus.

The LANCHART Cph Corpus contains interviews with 22 informants from 1986 and 1987, collected in the Copenhagen Project on Urban Sociolinguistics (Gregersen & Pedersen (1991)), and re-interviews with the same informants from 2005 and 2006. What I am calling the LANCHART Cph Corpus actually only consists of speakers that have been classified as middle class (Gregersen (fthc.), pp. 10-11), with 12 of them belonging to generation 1, i.e. speakers born between 1942 and 1963, and 10 of them belonging to generation 2, speakers born between 1964 and 1973. Within each generational group the speakers are also stratified for gender, with half of each group being men and the other half women. The labels Generation 1 and Generation 2 merely reflect that the former constitute (a sample of) the oldest speakers interviewed in the original Copenhagen Study on Urban Sociolinguistics (Gregersen & Pedersen (1991)), i.e. it is in no way intended to imply that these 12 speakers constitute a sample of what could be termed the first speakers of a distinctively Copenhagen dialect. Not only has Copenhagen as an urban center existed for centuries, but a distinctive Copenhagen speech variety has also existed for a long time (cf. Brink & Lund (1975)). Thus, the labels are merely practical designations to be used in the apparent time analysis of the original recordings of the LANCHART Cph Corpus. Below is a table of the informants with information about year of birth and gender.

Table 2.2a – Informants in the LANCHART Cph Corpus, Generation 1 (older)

Informant code	year of birth	gender
LAL	1949	F
CEL	1954	M
PFR	1955	F
EAF	1958	F
MJE	1957	M
LKR	1962	F
CNI	1959	M
TNI	1952	M
ASA	1951	F
HTH	1957	F
MFL	1958	M
PTK	1946	M

Table 2.2b – Informants in the LANCHART Cph Corpus, Generation 2 (younger)

Informant code	year of birth	Gender
MPT	1973	M
MIP	1971	M
NOR	1971	M
JOR	1968	M
SKO	1971	M
KKJ	1973	F
ASH	1972	F
JJE	1971	F
PKJ	1968	F
DBE	1967	F

In contradistinction to the dialogues in The DanPASS Corpus, the LANCHART Cph Corpus consists of rather lengthy recordings of almost completely free conversation. The structure of the interviews is quite similar when the original recordings are compared to the new recordings. Naturally, this was done intentionally, in order to ensure as much comparability as possible between the two recordings. The beginnings of the interviews differ in that the re-recordings all start with a rather formal questioning of the informant by the interviewer. This questioning concerns the gathering and confirmation of social background data like year of birth, occupation, marital status and the like. While serving some practical purpose, it also provides for a maximally formal situation which the interviewer and informant together may attempt to deviate from as the interview progresses. Since an integral part of the LANCHART project is to investigate the relationship between language use and situational setting, it is of course important to the interviewer to attempt to steer the "mood" of the conversation away from the formality of the questioning and in the direction of a more freely flowing exchange, rather than a strict question-and-answer session. In the original recordings from 1986-87, the interviewers also sought to manipulate the "mood" of the conversation with the purpose of eliciting as many shifts between non-casual and casual style as possible (Albris (1991), p. 49, pp. 54-5)). But the interviews did not begin with a fixed line of questioning nor was any other attempt made to ensure complete comparability across informants with regards to the content of the beginnings of interviews. Naturally, the extent to which such shifts take place in the structure of the conversation throughout the recordings varies from interview to interview. The structure has been coded and analyzed according to the conventions described in Gregersen, Nielsen & Thøgersen (fthc.), and this analysis is stored in the transcripts of each interview. The relation between situational setting and language use will not be included in the present investigation. However, the codings may serve the purpose of delimiting passages which are maximally comparable, i.e. as similar as any two conversational situations get when they are spaced 20 years apart, so that the differences that are found between original recording and re-recording are less likely to be attributable to stylistic differences (in a wide sense), and more likely to reflect differences over time. As mentioned above, the recordings contained in the DanPASS dialogues may all be classified as the exchange of information throughout, and therefore the analyses carried out on the LANCHART material should also be limited to this particular discourse context.

The transcripts of the sociolinguistic interviews in the LANCHART Cph Corpus are not as richly annotated as those of the DanPASS Dialogues. While the analysis of the structure of the interviews is quite elaborate, only selected phonetic phenomena are directly annotated in the transcripts, namely the sociophonetic variables described in Gregersen, Maegaard & Pharao (fthc.). In addition, however, both an automatically generated part-of-speech tagging as well as an automatically generated phonetic annotation of the entire corpus is provided and contained in each transcript. The PoS-tagging was again generated by Peter Juel Henrichsen, using the same tagger that was used on the DanPASS Corpus, but the tagging *has not* been hand corrected in the LANCHART Cph Corpus. Even a cursory inspection of the correctness of the automatic tagging reveals several points where the tagging is unreliable. And naturally, the automatically generated phonetic annotation does not provide an accurate transcription of the speech signal contained in the recordings, since it is based on grapheme to phoneme correspondences. The phonetic transcription exists mainly in order to provide a basis for searching for particular abstract segments without having to bother with the ambiguities inherent in basing such a search upon the orthography. The convention for the automatic phonetic annotation is a highly distinct and conservative pronunciation, e.g. there are no schwa-assimilations and length is preserved on underlyingly long vowels before [w j ð], an environment where long vowels have been subjected to shortening for more than a century, although with the shortened variants becoming dominant only recently (i.e. in the latter half of the 20th century) (Brink & Lund (1975), pp. 221-9). While not as abstract as the phonological representation in the DanPASS dialogues, the characteristics of the convention applied make it easy to search for occurrences of particular phonemes as they would occur in a morphologically informed phonological representation.

The LANCHART Cph Corpus, then, provides a possibility for further investigating the variation in the Danish consonants as it occurs in free conversations in a socially homogeneous sample of adult Copenhagen speakers. The following types of investigations may be conducted:

1. Apparent time study of variation and change in the original recordings
2. Real time panel study of the variation through comparison of original and new recordings

The new recordings made in 2005 and 2006 for the purpose of the real time study of linguistic variation and change unfortunately do not contain recordings with a Generation 3 for Copenhagen, i.e. informants of the same age as Generation 2 in the original sample from the mid 1980s. Therefore it will not be possible to study to whether individual speakers who significantly change their proportions of innovative variants of a consonantal variable between the original and the new recordings do so in the direction of change for the speech community. This also means that it will not be possible to study whether such life span changers shift their proportions of an innovative variant to a lesser extent than younger speakers. But it will be possible to study how many of the speakers in the sample that do continue to participate in phonetic change during the lifespan and to study changes in the influence of phonetic, extra linguistic and usage based factors in such cases of life span change.

The LANCHART Cph Corpus does not easily allow for the same level of detail with respect to linguistic factors and their influence upon the patterns of variation. However, the results from the analysis of the DanPASS corpus can be used to guide the studies based on the LANCHART Cph Corpus, in order to control for the factors that have been found to be relevant for the variation in consonants.

The corpora thus constitute samples of contemporary spontaneous Copenhagen Danish speech of adult middle class speakers. This variety is chosen because it forms the background for the most extensive descriptions of Danish phonology (e.g. Basbøll (2005), Grønnum (2005)), and because it in effect has been the standard reference variety for studies of the speech of other regional and social groups (cf. Kristensen & Jørgensen (1994), Nielsen & Nyberg (1993)).

2.2 Why Middle Class Speech?

Central to the study of phonetic variation and change in Danish since the 1970s is the history of Standard Danish by Brink and Lund from 1975, which focuses on the relation between the sociolects of Copenhagen and their influence on changes in other parts of the country, i.e. on other dialects of Danish. Brink & Lund (1975) convincingly shows that the majority of phonetic changes during the first half of the 20th century have originated in the speech of the working class and since been adopted by the middle class in Copenhagen and by speakers in other towns. This latter pattern has to some extent been corroborated by other studies of variation, most clearly in the study of the Næstved regiolect in Kristensen & Jørgensen (1994). While previous investigations suggest that working class speech would be the most likely type of data in which to find innovations, they also indicate that any innovations found in middle class speech are likely to pertain to a wider range of the population than the speech of working class speakers. Similarly, while many studies indicate that phonetic innovation originates in the speech of adolescents, and hence would suggest that one should study this younger age generation for signs of on-going change, this finding also indicates that any sign of change found among older speakers will be innovations that again pertain to a larger group of the population. All in all, by studying phonetic variation among middle class adults, it is less likely that the results will exaggerate the degree of change.

2.3 THE PHONETIC DATA

The analyses of phonetic variation presented here are all based on phonetic transcription of the acoustic speech signal. In the case of the analyses of the variation attestable in the DanPASS Dialogues, entire conversations have been transcribed (see above and Grønnum (2009)) In the case of the studies based on the LANCHART Cph Corpus, transcription has been restricted to coding of tokens of specific phonetic variables with a limited set of variants in each case. Generally, each phonetic variable 'X' has been classified as being manifested as one of two variants, 'y' or 'z'. In cases of doubt coders were allowed to use either an intermediate classification of the token, 'y/z', or as a realization belonging to a different continuum than the one intended in the variable, indicated by an asterisk. The asterisk was also used to encode tokens where the realization was either imperceptible or where the coder felt too uncertain about the token to provide a classification, mainly due to circumstances of the recording, e.g. laughter, background noise, or overlapping speech.

Auditory classification is difficult because we in daily communication focus on content rather than form. Obviously, we also notice differences that convey non-lexical meaning, like speaker dialect (Clopper & Pisoni (2006)), and affiliation with different subgroups of a speech community (Kristiansen (fthc.), Maegaard 2008), Campbell-Kibler(2007)). If we did not, the sociophonetic endeavor would be futile from the outset. In a sense, our hearing is to a large extent "phonological", we listen for the features that are important for extracting the lexical content, while ignoring others. While correlations between speech and non-lexical or social meaning is obviously part of speakers' knowledge about their language (cf. Docherty & Foulkes (2000)), it is difficult to become consciously aware of these differences. The ability to reliably identify non-lexical contrasts requires training.

In the case of DanPASS, phonetically trained student transcribers were used and disagreements between transcribers were settled in conference with an expert. In the case of the LANCHART transcriptions, the classifications were always done by me and a student helper trained in linguistics and phonetics at the BA-level. For the LANCHART transcriptions, all cases of disagreement between me and the other coder were discarded from the dataset used in the statistical modelling.

Acoustic measures have not been included in the present investigation. While this means that subtle variation may be lost, and that the classification is not as objective as one based on physical measures, it does increase the likelihood that the variation described is perceptible, although it may not be obvious to language users that are untrained in phonetic transcription. Neither transcription nor coding of phonetic variables are tasks that are identical to the on-line comprehension of speech. This is partly due to the different circumstances of transcription and on-line speech processing. In daily communication, utterances can only be heard once. When performing transcription and encoding of variables in recordings, transcribers have the possibility of hearing particular stretches of speech as many times as they want to and with as much contextual information as desired – possibilities that are not available in live conversation. But the difference is also due to the different purposes between on-line speech comprehension and transcription. As mentioned, the primary goal of daily conversational interaction is, arguably, the conveying of content. Furthermore it is certainly the content that most listeners pay *conscious* attention to. In transcription and coding, listeners are required to pay conscious attention to the subtle variations in the articulation of the recorded speech. This is precisely why these procedures are useful: they provide the opportunity for registering details of the speech signal to which we are exposed in daily on-line communication, although we may not consciously notice them.

The method of auditory classification may not be as objective as that of acoustic analysis, nor as fine grained (when a high degree of inter-coder reliability is desired). Particularly in the case of encoding a phonetic variable within a continuum by means of discrete variants, there is going to be loss of information. However, the process is less time consuming than acoustic analysis, and may give a good first indication of the quantitative distribution of segmental variants. Once patterns of variation have been established with reference to the perceptual categorization involved in transcription and coding, acoustic analysis can be used for further, more detailed investigation of the phenomena. Although it is beyond the scope of the present investigation to perform acoustic analysis of the phenomena studied, I would like to emphasize that it is not because I think that sociophonetic research should be restricted to clearly perceptible differences in the speech signal. Rather, any difference, whether perceptual or acoustic, that can be seen to be transmitted must be relevant for sociophonetics, in particular when the aim is to study possible on-going change. Nor do I think that the fact that the variation has been detectable by transcribers and coders necessarily ensures that it is a difference which is used, at whatever level of consciousness, by language users to convey a particular social meaning. What the correlations between perceptually identified variants and socially defined groups of speakers *can* tell us, is that speech variation is patterned and non-random to the extent shown by the correlations. But even the strongest correlation between a phonetic variant and a social category like e.g. 'male' or 'working class' does not tell us that the variant in question "means" maleness, or that it "means" working class. Not only is the link between speech variation and social meaning probably more complex than such an interpretation would suggest, but it is also possible that the correlation is caused by a factor that happens to co-vary with maleness, both in the sample studied and in the speech community. If all male working class speakers in our sample happen to also be passionate soccer fans, it is entirely possible that a particular segmental variant, raised /a/, say, is "used" to convey the identity of being a soccer fan. Perhaps this segmental variant is also used by female soccer fans or by middle class soccer fans, but we cannot know this on the basis of classifications of speech production patterns alone. Any study that aims to analyze the social meaning of variation must also include perceptual reactions to the patterns of variation, that have been uncovered in the study of the speech production habits of socially defined groups of speakers. This, too, lies beyond the scope of the present investigation.

3. Statistical methodology

The analyses of the phonetic variation of selected consonantal variables in contemporary Copenhagen Danish spontaneous speech will be presented in the form of statistical modelling. Because the analyses are based on distributions of discrete variants as determined by auditory classification and segmental phonetic transcription, the modelling will be carried out using multiple logistic regression. This method is widespread in the international sociolinguistics literature, where it has been used to model and evaluate the concept of the variable rule (Sankoff & Labov (1979)). Dedicated software for logistic regression analysis of linguistic data has been used, predominantly VARBRUL, GoldVarb and more recently Rbrul (see Johnson (2009) for an overview). However, a survey of the literature on Danish sociolinguistics reveals that logistic regression has never been employed in the study of linguistic variation in any variety of Danish. Therefore, this chapter gives a basic introduction to the statistical methods that will be used to discover patterns in the distribution of variants, as well as general guidelines for interpreting the patterns on the basis of the statistics. I will give an outline of the steps of the analyses in general terms. The results of these steps will be presented in the analyses of each phonetic variable. For illustration purposes, I show an analysis of preliminary data on the variable (wəð), that is deletion of distinct [w] before [ð̞], in this chapter. The data used here are only given to make the introduction less abstract. The actual analysis of the variable is presented in the studies of apparent and real time change

3.1 Why logistic regression?

Like other methods for the analysis of categorical data, logistic regression allows us to quantify the relationship between the behavior we are studying and the factors we have hypotheses about. And like other methods used in categorical data analysis, it is based on the examination of probabilities. Unlike other techniques, logistic regression (and regression in general) does not focus on examining whether distributions are random or not. This is, of course, a part of the regression analysis, and an important one, but where tests like Chi-square and Fisher's exact test are designed to determine whether it is likely that the distribution of categorical responses in relation to one particular background factor is random or not, regression techniques attempt to predict the outcome of the process that is being modeled on the basis of probabilities. Another drawback of statistical tests like Chi-square and Fisher's exact test is that they do not include a measure of how different the observed distribution is from a random distribution. The level of significance alone does not provide an accurate measure of the size of the difference. While we are interested in knowing whether an observed difference is likely to be linked to a background factor, we also very much want to know something about the magnitude of the effect of that factor, or in other words, the strength of its influence. Such a measure is given directly in regression analysis. Logistic regression, then, is useful for modelling probabilities, just like the traditional tests for distributions of categorical data. In keeping with the view that the explanation of phonetic variation does not necessarily involve finding absolute conditions for when or where a particular segmental variant occurs, logistic regression models the probability that a segmental variant will occur given one or more factors. It is of course an empirical question whether there are absolute conditions on the occurrence of any given segmental variant, but the stance taken here is that there are *tendencies* for particular segmental variants to occur in specific contexts and at particular times, rather than unequivocal rules for the application of reduction processes. That is, contextual factors, linguistic as well as extra-linguistic, may promote different types of variation, but this influence from the factor is not absolute or deterministic: a speaker may choose, consciously or subconsciously, to override what can be construed as default implementation mechanisms of the speech production system. In other words, I do not expect to be able to set up categorical rules for phonetic or phonological processes. Rather, the goal of this study is to find the probability of particular processes applying given a set of circumstances, and thereby to set up gradient or variable rules, i.e. rules that are inherently variable.

For example, while the presence of a following [ð̩] might overwhelmingly invite deletion of [w], there is no presupposition that this will always happen, i.e. that [w] will always disappear in casual speech when followed by [ð̩]. First of all, there are numerous other factors which are hypothesized to be involved, such as speech style, articulation rate, preceding vowel quality and word frequency. These factors may themselves have an adverse influence on the deletion of a particular token of a [w] irrespective of whether a [ð̩] follows or not, in a sense overriding the tendency from this component of the context of the individual tokens of [w]. However, even in a situation where all (known) factors are in a state where we would expect deletion, there is still the possibility that the speaker acts differently. In other words, if we see the factors as a set of contextual parameters, we should not expect to find exact settings for these parameters that will *necessarily* produce a specific outcome. We should only be able to find which settings *favor* particular outcomes.

Therefore, the hypothesis here, following the tradition of quantitative variationism (e.g. Labov 1966(2006)) and usage based phonology, is merely that a given factor invites the application of a process, not that it demands it. Factors may act in concert and in effect "boost" the articulatory processes that they promote independently. This is of course also an empirical question, but, continuing the example, it is entirely plausible that other factors which are hypothesized to promote [w] deletion, like high articulation rate and high word frequency, may increase the likelihood of deletion of a [w] followed by a [ð̞]. It is precisely the object of the quantitative study of contextual factors to examine such relationships and, if possible, to tease apart the contributions from individual factors. This is possible with multiple logistic regression.

3.2 ANALYZING DISTRIBUTIONS OF CATEGORICAL DATA

When dealing with categorical data, or categorical representations of non-discrete phenomena like speech, there are restrictions on the kinds of statistical tests that can be used in order to explore the possibility that an observed distribution is unlikely to have arisen randomly. Such tests are desirable in cases where we want to be able to generalize our findings beyond the sample of speech and speakers we have analyzed, and when we want to be able to distinguish between distributions that are highly likely to be random and in that sense are hardly patterns at all, distributions that tend toward a regular skewing with respect to a particular background factor and thus form a kind of pattern, although they do not reach the pre-specified level of significance, and patterns which are significantly different from each other at the conventional threshold of significance, i.e. $p < 0.05$. Whereas the first type will simply be interpreted as a lack of effect of the particular background variable upon the dependent, the latter two may be interpreted to exert some influence, and will be referred to as tendencies and effects, respectively. Thus, an effect is a factor that yields a significance level of less than 0.05 in the statistical test, while a tendency only approaches a significance level of 0.05, and any factor with a significance level above 0.1 will be discarded.

Since categorical data are measured primarily on a nominal scale, although sometimes on an ordinal scale, there are a number of measures that are not available and which therefore prevents the use of parametric statistical methods. The most widely used statistical method for the comparison of distributions of categorical data across contexts or conditions, is the Chi-square test. This test allows us to determine the probability that an observed difference between two distributions could have occurred by chance. The lower this probability the more likely it is that the difference between the two samples, i.e. the parameter on which they are different, holds predictive value.

The major problem with the Chi-square test, is that it does not allow the researcher to examine several background factors simultaneously. That is, the distribution of a dependent variable may be different to a statistically significant degree for values of two separate independent variables, but we have no way of knowing whether the two independent variables explain partly overlapping sections of the data. For example, we might have counted the different realizations of [w] before [ð̞] and also have registered the gender and age of the speakers who produced the tokens. Subsequent Chi-square testing may reveal that there is a statistically significant difference regarding deletion of [w] when we compare men and women, *and* that there is a statistically significant difference between older and younger speakers. But we have no way of telling whether part of the difference found for gender is related to the difference found for age, nor vice versa. The Chi-square test does not allow us to discover interactions nor to directly discover the effect of one variable when controlling for another. We could of course devise an independent variable "age+gender" with the values 'younger-men', 'younger-women', 'older-men', 'older-women' and test the distributions related to this complex variable. We could also test the difference between men and women in the two age groups separately and the difference between the two age groups for the two genders separately. This could tell us whether there was overlap between the distributions in the first general analysis. However, such serial analysis becomes increasingly cumbersome and complex for each time we add an independent variable – a factor whose influence we are interested in understanding. In multiple logistic regression it is possible to incorporate all factors at once, and to check for the possibility that two or more factors account for overlapping parts of the distribution of the dependent variable.

3.3 Basic procedure – from raw output to final model

This section presents the general outline of the model construction procedure that will be followed in the analyses of phonetic variation. This does not pretend to be an exhaustive introduction to multiple logistic regression and mixed-effects modelling. Rather it is a presentation of the principles employed in the dissertation, and hence gives the scaffolding for how the observations in the corpora are evaluated and interpreted. Here I give the general model architecture and criteria used in model criticism during construction, pruning and selection of the final model.

The construction process here mirrors the approach prevalent in much of sociolinguistics, also the studies of Danish that are the pre-cursors to the present investigation. As will be discussed, I am not altogether in favor of this procedure, since it runs the risk of capitalizing on spurious effects and thereby oversimplifying the data through the use of some times coarse grained *a priori* categorizations. I could have chosen to abandon this approach entirely and to have proceeded only in the order I find the most sensible. However, since this dissertation is partly a study of the ways in which the use of relatively new statistical methods may inform our present understanding of phonetic variation in Copenhagen Danish, I find it worthwhile to show the differences between the traditional approaches and the possibilities that arise from the use of regression modelling, in particular the use of mixed-effects techniques. This is best done by following the conventional sequence of steps in the statistical analyses of Danish sociophonetics.

The discussion of the virtues and drawbacks of the different approaches is postponed to the discussion of results, although I will give general arguments for advocating an alternative approach, which in the original sociophonetic study of the Copenhagen data from the 1980s is referred to as "going the other way" (Holmberg (1991)), i.e. by grouping speakers according to their phonetic behavior and subsequently examining their social characteristics. It should be noted that such an analysis was in fact carried out in the original Copenhagen study (Gregersen & Pedersen (1991)). However, the analysis yielded great diversity in the social make up of the groups and was not pursued beyond the variable of /a/ raising in that study. Instead, our knowledge of the connection between social characteristics of speakers and phonetic variation and change in Copenhagen Danish rests mainly on patterns at the macro-social level under the assumption that the groups defined by these macro-social classifications are sufficiently homogenic. As I hope to show, mixed-effects regression modelling allows for greater confidence in interpreting the results with respect to such background factors without assuming such homogeneity.

3.4 TRANSFORMATIONS OF THE RAW DATA

Although the classification of variants of the phonetic variables aims at a dichotomous classification in order to model the application of clearly defined processes, the classification has never been truly binary. The specifics are given in the section 2.3 "The phonetic data" above and for each variable analyzed, but relevant for the introduction here are the criteria involved in turning the multilevel dependent variable into a binary variable, which is necessary for performing multiple logistic regression. All tokens that have been classified as "faulty" i.e. with an asterisk are discarded immediately. For classifications of dichotomous variables, like the ones investigated in the LANCHART Cph Corpus, a decision must be made with respect to tokens where the coders disagree and tokens where coders agree that the variant is indeterminate between the two limiting values. For example for [w], a specific token may have been encoded by one coder as a case of deletion and by another coder as a token of realized [w]. Such tokens are discarded from the dataset. For the DanPASS Dialogues corpus a different situation obtains. Here, entire conversations have been phonetically transcribed. Hence, the tokens of a variable that have been extracted into a dataset are not inherently binary. Instead they may take on several values, values which reflect different types of processes. Here then, the individual tokens must be interpreted and their classification with respect to particular processes determined. Note that unlike for LANCHART there are no "faulty" classifications in the DanPASS Corpus, i.e. there are no sequences which haven't been transcribed. Nor is it possible to see whether a particular token has been the subject of diverging transcriptions. What is necessary is to determine which tokens should count as application of a process and which should not. Naturally, this depends on the process being modeled, and hence the description in each case will be given in the analysis of each variable.

As with any dichotomization of inherently continuous phenomena, information is lost in every case. Even more information is lost in the analyses of the DanPASS data, since I am forced to dichotomize a continuum of variants. Naturally, the coding procedure of the variables in the LANCHART Cph Corpus has the same effect, i.e. here the problem of loss of information lies in the coding of the variation, not in the treatment of the raw data. The difference mainly then is the granularity of the data available, and the subsequent treatment of the information. By choosing logistic regression (or any other categorical technique) I must eliminate some detail a priori. While I find such simplifications of the data warranted and will argue for the division in each case, it is a somewhat unfortunate, but necessary step.

Once the procedures for dichotomizing the dependent variable have been carried out, the data together with information on the background variables for each and every token is subjected to the steps involved in logistic regression modelling.

3.5 LOGISTIC REGRESSION – ANALYSIS OF VARIATION IN A CATEGORICAL VARIABLE

In logistic regression modelling the goal is to model the effect of one or more independent or background variables on a single, binary dependent variable (the introduction given here is based on Baayen (2008) except where otherwise indicated). The basic method is reminiscent of linear regression models, but differs from this method by modelling effects on variables that are binomially distributed rather than normally distributed. Actually, the method uses log transformed odds rather than probabilities, because: 1) probabilities are bounded, whereas odds range from 0 to infinity, and 2) logarithmic transformation removes the difference between values at the center of the odds distribution and at the extremes. Logistic regression, in effect, is linear regression in log odds space (p. 437, Jaeger (2008)). Where linear regression attempts to directly predict the value of the continuous dependent variable given a particular independent variable, logistic regression attempts to predict the log odds from which it is possible to derive the probability that the dichotomous dependent variable will take on the value that is the outcome of the process one is trying to model, given the value of a particular background variable. Logistic regression, then, like Chi-square, provides a measure of whether a background factor has a statistically significant effect on the probability of the desired value of the dependent variable occurring, and also, unlike Chi-square, provides information about the direction and magnitude of the effect.

Here, I give an example, first of the result of a Chi-square test of the preliminary data on [w] deletion. To investigate whether there is evidence of a change in apparent time, I conduct a Chi-square test of the distribution of the variable (wəð), which has been coded for deletion. After removing uncertain tokens, I get the following distributions of the two variants 'realized' and 'deleted' with respect to the independent variable 'age group':

	Age group	
[w]	Older	Younger
Realized	61	19
Deleted	189	105

The Chi-square test shows that this difference is significant, $p = 0.0438$. Transformation of the raw counts to proportions shows that the difference is in the expected direction: the group of older speakers delete [w] a little over 75 % of the time, the group of younger speakers delete [w] almost 85 % of the time. According to the Chi-square test, then, the difference between generations, although slight, is statistically significant, and in the expected direction: hypothesis confirmed.

3.5.1 Simple logistic regression

To illustrate the similarity between Chi-square tests and logistic regression, I have constructed a simple logistic regression model, i.e. a model with only one factor[2]. This simple model will allow me to introduce some central measures to be used in the evaluation of models, and to show how a more comprehensive understanding of the effect of the factor 'age group' is attained immediately, i.e. without transforming the counts by hand. Because there is only one factor in the model, there is no need to check for collinearity or interactions. The first measures to notice are given in the table below.

p	C	D_{xy}
0.0395	0.56	0.12

While the model as a whole is significant, because $p < 0.05$, the measures of goodness of fit, C, and Dxy, are very low. C is an index of concordance between the predicted probability and the observed response. According to Baayen (2008, p. 204) a value for C of 0.5 indicates that predictions are random, and a value of 1 indicates perfect prediction, i.e. complete agreement between predicted and observed responses. Values above 0.8 indicate that the model "may have some real predictive capacity" (Baayen (2008), p. 204). Dxy is a related measure giving the rank correlation between predicted probabilities and observed responses. Dxy varies between 0 and 1, and the higher the value, the greater the predictive capacity of the model. Clearly, in order to provide a good model of the process of [w] deletion, we need more background factors. Before constructing a more complex model, however, let us look at the coefficients of this simple model. The coefficients tell us about the direction and size of the effect that the factor has on the dependent variable.

[2] All of the statistical modelling has been conducted using the statistical software package R version 2.7.2

Factor	Coef	S.E.	Wald Z	P
Intercept	1.1309	0.1473	7.68	0.0000
agegroup=young	0.5786	0.2896	2.00	0.0457

We are mainly interested in the values in the columns marked Coef (for coefficient) and P for the factor we are exploring. The intercept will never interest us, since we are mainly interested in how the factors influence the likelihood of the modeled process applying, and I will not delve into the Standard Error and Wald Z-scores. The intercept represents the log odds for application of the modeled process, when all of the factors in a model are at the reference level. In the current example it represents the log odds for deletion of [w] by speakers in the age group 'older', averaged over all individuals. The log odds that a speaker from the age group 'younger' will delete [w] is obtained by adding the coefficient for this factor level, 0.5786, to the intercept, 1.1309, i.e. 1.7095, which may be converted into a probability for deletion of [w] with the formula: p(outcome) = $e^{intercept+coefficient} / 1 + e^{intercept+coefficient}$, where the superscript 'intercept+coefficient' are variables standing for the actual values of the intercept and the coefficient for the factor level in the model, and the transformation consists of raising (the mathematical constant) e to the power of these two values combined, and thus transforming log odds to odds. The division with 1 + the odds, converts odds to probabilities. In the current example, the probability that younger speakers delete [w] is: $e^{1.7095} / 1 + e^{1.7095} = 0.85$, corresponding to the proportion of tokens of (wəð) in which [w] is deleted by the younger group of speakers as a whole (the probability has been rounded off to the nearest second decimal). Notice that the p-value as is expected is nearly identical to the p-value in the chi-square test. This is because the method used for evaluating significance in the Chi-square test is actually derived from the method used for deriving the significance of factors in logistic regression (Johnson (2008) p. 162) – the two tests are just two ways of measuring the same thing. Since the factor is significant, it is interesting to examine what was set as the reference level of the factor and how the other level of our two level factor is related to it. The reference level is the level that the model assumes is the default (and in a simple model, the coefficient for the reference level of the factor equals the intercept). Here, the reference level was 'old' so the summary of the model provides us with a coefficient for the level young. The value of the coefficient tells us how the factor level influences the odds of the process of [w] deletion applying. So when the speaker belongs to the age group 'younger' in our sample and if the value is positive, it means that the slope rises as the value of the factor goes from the reference level to the level given in the summary. Here, that the log odds of a younger speaker deleting a [w] are higher compared to the odds that an older speaker would do so. The value in log odds given in the column "Coef" is not intuitive and therefore hard to interpret, beyond the fact that positive values indicate a rise in probability and negative values indicate a fall when compared to the reference level. To better understand the magnitude of the effect, we can convert the log odds back into the probability. For the current model we get a probability that a younger speaker will delete [w] of 0.84 - since probabilities range from 0 to 1, it is easy to see that in this case the probability directly reflects the frequency of deletion for the group of speakers indicated by the level of the factor.

3.5.2 Multiple logistic regression

So far I have shown that there isn't much difference between Chi-square testing and a simple logistic regression model. The value of logistic regression lies in the ability to model several factors simultaneously. If we wanted to control for linguistic factors in our Chi-square test of the difference between the two age groups, we would have to perform our test on a subset of the data, e.g. only unstressed tokens of the (wəð) variable. With multiple logistic regression we can simply specify that the model should evaluate and take into account the effects of any and all factors that we have information about. The factors can be either categorical or continuous. Let us look at what happens when we add the factor 'stress' to the model, that is when we tell the model that we would like to know whether the fact that a token is stressed or not has an influence on the probability that [w] will be deleted in our sample.

This two-factor model has the following characteristics:

p	C	Dxy
0.0028	0.616	0.232

Unsurprisingly, adding more information to the model improves it. All 4 measures of goodness-of-fit are better than in the simple model, although they are still not great. Let us look at the effect of stress and see whether the difference between age groups persists.

Factor	Coef	S.E.	Wald Z	P
Intercept	1.1309	0.1473	7.68	0.0000
stress=+	-0.7016	0.2573	-2.73	0.0064
agegroup=young	0.6270	0.2929	2.14	0.0323

The p-values tell us that both factors are significant, and the coefficients show that in tokens specified as + for the factor 'stress', i.e. stressed tokens, [w] is less likely to be deleted. We thus have evidence that even when we control for stress, there is an age difference in the tendency for [w] deletion, which may be interpreted as evidence for an on-going change in apparent time. One of the goals in this dissertation is to examine the effect of word frequency upon deletion processes in Copenhagen Danish. The details about this factor are described in the section "Usage based factors", but suffice it to say here that the measure for frequency is based on spontaneous speech and has been log transformed and included for each token in the preliminary dataset. In order to illustrate the incorporation of a continuous factor, I will show the model characteristics of a yet more complex model, by adding the factor 'logfrequency' to the model containing 'stress' and 'agegroup'.

3.5.3 Collinearity

Before showing the goodness-of-fits measures and the coefficients of the factors, I want to introduce a test to examine whether there is a risk of collinearity in the model. Collinearity occurs when two or more factors are strongly correlated and hence account for some of the same parts of the variation that is being modeled. When this is the case, it becomes impossible to study which of the factors is the decisive one. For example, lack of stress and word frequency may be strongly correlated, such that the likelihood for a word being unstressed always increases with word (log) frequency, and if they are very strongly correlated, we cannot tell whether an association between high probability for deletion of [w] in high frequency words should be ascribed to and interpreted in terms of word frequency or degree of stress. Some degree of correlation may often obtain between factors, but since the regression procedure can handle some degree of correlation, it is important to check the strength of the correlation, i.e. to check for collinearity. This can be done by calculating the variance inflation factor of a model. The values for each factor of the more complex model are given below.

stress=+	logfrequency	agegroup=young
1.006585	1.008161	1.014343

The variance inflation factor for factors in logistic regression models should ideally not be higher than 2 – if so, there is a risk of collinearity which may lead to spurious effects, i.e. false significances and estimated coefficients. Statisticians disagree on the threshold for the variance inflation factor, and the boundary of 2 is a conservative one, but an upper limit of 10 is generally agreed upon, since for models where a factor exhibits a variance inflation factor of 10 "collinearity cannot be ruled out from the model completely" (Jaeger & Kuperman (2009),p.40).

None of the three factors have high variance inflation factors, so we can proceed to the model diagnostics. This will help answer the question whether adding information about word (log)frequency improves our modelling of the probability that [w] will be deleted.

p	**C**	**Dxy**
0	0.855	0.711

Both measures of goodness of fit are improved, indicating that word frequency is an important factor in modelling [w] deletion. In fact, the value for the measure C is above 0.8, so the model can be assumed to have real predictive capacity, although it cannot provide perfect prediction, i.e. there are one or more unknown factors which accounts for part of the variation. The question now is whether age of the speaker is still relevant in the modelling of [w] deletion. If 'agegroup', still emerges as a significant factor, we can say that apparent time change is relatively robust. Below are significance levels and coefficients for all factors in the model with frequency.

Factor	Coef	S.E.	Wald Z	P
Intercept	8.5169	0.9405	9.06	0.0000
stress=+	-0.6676	0.3053	-2.19	0.0288
logfrequency	1.9035	0.2391	7.96	0.0000
agegroup=young	0.6774	0.3448	1.96	0.0495

The table shows all three factors to contribute to the model at a statistically significant level. Obviously, we would still expect the model to predict that unstressed syllables favor deletion in comparison to stressed syllables, and in order to have an indication that the differences between age groups is a sign of continued on-going change, that younger speakers are more prone to deletion than older speakers. We would also expect the tendency for [w] deletion to increase with log frequency, in accordance with e.g. Bybee (2002). The coefficients of the factors all go in the expected directions: the coefficient for the level 'stressed' for the factor 'stress' is negative, so [w] deletion is less likely to occur in stressed syllables compared to unstressed syllables. Conversely, the positive value for the level young in the factor 'agegroup' means that [w] deletion is more likely to occur in the speech of a speaker from the younger generation than in the speech of a speaker of the older generation. The positive coefficient for the factor logfrequency indicates that the likelihood of the process applying, i.e. of [w] being deleted before [ð̩] increases with the log frequency of the word. Thus all three factors influence the process as expected. However, we cannot tell just how different each level of each factor is from the reference level. Instead of converting the log odds to probabilities by hand, we can convert and plot them in graphs showing the probability for [w] deletion (ranging from 0 to 1) on the y-axis, using the functions 'plot' and 'plogis' in R. This has been done to construct the graphs below.

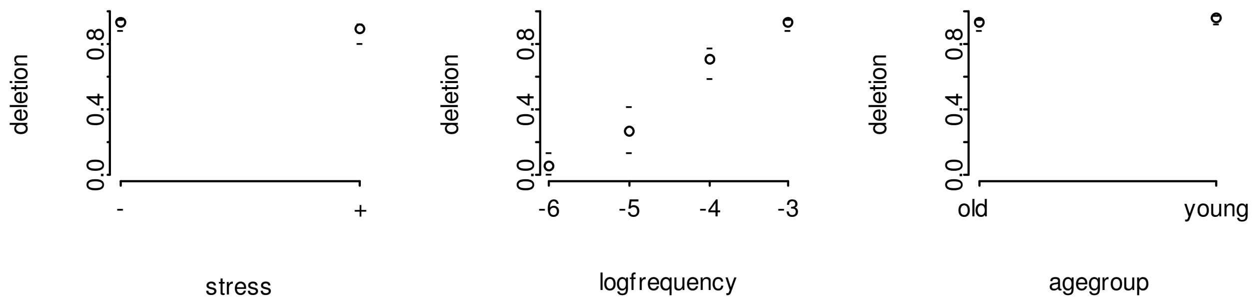


Figure B.1 – Model of [w] *deletion in apparent time (1980s sample). Circles mark the probability and confidence intervals are given by lines above and below the mark for probability*

As can be seen from the graphs, word form log frequency clearly has the biggest effect on the process of [w] deletion.

3.6 Mixed effects models

We should recall that we are so far dealing with tendencies at the group level, and somewhat large groups at that: 12 middle aged speakers, and 10 young adults. Following the tradition in sociolinguistics we have assumed sufficient homogeneity within groups, and accepted the concomitant limitations on the interpretations of our statistical analyses. Hence, the current model knows nothing about the possible influence from individual speakers, because I have not given it any information that there might be a difference between speakers. Rather than actually including 'subject' as a factor in the model, and thereby implying that individuals are sufficiently representative that they can be meaningfully used to model any other person of the same age and gender, an assumption that is clearly unreasonable, we shall see whether 'stress', 'logfrequency' and 'agegroup' are retained in a mixed-effects model where 'subject' is included as a random factor. The combination of fixed and random factors, like frequency and speaker respectively, allows us to examine whether the effects observed in the pure fixed effects multiple logistic regression models we have constructed so far, still hold when we examine the contribution from random factors. By including subject as a random factor in the model, we can take into account the contribution from each individual informant to the distribution of the probabilities of [w] being deleted. The mixed-effects model of the preliminary data on [w] deletion in the 1980s recordings has the following characteristics for the fixed effects:

Factors	**Estimate**	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-0.7706	0.7166	-1.075	0.28223
stress +	-0.2862	0.2390	-1.198	0.23113
agegroup y	-0.6601	0.4895	-1.349	0.17747
logfreq	0.5391	0.1843	2.925	0.00344 **

The names of the measures are a little different but correspond exactly to the measures seen in pure fixed-effects models previously (i.e. the estimate is the same as the coefficient and Pr(>|z|) is the same as P.)

As we might have expected, both the factors stress and age group are knocked out of the mixed-effects model – their *p*-values are both above 0.1. Log frequency, however, is retained. It would appear, then, that what looked like a small but statistically significant effect of age is an artifact of a few individual speakers.

The introduction given in this chapter is intended to show that, where warranted, mixed effects logistic regression modelling should be the preferred choice for exploring the effect of background variables on phonetic variation. The technique allows us to be (more) confident that the factors which do emerge as statistically significant are not the effects of the behavior of a few individuals. Thus, the actual analysis to be conducted in the following chapters will always control for the effects of individual informants, by incorporating the factors subject as a random effect.

Phonetic data from corpus studies is not only unbalanced with respect to the number of tokens of a variable produced by each speaker, but also with respect to the number of tokens that occur in instances of particular words. Some words occur much more frequently than others, and if these high frequency words have properties that set them apart from other words, this property may emerge as a significant factor when the model is fit to the dataset. Baayen (2008), pp. 278-281, on the basis of data from Bresnan et al (2007) provides an example of how such effects arising from overrepresented words can be controlled for by introducing words as a random effect, and Johnson (2009) also argues for inclusion of the word or word form as a random effect, since doing so effectively controls for the influence of highly frequent word forms in the dataset, since only factors which have an effect on the process studied that is greater than the effect which can be observed for individual word forms will emerge as statistically significant in the model (Johnson (2009), pp. 377-378). This also means that unbalanced datasets may be used without having to divide them into particular subsets due to the overrepresentation of a few word forms. There may be other reasons for dividing the datasets for each variable and even for excluding some word forms, but this will be discussed for each variable in turn.

3.7 General format for the presentation the analyses

As argued above I find good reason to provide a mixed effects modelling of the phonetic variables studied in the DanPASS and LANCHART corpora, and to include both subject and word form as a random effect in the models, because the datasets are often very unbalanced with respect to these two background variables (and I will have occasion to show this for some of the variables studied). The factors that are to be explored in the statistical modelling are described in chapter 5 “The Factors” and will be given for each variable in turn together with the reference level for each factor. It is important to include the reference level for the factors, since this is the value of the factor to which the other levels are compared when the model is fit to the dataset. For example, when modelling the effect of stress on [w] deletion above, the reference level was set to “stressed”, and therefore the significance of the factor level “unstressed” and its positive estimate allows us to infer that the two levels are significantly different from each other, i.e. there is statistically significant difference between the probability for deletion of [w] before [ð̞] in stressed and unstressed syllables, and deletion is more likely in unstressed syllables than in stressed syllables.

The results of fitting the regression models to each variable will be presented in the form of tables of coefficients, as was done above, but only factors that emerge as significant will be presented in the tables of coefficients, and only factors that do not enter into interactions with other factors. Interactions will be presented graphically in the form of plots of probabilities, like the ones above[3], of one factor in both (or all) of the conditions of the factor with which it interacts. This provides a better way of understanding the interactions rather than calculating their effects from the coefficients. Naturally, significances of the different levels of the interactions will be taken into account in the interpretation of the results, just as for the significances between levels of fixed effects factors that do not interact with other factors.

[3] But produced with the function plotLMER.fnc which handles the transformation of log odds into probabilities for mixed-effects models (Baayen (2008))

4. The Phonetics and Phonology of Danish Consonants

In an investigation of synchronic variation and (putative) sound change in the consonants of contemporary spoken Copenhagen Danish, it is necessary to delimit the units used in the description and analysis of speech, and to take a stance on how the different units are related to each other. I will assume that there are 3 relevant levels of representation in the description and analysis of the segmental units of speech. The 3 levels, in decreasing degree of abstraction from the speech signal, are: morphophonemes, phonemes, and phones.

These 3 levels are not an exhaustive list of the units currently used in phonetics and phonology – with the most notable exceptions being distinctive features and articulatory gestures (cf. Clements & Hume (1995) and Browman & Goldstein (1990)). However, there are historical, practical and conceptual reasons for including just these three levels of representation in the analysis of segmental variation of Copenhagen Danish.

This section provides an overview of previous analyses of the consonant system of Copenhagen Danish in the latter half of the twentieth century, with a particular focus on the levels of representation utilized in these previous studies. This leads to an overview of the inventory of consonants in contemporary Copenhagen Danish, and how the present study of their phonetic variation can be related to different interests in studying the phonetics and phonology of this particular lect.

4.1 Previous analyses of consonants in Copenhagen Danish – phonetics and phonology

Historically, Danish phonetics and phonology, both within general linguistics and in the dialectological tradition, have been practiced in a structuralist framework. The structuralism practiced in Danish phonetics and phonology has quite naturally been closely associated with European structuralism, i.e. the work of De Saussure, Trubetzkoy and Hjelmslev. The focus of European structuralism was on paradigmatic contrast, i.e. oppositions between units within a closed set with reference to other units that are attested in the same context or position. This is in contradistinction to the American structuralism practiced by the likes of Bloomfield. In this tradition, the focus was on the syntagmatic oppositions between the contrastive units of the system, i.e. their distribution in the stream of speech (Fischer-Jørgensen (1995)). A major difference between European and American, or between paradigmatic and syntagmatic, structuralism in the areas of (phonetics and) phonology, has been the role of morphology in the analysis of contrasts and establishment of inventories of contrastive units. In the strictest practices of (American) syntagmatic structuralism, no reference to morphology was allowed in the description of sound systems – only surface contrasts and distributional properties were allowed in the analysis (Fischer-Jørgensen (1995)). In the (European) paradigmatic structuralist schools morphology was incorporated in the phonological analysis early on, in particular with respect to the resolution of neutralizations of contrasts. This lead to two levels of representation, a phonological and a morphophonological one, which are maintained in some contemporary work on Danish phonology, notably in Basbøll (2005), and which is also implicit in some strands of Generative Phonology, e.g. Lexical Phonology (Kiparsky (1982)), where morphological boundary specifications are incorporated in the cycle of rules applied to abstract representations of word form (cf. the trisyllabic laxing rule in Kiparsky (1982)). Other analyses of Danish phonology have tended to focus only on the surface contrasts or what I am calling the phonological level of analysis (Brink & Lund (1975), although they (emphatically) do not refer to this level as a phonological representation), and in some cases the distinction between phonology and morphophonology is erased, such that no explicit distinction is made in the representation between phonemes and morphophonemes, although the principles of analysis are clearly delineated and explicitly refer to the morphological level (cf. mainly Grønnum (2005) and (2007), but also Pharao (2003) and unpublished).

In the analysis of Basbøll (2005) there are 21 non-syllabic segments in Modern Standard Copenhagen, the lect which provides the conservative norm for the present study (Basbøll (2005, pp. 60-65). This set of segments is reduced to 17 consonant phonemes in Basbøll (2005)'s table 2.11 on p. 73, reproduced below.

TABLE 2.11. Non-syllabic phonemes in Modern Standard Danish.

	ASP	VL UNASP PLOS	VL FRIC	NASAL	VOI CNT NON-LAT	LAT
LAB	/p-/	/b/	/f/	/m/	/v/	
ALV	/t-/	/d/	/s/	/n/	/-ð/	/l/
PAL					/j/	
VEL	/k-/	/g/		/-ŋ/		
UV/PHA					/r/	
	/h-/					

The four non-syllabic segments which are not phonemic in Basbøll (2005)'s analysis are:

[ç i̯ ɐ̯ u̯]. The latter three are the final allophones of /j r v/ respectively, and the alveopalatal sibilant (Basbøll's classification) can be analyzed as an allophone of /s/ which has been fused with /j/. In Basbøll (2005), the set of 17 consonant phonemes are reduced to a set of 16 morphophonemes as shown in his table 2.13 on p. 76, reproduced below.

TABLE 2.13. Non-syllabic morphophonemes in Modern Standard Danish. |-ŋ| is in parentheses because it is restricted to French words (as *satin* [saˈtˢeŋ]) and the weak suffixes *-ing*, *-ling*, *-ning* (e.g. *maling* [ˈmæːleŋ]), whereas other instances of /ŋ/ represent the morphophonemes |ng|.

<table>
<tr><td>|p|</td><td>|b|</td><td>|f|</td><td>|m|</td><td>|v|</td></tr>
<tr><td>|t|</td><td>|d|</td><td>|s|</td><td>|n|</td><td>|l|</td></tr>
<tr><td></td><td></td><td></td><td></td><td>|j|</td></tr>
<tr><td>|k|</td><td>|g|</td><td></td><td>(|-ŋ|)</td><td></td></tr>
<tr><td></td><td></td><td></td><td></td><td>|r|</td></tr>
<tr><td>|h-|</td><td></td><td></td><td></td><td></td></tr>
</table>

In Basbøll (2005)'s morphophonological analysis, /ð/ is a variant of |d|. All other phonemes are also morphophonemes but their distributions differ – only |h| and |ŋ| are defectively distributed. The set of consonantal morphophonemes in Basbøll (2005) corresponds exactly to the consonantal phonemes of Grønnum (2005)[4]. The latter analysis is based on morphological alternations, and is intended to reflect speaker competence, but not necessarily the representations relevant for on-line speech production (Grønnum (2005), p. 319). The lects described in Basbøll (2005) and Grønnum (2005) are largely identical, and the inventories of consonantal segments are the same in the two analyses.

Brink & Lund (1975) document the use of 36 consonantal segments in Danish throughout the twentieth century (according to their table of phonetic symbols on p. 60). However, these segments are not all attested in modern Copenhagen Danish speech, nor are they expected to be. Brink & Lund (1975) refute phonological and morphophonological representation altogether and restrict the representation of speech to a relatively fine grained system of segmental transcription (as well as length, stress and stød). The level of detail in the phonetic transcription is established on the basis of commutation. Where Basbøll (2005) and Grønnum (2005) restrict their use of commutation to individual word forms as produced in isolation for the delimitation of the inventory (with subsequent analysis into phonological and/or morphophonological levels), Brink & Lund (1975) also include commutations in (short) utterances. This extension of the domain of commutation excludes the possibility for reduction of the inventory by reference to position in the word or syllable. An example: [ɐ̯] can be interpreted as an allophone of /r/ in both Basbøll (2005) and Grønnum (2005), because it only occurs in syllable final position and because it alternates with [ʁ]. This is impossible according to the principles of Brink & Lund (1975) due to phrases like [ˈseːˀɐ̯iˈvæːˀlən] and [ˈseːˀʁivæːˀlən] which mean different things ("*do you see the whale?*" and "*behold the rival*" respectively). Hence, [ɐ̯] and [ʁ] are contrastive and serve communicative purpose according to Brink & Lund (1975).

[4] With the exception of |ŋ| which is analyzed as an allophone of |n| before velar stops in Grønnum (2005)

This approach focuses on the segments themselves as the relevant units of analysis, and excludes the possibility for establishing morphologically based correspondences between segments, one of the main aims of Grønnum (2005)'s analysis and the morphophonological analysis of Basbøll (2005). The distribution of segments can be established, exactly as it is in the phonological analysis of Basbøll (2005), where several phonemes are revealed to be in defective distribution. But segments cannot be interpreted as variants on the basis of position in the syllable, word or phrase. Only meaning is allowed to determine whether two segments are variants of the "same sound" or whether they constitute discrete units in the segmental inventory. This obviously does not mean that Brink & Lund (1975) do not allow abstraction. It is explicitly acknowledged that some variation should be abstracted away in a phonetic transcription, mainly variation that is a function of the individual speaker and coarticulatory variation. But abstraction is not pursued beyond the level of communicatively distinct utterances, i.e. utterances with distinct lexical content.

4.2 SEGMENTAL UNITS OF SPEECH

The following three levels of representation will be included in the present study of phonetic variation: morphophonemes, phonemes and distinct phones. Each level is defined below.

morphophonemes are the set of contrastive units that can be established with reference to morphological boundaries. I.e. if two phonetically distinct segments can be shown to occupy the same position in the representation of a stem and a derivative or inflection of this stem, then these two segments are viewed as variants of the same morphophoneme.

phonemes are the set of contrastive units that can be established on the basis of surface contrasts only. This means that some phonemes may be defectively distributed relative to the morphophonemes they are associated with, either with respect to their position in the syllable or word, or relative to other segments.

distinct phones are the minimal set of language-specific phonetic categories and they constitute the set of sounds that speakers of contemporary Copenhagen Danish may classify as separate phonetic categories. The inventory is expanded relative to both morphophonemic and phonemic inventories, since phones which are in free variation and even phones which may be ascribed to the same phoneme on the basis of distributions are kept separate at the level of distinctive phones. For example, 'skrevet' *written*, which is morphophonemically |ˈskrɛːvəd| and phonemically /ˈsgrɛːvəð/, will be represented as [ˈsgræːwəð] at the level of distinctive phones. The pronunciation [ˈsgræːvəð], which is possible in very distinct speech, and the fact that [w] only occurs in coda position in the phonological syllable are arguments which allow the phonologist to classify [w] as an allophone of /v/. But this does not entail that the language user does not categorize them as distinct (albeit related) sounds. The claim is that speakers are aware of the positional allophones of phonemes, and that these may constitute separate cognitive categories despite the lack of a lexical contrast. Such a view seems plausible given the fact that speech does (so much) more than convey the lexical and grammatical content of a message. As suggested in Ladd (2006) this awareness of non-lexical distinctions of segments can be incorporated into an interface representation that mediates between phonemes and the non-discrete physical signal that is overt speech.

Using this level of representation is not, however, unproblematic. No investigation exists that can show which sounds are perceived as sufficiently different from each other to constitute separate categories for speakers of contemporary Copenhagen Danish. And the distinction may not be equally clear in all cases and may differ between individuals. Rather than relying on intuitions about what constitutes sufficiently distinct speech sounds, I will delimit the inventory to include positional allophones only. I will thus assume that language users can abstract away the coarticulatory variation present in the signal, but not the variation that can be ascribed to position in the phonological syllable. For example, I will not include two different types of [k^h], one before front vowels and one before back vowels, but I will include [w ɕ ɐ̯] in the inventory. The inventory is given below in table 4.1.

Table 4.1 – Distinctive consonant phones in contemporary Copenhagen Danish	
Plosives	p^h t^s k^h b̥ d̥ ɡ̊
Fricatives	f s ɕ h
Nasals	m n ŋ
Approximants	v ð
Liquids	l ʁ
Semi-vowels	ɐ̯ j w

The extensions relative to the phonemic level (compare table 2.11 from Basbøll (2005) above) are the alveo-palatal fricative [ɕ] and the semi-vowels [ɐ̯ w]. Note that [v] and [ð] in Danish are approximants articulatorily, but I use the IPA symbols for fricatives for convenience and in order to make the transcriptions more easily comparable to the existing literature on Danish phonetics and phonology.

The analyses of reduction will thus be concerned with the processes that apply to normative forms of lexemes in distinct pronunciation. However, I will also relate the analyses to the phonemic and morphophonemic representations in order to provide a discussion of the relevance of these levels of abstraction in a model of the speech production process. To what extent the different levels of representation may be posited to have psychological reality for language users will be discussed on the basis of the analysis of the role of word form frequency in the processes modeled, and in analyses of word specific phonetic patterns as attestable in the corpora used in the investigation.

A main goal of the dissertation is thus to describe and analyze quantitatively the reduction of selected distinctive phones in running speech, and to investigate the role of language use in these processes and the implications that these processes and their connection to the word form level have for a conceptualization of lexical representation.

The practical reasons for distinguishing between these three levels of representation stem from the nature of the annotations of the corpora used in the investigation. The DanPASS Dialogues include an abstract representation roughly corresponding to the representation that is the result of a morphologically based analysis (the exceptions are described in the section "Description of the corpora used"). The LANCHART Cph Corpus contains a phonetic annotation that is based on a highly distinct speech style (see the section "Description of the corpora used"). While some rules of deletion have been incorporated into the representation that is generated (on the basis of word frequency, according to Peter Juel Henrichsen (personal communication)), no features of surface manifestation are abstracted away with reference to morphological categories. Roughly, then, DanPASS can be described as containing a morphophonological representation, whereas LANCHART contains a phonological representation. This makes it possible to study the phonetic variation at the segmental level with reference to more abstract levels of representation, and to use these abstract representations as the basis for delimiting the variables that are to be studied synchronically and diachronically. In combination with the analyses of the influence of word form frequency, it will be possible to evaluate the relation between the levels of representation.

Conceptually, then, these three levels of abstraction are maintained in the analysis in order to be able to provide a discussion of the degree of abstraction in lexical representation. As mentioned in the introduction, several studies have shown an effect of word frequency on the spreading of a sound change. The correlation between high frequency and a tendency for reduction which is also found in synchronic studies of phonetic variation is interpreted as a causal one in recent theories of lexical representation, i.e. models of the lexicon developed under the theoretical approaches of usage based phonology and exemplar theory. In these approaches, high frequency words and word forms are more often reduced *because* they are used more frequently. There is an overwhelming amount of evidence in support of this interpretation, but there is also evidence to suggest that abstraction may still take place, and that it indeed does take place simultaneously with the storage of detailed representations in memory. Recent evidence from speech perception (Norris et al (2003)) convincingly shows that abstraction simply must occur, and certainly does occur, in the face of novel and ambiguous stimuli. These results are important for the current discussion of frequency effects in phonetic variation and change, in that, although they have nothing to do with frequency effects (these are (presumably) controlled for in the experiments), they touch upon fundamental issues of mechanisms in the construction of mental lexicons. Norris et al (2003) show that learned variability in the manifestation of particular segments affects both categorical perception and is generalized to new lexical items. Evidence from phonological experiments in Danish also show that speakers do abstract away from the speech signal and that at least some of them do so in a way that suggests that morphophonological relations are cognitively real to them in their representation of speech. Grønnum (1996) shows the abstract analysis of certain Danish vowels as a combination of a full vowel + an underlying /r/ to be entirely in accordance with speaker behavior in a speech production experiment. Pharao (unpublished) studied the representation of word final consonants and the psychological reality of an abstract, morphologically informed representation. The latter study showed that at least for some speakers, the correspondence between word final approximants and syllable initial plosives found on the basis of morphophonological alternations in the existing vocabulary was generalized to nonce words. Thus, experimental evidence from speech perception and production indicates that some form of abstraction over the speech signal must also be a part of the speaker's knowledge of his or her language.

It is the intention of the present study to investigate the role of usage in phonetic variation and change. However, given the current evidence, briefly reviewed above and in more detail in the section "Usage and phonetic variation", it seems to me unreasonable to expect that usage alone will be able to account for the observed variation. Nor do I assume that word frequency will necessarily be the most important factor for any given case of variability. However, it is also clear that it plays some part in the application of processes of reduction. The effect of frequency is not easily handled in a traditional rule based framework. It is the intention of the present investigation to elucidate the probabilistic nature of the rules of manifestation for selected consonantal phonemes and morphophonemes in contemporary Copenhagen Danish.

5. THE FACTORS

This chapter describes the phonetic, linguistic and extra-linguistic factors that are used in the modelling of phonetic variation in the DanPASS Dialogues and (to some extent) in the LANCHART Cph Corpus.

5.1 PHONETIC FACTORS

5.1.1 Segmental environment

The nature of connected speech is such that we should expect neighboring segments to influence each other. This is because, while the segment is a useful analytical unit and quite probably also a psychologically real unit in speech processing (Cutler (1992), Fromkin (1971)), it is clear that articulatory gestures overlap each other across segmental boundaries (cf. Browman & Goldstein (1990)). This overlapping nature of articulatory gestures implies that the phonetic environment must be taken into account in the study of segmental variability. Therefore, in the present analysis of variation in consonants in contemporary Copenhagen Danish, the factor of segmental context has been coded. Preceding and following segments are detected automatically and vowel quality, manner of articulation of consonants and place of articulation for all types of segments are coded.

5 factors are generated for each of the two neighboring phonetic segments in the speech stream where a token occurs. These factors contain information about the segmental environment.

The factors are:

1. Segment

2a. Manner of articulation for consonants

2b. Vowel quality

3. Place of articulation

4a. Simplified manner of articulation

4b. Simplified vowel quality

5. Simplified place of articulation

The coding is carried out on the *phonetic* environment, and classifications have been made in accordance with the terminology and parameters set up by the IPA (Handbook (1999)), and the principles of simplified classification are given below.

The first of the factors pertaining to segmental context is obtained by extracting from the phonetic transcription of the DanPASS Dialogues the segments immediately adjacent to a token of a variable. This was done semi-automatically by the use of customized Praat scripts after tokens of words containing a variable had been segmented by hand. The first factor, Segment, is complex and encodes both manner and place of articulation of a neighboring segment, and aspects of vowel quality in the case of adjacent vowels. The information about manner and place of articulation has also been included separately. This was done automatically by extracting the information encoded in the factor Segment and represented in the factors Manner and Place respectively. This enables the study of the influence from different types of segments, e.g. fricatives or back vowels, on reduction processes, rather than just being able to study the influence from particular phonetic segments. The factors were generated for both the immediately preceding and the immediately following segment, yielding a total of 4 additional factors for each token: previous manner, previous place, following manner and following place. Information about vowel quality is encoded in the factors representing manner of articulation in order to simplify the structure of the datasets.

The levels of the factor Manner/Quality are:

For vowels:

Unrounded, Rounded, Centralized,

For consonants:

Stop, Fricative, Nasal, Approximant, Rhotic.

Note that nasalization of vowels also occurs in the DanPASS Dialogues. However, this was not encoded in the factors, but taken into account in the semi-automatic classification of the neighboring segments.

The levels of the factor Place are:

For vowels:

Front, Central, Back,

For consonants:

Labial, Labiodental, Alveolar, Alveopalatal, Lateral, Palatal, Velar, Radical, Glottal.

As far as I have been able to determine, these sets of features for Manner/Quality and Place adequately cover the range of segments found in the DanPASS Dialogues. The semi-automatic coding took into account the possibility that a stød might occur in the syllable in which a particular token occurred. Stød is transcribed as[ʔ] in the corpus, even though it is not a glottal stop, but rather a laryngealization of part of the nucleus and coda of the syllable carrying stød (cf. Grønnum & Basbøll (2007)). Tokens of a variable that occurred in a syllable with stød were coded for this in the factor 'stod' in the datasets (see below), but the information was not encoded in the factors that represent information about manner or place of articulation.

Simplified encodings of manner/quality and place of articulation were also generated automatically, by conflating two or more levels of the detailed representation described above. The factor of simplified manner of articulation has the following levels:

Vowel – all vowels, irrespective of lip rounding and centralization

Stop – identical to the detailed representation

Fricative – identical to the detailed representation

Sonorant – nasals and laterals

Approximant – approximants and rhotics

The simplified factor for Place has the following levels:

There is no simplification for place of articulation for vowels.

For consonants:

Labial – both labials and labiodentals

Coronal – alveolars, alveopalatals, and laterals

Dorsal – palatals and velars and radicals

Glottal – identical to the detailed representation

These sets of simplified representations of the manner and place of articulation of adjacent phonetic segments will allow for a study of the influence from segments in cases where the data is too unbalanced to make more detailed analysis feasible, i.e. when specific types of vowel quality, manner and place of articulation are sparsely represented in the dataset. As is apparent from the description of the models developed for the individual phonetic variables, the simplified factors were almost always used, due to the heavily unbalanced nature of the datasets (see further the section "A descriptive study"). In cases where it is necessary to reduce the information even further, i.e. by conflating more levels, this will be explicit in the text. For all of the 5 factors encoding the articulation of adjacent segments, the level pause was also included in cases where a token was either preceded or followed by a pause.

By analyzing the influence from neighboring segments on the probabilities for reduction of the selected distinct phones, it will be possible to discover some of the articulatory processes that occur in connected speech in contemporary Copenhagen Danish.

5.1.2 Stress

Like segmental environment, prosodic factors are ubiquitous in the study of phonetic variation, including reduction. I will not treat the studies of the influence on vowel production by stress, but limit myself to the influence of stress upon consonant reduction. The general pattern found in the literature is clear: consonants are more often reduced and deleted in unstressed syllables, than in fully stressed syllables (cf. Labov (1972)), and conversely, stressed syllables show fewer reductions (Greenberg et al (2002)) and longer segments overall (Turk & White (1999)) than unstressed syllables. Raymond et al. (2006) included prominence as a factor in their investigation of /t,d/ deletion in spontaneous speech, and found a similar effect: both /t/ and /d/ was more likely to be deleted in non-prominent syllables than in prominent ones, although the effect was strongest for tokens in syllable onset (Raymond et al (2006) p. 77-8).

In light of this evidence from both experimental and corpus studies it was deemed necessary to include the factor stress in the modeling of reduction and deletion of consonants in contemporary Copenhagen Danish. Perceived rather than lexical stress is included in the analyses of the DanPASS Dialogues, because lexical stress may be completely lost in certain phrasal constructions and derivations (cf. Rischel (1983)). The investigation of the factor stress will therefore provide details about the relative segmental richness of syllables that also carry a prosodic cue to prominence, i.e. stressed syllables, as compared to syllables without a prosodic cue to prominence, i.e. unstressed syllables.

5.1.3 Speech rate

Speech rate has been shown to influence segment deletion in other languages. Guy (1980) claims that word final /t,d/ deletion in American English is more prevalent at higher speech rates, as determined impressionistically. The connection between speech rate and shortening of segments has been established in work within Articulatory Phonology. Byrd & Tan (1996) show that faster speech has shorter segments and more gestural overlap than moderate and slow speech does. Fosler-Lussier & Morgan (1999) found that segment deletion increased from 9.3 % in slow conversational speech to 13.6 % in very fast conversational speech (p. 144). Raymond et al. (2006) found speech rate to be significant also for word internal /t,d/ deletion, and furthermore, that different subsets of speakers were prone to delete word internal /t,d/ at different speech rates: younger speakers deleted segments at slower mean speech rates than older speakers did (Raymond et al (2006), p. 72).

The basis of the measurement for speech rate used here is syllables pr. second, also shown by Fosler-Lussier & Morgan (1999) to be a better measure than phones pr. second, since whole syllables are less likely to delete than individual phones (Fosler-Lussier & Morgan (1999), p. 140). Furthermore, since the object of investigation here is consonant reduction and deletion, it is more likely that the link between the process studied and the factor speech rate will be somewhat independent of each other when the measure is based on syllables rather than phones, since consonants can be deleted in the speech stream without reducing the number of realized syllables. Since the measurement is based on realized rather than underlying syllables, the articulation rate will be influenced by the application of the processes studied, since both deletion and reduction will tend to shorten syllable durations. It must be kept in mind that the link between the processes of reduction and deletion and articulation rate means that it is not possible to say whether high speech rate *causes* reduction and deletion, if indeed the processes, as described for other languages above, are more likely to occur at higher speech rates than at lower speech rates. It is only possible to say that reduction and deletion is more likely to occur at higher rates of speech.

The domain of the measurement used to represent articulation rate in the DanPASS Dialogues is a local one, namely the articulation rate of the utterance in which the token occurs. It was obtained by counting the number of syllabic segments, i.e. vowels and syllabic contoids, in the prosodic phrase in which the token occurred and dividing this number with the duration of the phrase in seconds, yielding syllables pr. second. This was done automatically, using a customized praat-script, for each individual utterance in the DanPASS Dialogues, and the information subsequently imported into each dataset, by matching each token to the articulation rate of the prosodic phrase in which it occurred. The prosodic phrases were delimited by the transcribers of the recordings. The prosodic phrases were delimited on the basis of perceived intonational and rhythmic coherence, and conversely, boundaries between prosodic phrases were placed at point where there was a perceived rupture in rhythm and in the global intonation contour (Nina Grønnum, personal communication).

5.1.4 Position in the syllable and position in the word

Since the purpose of the present study is to account for the variation that can be observed in connected, running speech as compared to the distinct pronunciation of words spoken in isolation (which forms the basis for the morphophonological and phonological representation), the analysis and presentation of results will be carried out with reference to “distinct allophones” rather than morphophonemes or phonemes. However, in order to be able to relate the phonetic variables based on distinct pronunciation to their phonological and morphophonological representations, reference will be made to the position of the variable in the phonological syllable and in the word. Both the syllable and the word are included, because while word initial and word final position are unequivocally also syllable initial and syllable final position respectively, the syllabification of word medial consonants depends on the nature of the following vowel in the phonological analysis of Danish (cf. Basbøll (2005) and Grønnum (2005)). Word medial singletons are assigned to the onset of the following syllable if the following vowel is a full vowel, and to the coda of the preceding syllable if it is a non-full vowel, i.e. schwa. This difference facilitates the analysis of the distribution of consonants, and makes it possible to generalize the morphophonological representation with reference to the syllable. The phonological syllable, then, is a descriptive construct and not (necessarily) cognitively real for speakers of Danish. The distinction is particularly important in the analysis of what can on morphophonological grounds be represented as underlying plosives, since the allophony is highly sensitive to position in the syllable. In addition to analyzing the variation with reference to the phonological syllable, I will also relate the findings to position in the phonetic syllable. Word medial consonants that are phonologically in coda position may be syllabified to the onset of the following syllable according to the principles for phonetic syllabification that can be derived on the basis of a number of experiments described in Grønnum (1999), as long as this syllabification does not lead to onset clusters in the following syllable that do not also occur in word initial position. In effect, the phonetic syllabification of word medial consonants may move tokens from a stressed to an unstressed syllable. I will examine the role of phonetic syllabification and the consequences for the role of stress for the variables that can occur in the phonological coda word medially.

5.1.5 Position in the prosodic phrase

Studies by Keating and colleagues (cf. Keating et al. (2003)) have shown evidence for domain initial strengthening, i.e. the phenomenon that consonants are articulated more strongly, e.g. with longer duration of the stop closure, when they are in utterance initial position than when they are in phrase intial position and more strongly in phrase initial position than in word initial position. This difference has been found in a number of languages for laboratory speech, but it is clear that they may serve a purpose in any style: they can act as segmental cues to prosodic constituency. Cho et al (2007) show that listeners do in fact use this cue in lexical disambiguation, suggesting that the effect is sufficiently consistent in order for listeners to generalize it and use it practically in spoken word recognition.

In the light of the experimental evidence, it seems to me to be interesting to examine whether there might be evidence of an effect of position in the phrase on the articulation of consonants in contemporary Copenhagen Danish. In order to control for the effects from other prosodic domains, the investigation will be carried out for subsets of the data only. Tokens in syllable onset will be studied for possible phrase initial effects, and tokens in the coda will be investigated for possible phrase final effects. Note that Tøndering (2008) has shown evidence for final lengthening in the spontaneously spoken monologues of DanPASS (Tøndering (2008), p. 82 ff.). It is possible that an effect of phrase final lengthening will inhibit the tendency for reduction in the same way that speech rate may influence reduction processes.

While domain initial strengthening and final lengthening will not be studied directly, the modeling of reduction processes will take the possibility of these effects into account, and it will be possible to study to what extent they are relevant for the understanding of reduction processes in connected speech.
The factor is coded for each token by getting the number of words in the phrase in which it occurs, and getting the number of the word in which the token occurs. If the total number of words = word number, the token is phrase final, and if the word number = 1, the token is phrase initial. If there is only one word in a phrase, the token is coded as a single word phrase. All other tokens are coded as non-marginal.

5.1.6 Stød

The stød is a syllable prosody peculiar to certain varieties of Danish, and it occurs in contemporary Copenhagen Danish. Phonetically, it is a kind of creaky voice, and can only occur in heavy syllables, i.e. syllables containing a long vowel or a short vowel followed by a sonorant (Grønnum & Basbøll (2007)). In addition the syllable must have more than null stress in order to contain stød. Stød is distinctive on the surface, but may to a large extent be predicted from morphological structure and can be interpreted as functioning as a signal for morphological structure (for a detailed account of the distribution of stød see Basbøll (2005)). Stød is included as a factor in the present investigation of consonant variation in order to investigate whether the requirement that syllables with stød must be heavy influences speaker behavior such that they avoid segmental reduction in syllables with stød. Furthermore, syllables with stød are perceived as more prominent than syllables without stød, all things being equal (Tøndering (unpublished)). As a correlate of prominence then, it is possible that speakers will tend to preserve segments in syllables with stød even in spontaneous speech. If so, it is to be expected that stressed syllables with stød are even less likely to contain reduced variants of segments than stødless, stressed syllables.

5.2 EXTRA-LINGUISTIC FACTORS

5.2.1 Age

Age is used in two ways in studies of phonetic variation. In apparent time studies, age is used as a predictor of change. However, it may also be studied as a factor in a strictly synchronic interpretation of comparisons between samples of different generations without the implication of a change in progress. This is rarely done in the variationist literature, but see Rose (2006) and Rathje (2008) for examples of analyses of age related speech variation.

In the present study, age will be included as a synchronic factor in two ways: a continuous factor based on the decade of birth of the informant, and membership of the two predefined age categories older and younger. This is possible in both the LANCHART Cph Corpus and the DanPASS Dialogues, and speakers were assigned to each age group as shown in the section "Description of the corpora used". The continuous factor will be able to show a general age effect at a more fine grained level than the dichotomous classification imposed by dividing speakers into two generations. The dichotomous split will provide the opportunity to study the relative homogeneity of age co-horts, to the extent that it is possible to analyze the behavior of individuals in the groups. The hypothesis is always that younger speakers will delete and reduce more often than older speakers.

5.2.2 Gender

A general finding from dialectological and quantitative sociolinguistic research is the increased tendency for female speakers to use standard variants in their speech when compared to age-matched males (Milroy & Gordon (2003), p. 101). Or rather, this was the assumption until Holmes (1997) which reviews the literature on gender and sound change, and finds a somewhat more complex pattern. A major difficulty with the hitherto generally accepted interpretation, i.e. that women orient themselves more towards the prestige norm than men do, is that the prestige norm is not always easily established. One example is the clear difference in attitudes towards Cajun English as displayed by speakers of different ages (Dubois & Horvath (1999)). Older speakers were seen to orient themselves towards an external norm for English (presumably General American), whereas a recent revival of Cajun culture had led younger speakers to orient themselves more towards Cajun English – in other words, an on-going dialect leveling process appeared to have been reversed by the younger speakers. A particular problem emphasized by Milroy & Gordon (2003) is the assumption that social class is always more important than gender in a change in progress. A handful of analyses reveal the importance of gender over class in the structuring of phonetic variation, notably Horvath's (1985) reorganization of Labov's data on [ð]-stopping in New York City. Whereas Labov (1990) gives priority to class in his analysis of the data, Horvath (1985) shows that gender is in fact more important in structuring the variation, at least in the statistical sense, since gender is a better predictor of application of the process (realizing standard [ð] as [d] word-initially). Horvath's (1985) re-analysis is important because it turns the traditional approach in quantitative sociolinguistics on its head: rather than comparing the behavior of groups established on external grounds, Horvath (1985) groups speakers according to their linguistic behavior to discover which traits are shared by speakers who exhibit similar linguistic behavior. A similar approach will be followed in the study of phonetic change included in the present investigation by examining the behavior of individual speakers. But a traditional approach will also be followed by including the binary factor gender in the statistical modelling. The hypothesis is that women will exhibit fewer reductions, since reduction is assumed to be a non-standard feature of speech.

6. The descriptive study of the role of linguistic and extra-linguistic factors in consonant reduction

The purpose of the modeling of phonetic variation in this dissertation is two-fold: The primary goal is to examine the extent to which language use influences processes of reduction and deletion, which is done by examining the influence of word-form frequency on reduction and deletion of consonants. However, in order to adequately assess the influence from word-form frequency it is necessary to include a number of other linguistic and extra-linguistic factors which, on the basis of studies of phonetic variation in other languages, can be seen to influence the tendencies for reduction of consonants. These factors must therefore also be included in the statistical models of reduction and deletion. Since many of these factors have not previously been studied with respect to consonant reduction in contemporary Copenhagen Danish, the modeling will provide an opportunity to investigate to what extent factors like articulation rate, segmental context, and stress affect propensity for reduction of consonants in contemporary Copenhagen Danish. Rather than simply controlling for their influence, I will study in greater detail the role of linguistic and extra-linguistic factors in the patterns of consonant variation. The factors are described in detail in chapter 5 "The Factors" and the specific linguistic and extra-linguistic factors that are of major interest to particular phonetic variables will be given for each of them in turn.

The linguistic factors are studied in order to gain an understanding of how the process of speaking itself influences variation in the speech signal. For phonetic variation, these factors naturally concern parts of the articulatory process. Extra-linguistic factors are included in order to study to what extent speaker behavior is homogeneous with respect to the constraints imposed by linguistic factors. In the present investigation of variation in the production of selected consonants in spontaneous Copenhagen Danish a set of linguistic factors will be included in the study of synchronic variation. These factors are: segmental environment, stress, stød (where relevant), speech rate, position in the word and position in the prosodic phrase, and, for one variable, the morphological category. The extra-linguistic factors are age and gender, which are included to investigate the commonly held assumptions that younger speakers produce more reductions and deletions than older speakers, and that men produce fewer standard forms than women. This sub-study, then, includes extra-linguistic factors in order to contribute to an understanding of their role in shaping a standard norm. Factors of usage, like word frequency and occurrence in the conversation, are treated in the chapter on phonetic variation and language use.

The synchronic study is carried out using the DanPASS Dialogues, where the processes of monophthongization and of reduction and deletion of plosives have been analyzed. The analyses are carried out at the level of surface manifestations in distinct speech, essentially relating the variability in spontaneous speech to the standard pronunciation of words in isolation, and it is in this sense that the processes may be viewed as reductions and deletions. To what extent these processes should be viewed as on-line operations on a full underlying form in the process of articulation will be discussed in the chapter on phonetic variation and language use. Tokens of the variables were identified using the abstract (morpho)phonological representation in the DanPASS Dialogues. The following (sequences of) morphophonemes were used to identify the tokens: |aj|,|ɔv|, |p|, |k|and word final |t|.

Since the realization of the variables in distinct pronunciation forms the basis for studying the reduction processes, the morphophonemes |p| and |k| must be split into two phonetic variables each on the basis of their expected realization in distinct pronunciation. For clarity of exposition, the analyses of all of the morphophonologically delimited variables are therefore presented with reference to their expected manifestation in distinct pronunciation. The resulting seven phonetic variables are: (ɑj), (ɒw), (p), (b), (k), (g) and (d). I follow the tradition from quantitative sociophonetics and use regular brackets to refer to the phonetic variables, since they do not necessarily correspond to phonemes. Each variable is defined and treated in turn in the following.
After the presentation of the results of the analyses of variation, the distribution of variants is related to the underlying representations allowing for a discussion of the implications for the morphophonological and phonological analysis of contemporary Copenhagen Danish.

6.1 MONOPHTHONGIZATION OF (ɑj)

The (ɑj) variable concerns the monophthongization of the diphthong [ɑj]. The [ɑj] diphthong is one of only two falling diphthongs ending in [j] with a *short* vowel nucleus (the other being [ʌj] [5]) in distinct pronunciation. Diphthongs in Copenhagen Danish are analyzed as biphonemic sequences, in the case of [ɑj], a short back vowel followed by a glide, since both the nucleus and the glide are commutable with other vowels and consonants respectively. This can be seen for the [ɑj] diphthong from the minimal pairs 'mig, møg' [mɑj mʌj] *me, dung* showing commutation of the nucleus, and 'haj, ham' [hɑjˀ

[5] One word, 'huje'[ˈhujə] *to cheer*, contains the only attested token of the [uj] diphthong, which is not considered further here.

hɑmˀ] *shark, hide (n.)* showing commutation of the glide. The non-syllabic component of the diphthong is always assigned to coda position in the phonological syllable, but need not be in absolute syllable final position. However, even in complex codas as in 'sejl, fejl, spejl', [sɑjˀl fɑjˀl sb̥ɑjˀl] *sail, error, mirror* there is no difference in the allophony of the glide in distinct pronunciation: it is always [j]. The diphthong may be realized with or without stød, e.g. as in 'maj, sej, haj' [mɑjˀ sɑjˀ hɑjˀ] *May, tough, shark* and 'mig, sig, hej' [mɑj sɑj hɑj] *me, oneself, hi.* There are no words with obligatory [ɑːj] (i.e. a long syllabic component in the nucleus of the diphthong) in distinct pronunciation (Grønnum (2005), p. 295), but note that there is variation in the duration of the stressed vowel in disyllabic words of the structure [(C)VCə(C)], when the intervocalic consonant is a semi-vowel or approximant (Grønnum (2005), p. 331). This means that words like 'veje, seje, pleje' [ˈvɑjə, ˈsɑjə, ˈplɑjə] *roads, tough (pl.), nurse,* may be realized as [ˈvɑːjə ˈsɑːjə ˈpʰ̥lɑːjə] or, more likely, [ˈvɑːj̩ ˈsɑːj̩ ˈpʰlɑːj̩] i.e. with deletion of schwa, a common process in running speech (cf. Basbøll (2005)). The conditions for this lengthening have not been studied, and will not be the focus of the present study.

The [ɑj] diphthong has previously been studied with respect to the variation in the quality of the vowel – the pronunciation has varied between a front and a back low (unrounded) vowel, and may still do so even in the speech of adolescent Copenhageners (Maegaard (2007), pp. 187-193). However, the propensity for monophthongization of this particular diphthong has not been studied. Brink & Lund (1975) show that diphthongs ending in [j] become monophthongized before [ð̩] irrespective of the quality of the preceding vowel, e.g. 'liget, meget, tøjet', [ˈliːˀjəð, ˈmɑjəð, ˈtˢʌjˀəð] *the corpse, very, the clothes* are realized as [ˈliːˀð̩, mɑːð̩, tˢʌðˀð̩] and monophthongization has become categorical in this context for speakers born in the 1930s or later (Brink & Lund (1975), p. 351). Monophthongization of falling diphthongs ending in a [j] in Copenhagen Danish is known to be prevalent for diphthongs with a long vowel as the syllabic component: it is obligatory after high front vowels, but also occurs after low front vowels ([j] does not occur after long back vowels)(Grønnum (2005), p. 295). For example, 'sige,syge' [ˈsiːjə, ˈsyːjə] *to say, sick (pl.)* are always realized as [ˈsiːi, ˈsyːy] whereas 'læge, bage' [ˈlɛːjə, ˈbæːjə] *physician, to bake* may be pronounced with the glide as in [ˈlɛːjə, ˈbæːjə] or without it, as in [ˈlɛːɛ, ˈbæːæ].

This study analyzes the monophthongization of [ɑj] in all contexts. In the corpus, there is a total of 2721 tokens of words which would be expected to contain [ɑj] in distinct pronunciation in contemporary Copenhagen Danish. However, some of these words show variability with respect to the quality of the vowel, and for some speakers the quality of the vowel is never [ɑ] before [j]. As mentioned above, the vowel may also be realized as front [a], and this realization was in fact the predominant variant for speakers born between 1870 and 1940 (Brink & Lund (1975), pp. 90-91), but since then the variant [ɑ] has become dominant. Given the age of some of the speakers in the group of older informants in the DanPASS Dialogues, it is to be expected that some of them will have categorical [a] in the diphthong. For these speakers it is unreasonable to expect a back [ɑ] in the words where the diphthong occurs. Only one speaker, number 13, has categorical [aj], whereas others vary between [a] and [ɑ] in the diphthong, and others still have only [ɑ] except in the pronoun 'jeg' [jɑj] *I* and the interjection 'nej' [nɑjˀ] *no.* In order to keep the preceding context constant, all tokens which contained a vowel different from [ɑ] have been excluded from the dataset. Restricting the variable (ɑj) according to these criteria reduces the total of tokens to 1333 in the DanPASS dialogues. These have been extracted and coded with respect to linguistic and extra-linguistic factors. Monophthongization is coded as a binary variable, where tokens without a glide or a high front vowel, either [i] or [ɪ], have been classified as cases of monophthongization. All other tokens have been classified as "standard", i.e. non-monophthongized. In 803 or 60 % of tokens, the glide is retained, and in 530 or 40 % it is deleted, showing that monophthongization of (ɑj) is a common process in running speech, if not prevalent[6].

First of all, I present the proportion of monophthongized (ɑj) in the individual word forms containing the variable. There are 64 word forms containing (ɑj), 46 of which occur less than 10 times in the corpus. Categorical monophthongization only occurs in 6 of these 46 infrequent items, and none of the 18 word forms that occur more than 10 times exhibit categorical monophthongization. The explanation for the categorical monophthongization observed for the 6 word forms appears to lie in the phonetic context: Four of them, 'drejede, drejet, eget, flodlejet' [ˈd̥ʁɑjəðə, ˈd̥ʁɑjəð, ˈɑjəð, ˈfloðlɑjəð], *turned (past),*

[6] In the dataset including all 2721 tokens , 67 % of tokens were monophthongized, showing that the process is prevalent in words which in contemporary Copenhagen Danish may be expected to have [ɑj]. The difference also suggests that monophthongization of this diphthong may be followed by a change in vowel quality, but since it is not possible to decide whether such a shift requires monophthongization to occur, or whether some speakers simply show variation in both the realization of the vowel and the glide independently, the analysis is restricted to tokens where the vowel is realized as [ɑ].

turned (past ptcpl.), own, the riverbed all contain the sequence [jəð] that triggers deletion of the [j] when the schwa is assimilated to the following [ð] irrespective of the quality of the preceding vowel. Schwa assimilation has applied in every token of 'drejet' (twice), 'eget' (once), and 'flodlejet' (three times), and once in 'drejede'. All tokens of these word forms, including the two of 'drejede' in which schwa has not assimilated to the following [ð], are retained in the dataset in order to test the strength of the influence from the following segment, in this case syllabic [ð̩]. The two remaining word forms with categorical (ɑj)-monophthongization, 'leger, stavefejl' [ˈlɑjɐ, ˈsd̥æːvəfɑjˀl], *play (vb.), spelling error,* occur only once. For these latter word forms it is of course impossible to say whether the monophthongization is a property of the word form specifically, or (which I consider much more likely) a result of phonetic factors such as intervocalic position of the glide, as in 'leger', or reduced stress on the syllable with (ɑj), as in 'stavefejl'. These are therefore also retained in the dataset for statistical modeling.

All word forms were thus retained in the dataset for the statistical modelling, and are given in Appendix A. Table 6.1 gives the word forms which occurred 10 times or more in the corpus, together with an English gloss, number of occurrences, and rate of monophthongization (i.e. percentage of monophthongal (ɑj) for each word form).

Table 6.1 – word forms containing (ɑj) (n > or = 10)

newword	English	sum	% monophthongal
nej	no	313	15
jeg	I	244	94
drejer	turn (pres.)	147	49
dreje	turn (inf.)	71	6
meget	very	68	96
mig	me	66	47
indianerlejren	the indian camp	52	25
dig	you	51	76
vej	road	41	2
byggelegepladsen	the playground	28	4
flodleje	riverbed	25	0
vejen	the road	25	8
indhegning	fencing	24	4
lejr	camp	22	14
byggelegeplads	playground	12	8
indhegningen	the fencing	12	0
tegner	draw (pres.)	11	0
tegnet	draw (past ptcpl.)	11	0

The more frequent word forms, i.e. those which occur 10 or more times in the corpus, show a wide variety of word classes represented among them, although the majority of the highly frequent word forms are function words, as expected. These most frequently occurring word forms vary greatly in their propensity for monophthongization of (ɑj). Some, like 'jeg, dig, meget' [jɑj, d̥ɑj, ˈmɑjəð] *I, you, very*, show near-categorical monophthongization of (ɑj) with reduction rates of 94 %, 76% and 96 % respectively. No doubt these word forms in particular contribute to the high instance of monophthongization overall in running speech. While the extremely high rates of monophthongization for the very highly frequent word forms 'jeg, dig, meget' will necessarily skew the dataset, and suggests that the citation forms of these words may be monophthongal (with respect to (ɑj)) rather than diphthongal, exclusion from the dataset is not warranted for any of them, since they do exhibit some variation. However, the unbalanced nature of the dataset with respect to word form warrants inclusion of the word form as a random effect in the statistical modeling. This will control for the influence from individual word forms in the evaluation of the significance of the factors included in the model and ensure that statistically significant generalizations are not based on properties that are associated with a few, overrepresented word forms in the dataset. Without the inclusion of word form as a random effect, factors may emerge as statistically significant in the modeling of (ɑj) monophthongization, even when they are tied to properties of a specific, overly represented, word form. Since the purpose of the investigations in this chapter is to assess the influence from phonetic, linguistic and extra-linguistic factors in general on the reduction of the phonetic variables studied, it is desirable to control for effects that may be attributable to individual word forms, so as not to make inferences that are biased by a small set of lexical items.

As mentioned in the section "Statistical methodology", by allowing random intercepts for the individual word forms in the modelling process, the equation used to predict monophthongization effectively controls for the influence of highly frequent word forms in the dataset, since only factors which have an effect on monophthongization of (ɑj) that is greater than the effect which can be observed for individual word forms will emerge as statistically significant in the model (Johnson (2008)). That way, the information from all word forms can be kept in the dataset, without the risk of making generalizations that rest solely on the properties of a few highly frequent word forms.

The strong prevalence of monophthongized tokens of (ɑj) in particular word forms indicates that the diphthongal form may be the exception in these cases, and that the monophthongal form may be the basic lexical representation. In other words, while reduction may be the relevant process in word forms like 'indianerlejr, drejer, vej' [enḓiˈæːˀnɐlɑjˀɐ̯, ˈḓʁɑjɐ, vɑjˀ] *Indian camp, turn (vb.), road*, it is possible that word forms with a high proportion of monophthongal realizations, like 'jeg, dig, meget', are expanded in certain contexts rather than reduced in most contexts. The interpretation of the pattern depends on the approach to lexical representation, and further discussion is deferred to the chapter "Language use and the phonology of spontaneous speech".

6.2 A MODEL OF THE MONOPHTHONGIZATION OF (aj)

Table 6.2 – Factors in the full model of (ɑj) monophthongization	
Fixed effect	**Reference level**
'token stress'	stressed (yes)
'following segment'	Pause
'local articulation rate'	n.a. (continuous)
'word boundary'	Yes
'stød'	stød (yes)
'position in phrase'	Non-marginal
'age'	n.a. (continuous)
'gender'	Female

Table 6.2 on the right gives the fixed effect factors that were included in the first model of monophthongization of (ɑj) with the reference level of each given in the second column, where relevant. The random effect factors included were subject and word form, in order to control for effects of particular speakers and words.

With the reference levels given in table 6.2, the mixed-effect multiple logistic regression procedure will model the probability of monophthongization of (ɑj) in unstressed syllables without stød by male speakers while including the effects of local articulation rate, position in the prosodic phrase and age of the subjects. The model will help determine the influence from position in the word and the presence of a following segment on the process of monophthongization. Since the preceding segment [ɑ] is held constant for all tokens, only information about the following segment is included in this model. This initial modeling of monophthongization includes only a rough indication of the following segmental context, since it is interesting first of all to determine whether an effect of the following segmental context can be found, i.e. whether monophthongization is more or less likely in pre-pausal position compared to when the diphthong is followed by a segment. Details about the segmental environment will be examined further in subsequent models, with more elaborate information about the following segment, if an effect is found. There are no reasons for serious concerns about collinearity (all variance inflation factors < 2.7).

6.2.1 Characteristics of the model

Table 6.3 gives the coefficients for the factors that emerged as statistically significant when fitting the model to the dataset for (ɑj). All possible two way interactions were tested, and a significant interaction of stress and position in the prosodic phrase emerged.

Table 6.3 - Coefficients of the significant factors in the model of (ɑj) monophthongization

Factor	Estimate	Std.Error	zvalue	Pr(>lzl)
(intercept)	-9.57265	1.02898	-9.303	<2e-16***
stødless	1.98765	0.32449	6.125	9.04e-10***
word internal	1.88842	0.42010	4.495	6.95e-06***
local articulation rate	0.25473	0.07263	3.507	0.000453***
unstressed	2.77217	0.46531	5.958	2.56e-09***
phraseinitial	0.86633	0.66732	1.298	0.194210
phrasefinal	-1.72506	0.50116	-3.442	0.000577***
singleword phrase	1.03082	0.71780	1.436	0.150981
unstressed: phraseinitial	-1.71425	0.67548	-2.538	0.011154*
unstressed: phrasefinal	2.41355	0.68640	3.516	0.000438***
unstressed: singleword phrase	-1.19195	1.77200	-0.673	0.501164

Recall that positive estimates indicate an increased probability for monophthongization of (ɑj) for the level given in table 6.3 relative to the reference level specified in table 6.2 above. Interpreting the effects of the simple factors first, it is evident that the probability for monophthongization increases with local articulation rate, i.e. the faster the speech, the more likely that 'mig' [mɑj] *me* will be realized [mɑ]. Monophthongization is more likely word internally than word finally, i.e. 'vejen' [ˈvɑjˀən] *the road* is more likely to be realized as [ˈvɑːˀn̩] than 'vej' [vɑjˀ] *road* is to be realized as [ˈvɑːˀ]. Stødless syllables are more likely to be affected by monophthongization than syllables with stød, so [mɑ] for 'mig' is more likely than [vɑːˀ] for 'vej'. The partial effects of stress and position in the prosodic phrase are more complex, due to the interaction between them. The interaction shows that the effect of position in the prosodic phrase is different for stressed and unstressed syllables. Inspection of the dataset revealed that only 4 unstressed tokens occur as singleword prosodic phrases, all of which consist of compounds where the syllable containing (ɑj) has secondary stress, namely the words 'byggelegepladsen' [ˈb̥yg̊əˌlɑjəpʰlasən] *the playground* and 'indianerlejren' [end̥iˈæːˀnɐˌlɑjʁɐn] *the Indian camp*. None of these tokens contained monophthongized (ɑj). It was therefore decided to reexamine the effect of position in the phrase by looking only at the subset of data which occur in prosodic phrases containing more than one word.

The interaction between stress and position in the prosodic phrase still emerges as significant. To show the interaction, I have plotted the partial effect of position in the prosodic phrase for each of the two conditions stressed and unstressed, displayed in the left and right graph respectively.

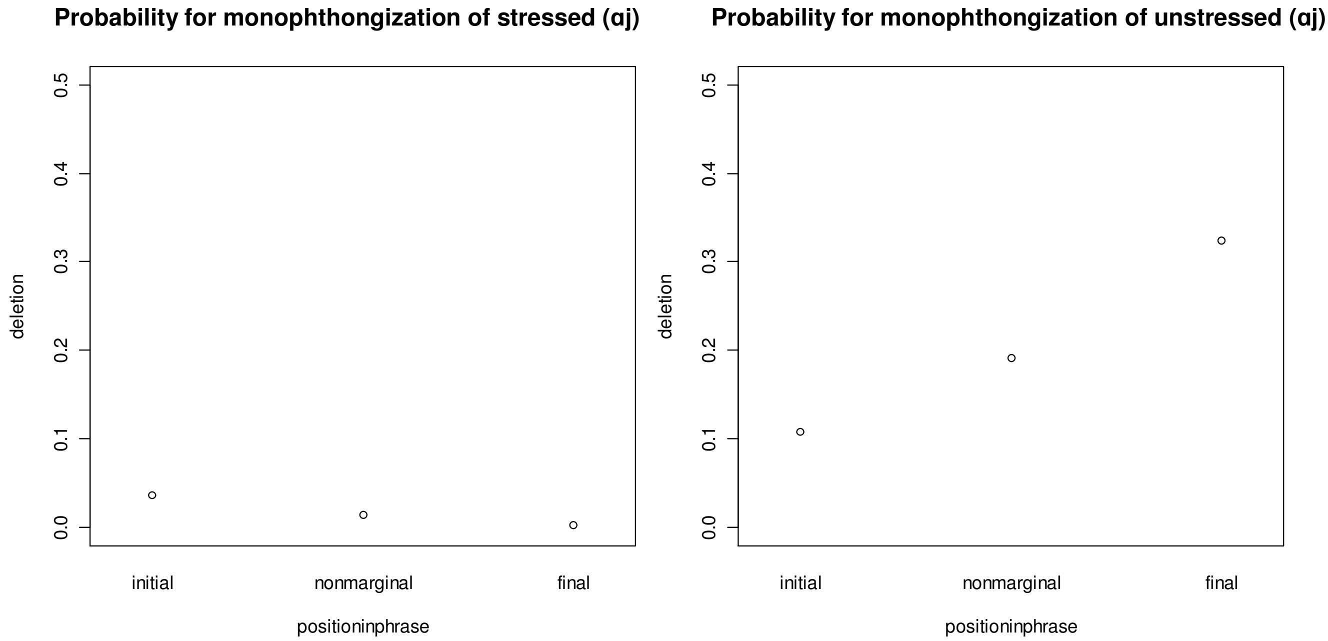


Figure 6.1 – The partial effect of position in the prosodic phrase in the two conditions stressed (on the left) and unstressed (on the right)

The probability for monophthongization is not significantly different for the three different phrasal positions when (ɑj) is stressed: The slight increase in the probability for monophthongization in phrase initial position is not statistically significant. Note that the results do not mean that (ɑj) is never monophthongized in stressed syllables – in fact (ɑj) is monophthongized in 28 % of stressed tokens – but that reduction is equally likely for stressed syllables irrespective of position in the prosodic phrase. There are significant differences for tokens in unstressed syllables: unstressed (ɑj) is most likely to be monophthongized in phrase final position, less likely to be monophthongized in phrase internal position, and least likely to be monophthongized in phrase initial position. The interaction shows that not only is there a difference between the probabilities for monophthongization with respect to stress, but the magnitude of the difference depends on the position in the phrase. Note that the probabilities presented here depend on the inclusion of the word as a random factor. This means that the model takes into account the different rates of monophthongization of the individual words that occur in each of the three positions in the prosodic phrase and adjusts the effect of position in the phrase accordingly. To explore the consequences of incorporating the word as a random factor, a model which did not take the variability between words into account was fit to the subset of tokens occurring in multiword prosodic phrases. I will focus on the consequences for unstressed tokens in the exploration of this simpler model. The interaction of stress and position in the prosodic phrase also emerges as significant in the simpler model, but in this model the probability for monophthongization of unstressed tokens of (ɑj) is *greatest* in phrase initial position (there is no statistically significant difference between non-marginal and phrase final tokens). The discrepancy between the two models can be accounted for by examining which words occur as unstressed in phrase initial position: in the present dataset there are 117 unstressed tokens of (ɑj) in phrase initial position, 61 of which occur in the word 'jeg' [jɑj] *I*, i.e. the word constitutes 52 % of the tokens in phrase initial position. The word 'jeg' has a much higher rate of monophthongization than the average rate for unstressed words. (ɑj) is monophthongized in 51 % of all words in which the variable is unstressed, whereas 97 % of unstressed tokens of 'jeg' are monophthongized, i.e. realized as [jɑ]. Furthermore, 'jeg' is monophthongized in all but 1 of the 61 tokens occurring in phrase initial position, and constitutes 73 % of the monophthongized tokens in this phrasal position. In other words, the tendency for monophthongization of unstressed tokens in phrase initial position is heavily influenced by the increased tendency for monophthongization of the word 'jeg'. This imbalance in the distribution makes it difficult to determine whether the high proportion of monophthongized tokens in phrase initial position has to do with the position in the phrase or with the words that occur in this position. Incorporating the word as a random effect factor adjusts the likelihood of monophthongization of (ɑj) by taking into account the difference between the average likelihood of (ɑj) monophthongization and the likelihood of monophthongization in the particular words that occur in phrase initial position. Thus, the effect found in the more complex model provides a better estimate of the effect of the factor position in the phrase, than the effect found in the simpler model, which disregards the differences between individual words.

In summary, then, monophthongization of [ɑj] can be seen to be influenced by the prominence of the syllable in which it occurs, with monophthongization in general being rare in syllables with stress and even rarer in syllables with stress and stød. There is also a general effect of articulation rate, as well as a tendency for (ɑj) monophthongization to be most likely in final position in the prosodic phrase and in single word phrases, when the token is unstressed.

6.3 MONOPHTHONGIZATION OF (ɒw)

The (ɒw) variable concerns the monophthongization of the diphthong [ɒw]. There are no words with a phonologically long vowel in the [ɒw] diphthong (Grønnum (2005), p. 296), with the exception of words with the sequence [-ɒwə], where the quantity of the stressed vowel is currently in variation (Grønnum (2005) p. 331).

The present analysis concerns the loss of [w] after original or standard short [ɒ]. Previous studies have found that the diphthong is only rarely monophthongized before [ð̩], unlike other diphthongs ending in [w] (cf. Brink & Lund (1975) and the section on [w] deletion). However, since the diphthong occurs in some highly frequent words where it is often monophthongized, the variable will be studied further here. All 2365 tokens of morphophonological |ɔv|, i.e. phonological [ɒw], the presumed lexical representation of (ɒw), were extracted from the DanPASS Dialogues. 510 or 22 % of the tokens are not monophthongized, whereas in 1855 or 78 % of the tokens, monophthongization applies. This distribution indicates that the monophthongization of (ɒw) is a very common process in running speech, and that the phonological representation as [ɒw] can be hypothesized to be a fairly conservative one for some word forms. In order to investigate this, the 23 individual word forms were examined for their propensity for monophthongization. The table below gives each word form (with an English gloss), the total number of occurrences in the corpus, and the percentage of monophthongized tokens. The word forms are sorted by number of tokens in the corpus.

Table 6.4 – word forms with [ɒw] (n > or = 10)

Word form	English	n	% reduced
og	and	1717	92
over	over	187	15
okay	okay	150	53
også	also	113	100
jernbaneoverskæring	railway crossing	45	45
skov	forest	37	3
oven	above	34	35
jernbaneoverskæringen	the railway crossing	33	33
hov	oops	12	08

6.3.1 Reducing the dataset

As was the case for (ɑj), the majority of these high frequency words in which monophthongization is prevalent are, of course, function words, like the conjunction ‘og’ [ɒw] *and,* the connective adverb ‘også’ [ˈɒwsʌ] *also,* the preposition ‘over’ [ˈɒwˀɐ] *over,* and the interjection ‘okay’ [ˈɒwˈkɛj] *okay.* Inspection of the propensity for monophthongization shows that (ɒw) in ‘også’ [ˈɒwsʌ] *also* is categorically monophthongized. Given the relatively frequent occurrence of this word in the corpus, it seems unlikely that the lack of variation is due to an insufficient number of tokens, but rather that this word form can no longer reasonably be said to contain the relevant phonological structure, and hence not the variable (ɒw). Therefore I have decided to exclude all 113 tokens of ‘også’ from the dataset of (ɒw) in the statistical modelling. There are a few other words, which can also be called into question for their suitability for inclusion in the dataset, namely: ‘og’ *and*, ‘okay’ *okay* and ‘hov’ *oops.* They vary not only with respect to the realization of the semi-vowel in (ɒw), but also with respect to the quality of the vowel. These words are discussed and the criteria for further reducing the dataset are given below.

The words ‘okay’ and ‘hov’ may both be realized with [ɔ] rather that [ɒ] in the stressed syllable, and if so do not constitute tokens of (ɒw) either, since there is nothing to suggest that such realizations are connected to the process of monophthongization. That is, the forms with [ɔ] are not the result of a putative concomitant process of raising[7]. I have examined the quality of the stressed vowel in these word forms for each token in the corpus, and their distribution is given in the table below.

Table 6.5 – number of tokens with (ɒw) for the word forms ‘hov’ and ‘okay’

Word form	Total number of occurrences	Number of occurrences with (ɒw)
Hov	12	10
Okay	150	41

As table 6.5 shows, ‘okay’ pronounced with [ɒ] in the first syllable is in fact not the dominant realization, where as [ɒ] in ‘hov’ is prevalent. For the conjunction ‘og’ *and*, realization of the word as simply [ʌ] is very common: it happens in 1456 of the 1717 tokens of the word, and an additional 15 tokens are realized as [ʌw], i.e. a diphthong with a different vowel quality than that expected in distinct speech. It is possible that this multitude of monophthongal tokens are indicative of what may historically be a new, alternate

[7] Historically, the [ɔ] in ‘hov’ *oops* may be related to a possible lowering of original [o] in ‘jo’ (Brink & Lund (1975), p. 184).

form of the lemma 'og' *and*, but which synchronically could be viewed as an independent lexeme. The occurrence, although rare, of the form [ʌw] shows that monophthongization is not a prerequisite for the vowel shift, and in this respect the forms of 'og' with [ʌ], regardless of whether it is followed by a [w] or not, are like the forms of 'hov' and 'okay' with [ɔ] in that they do not, at the surface level, constitute tokens of (ɒw) and should on these grounds be excluded from the modelling of the process of (ɒw) monophthongization.

Excluding tokens with a vowel quality in the nucleus of the [ɒw] diphthong different from [ɒ] reduces the dataset by (2 'hov' + 109 'okay') 111 tokens. If we further accept that a realization of the conjunction 'og' with a vowel quality different from [ɒ] is a realization of an alternate lexeme for the same lemma and therefore cannot constitute a token of (ɒw), the dataset is drastically reduced by (1717 tokens of 'og' - 164 tokens with [ɒ]) 1553 tokens, in which 'og' is realized with a vowel different from [ɒ]. This means that of the original 2365 tokens of distinct [ɒw] (2365 – (113 'også' + 2 'hov' + 109 'okay' + 1553 'og')) a mere 588 remain. Of these 588 tokens with genuine (ɒw), 171 tokens or 29 % are monophthongized. The proportion of monophthongized tokens by word form is given in table 6.6 for the word forms which occur 10 or more times in the corpus (the full dataset is given in Appendix A). Note that the dataset is still unbalanced, with a few word forms making up the largest proportion of the entire set, indicating that inclusion of the word form as a random effect is warranted.

Table 6.6 – Word forms with (ɒw) (n > or = 10)

Word form	English	sum	% reduced
over	over	187	15
og	and	164	41
jernbaneoverskæring	railway crossing	45	44
okay	okay	41	54
skov	forest	37	3
oven	above	34	35
jernbaneoverskæringen	the railway crossing	33	33
hov	oops	10	10

6.4 A MODEL OF THE MONOPHTHONGIZATION OF (ɒw)

To further explore the factors involved in the monophthongization of (ɒw), a mixed effects multiple logistic regression model was fit to the data with the fixed effect factors given in table 6.7 below. The second column gives the reference level for each factor, where applicable.

The mixed-effect multiple logistic regression procedure will model the probability of monophthongization of (ɒw) in unstressed syllables without stød by male speakers while taking into account the effect of local articulation rate and age of the subjects. The model will help determine the influence from position in the word and the segmental context on monophthongization of (ɒw). Inclusion of the random factors subject and word form will help control for the unbalanced nature of the dataset.

Table 6.7 – Fixed effects in the full model of (ɒw) monophthongization	
Fixed effect	**Reference level**
'tokenstress'	stressed (yes)
'segmental context'	Pause
Stød	Yes
Word boundary	Yes
'local articulation rate'	n.a. (continuous)
Position in phrase	Non-marginal
'decade'	n.a. (continuous)
'gender'	Female

The goodness of fit characteristics of the model are: C = 0.949 and D_{xy} = 0.899 indicating an excellent fit. The coefficients are shown in table 6.8 below for the factors that emerge as significant when the model is fit to the dataset of (ɒw). All possible interactions were tested, and two emerged as significant: stress interacts with both the following segment and stød. Neither age nor gender emerged as significant factors in the model, nor did position in the word. There are two simple effects in the model: monophthongization of (ɒw) increases with local articulation rate, meaning that 'okay' is more likely to be realized as [ˈɒˈkʰɛj] the faster the speech rate. Tokens in initial position in the prosodic phrase are less likely to be monophthongized than tokens in any other position in the phrase or in single word phrases, e.g. 'og' is less likely to be realized as [ɒ] in phrase initial position.

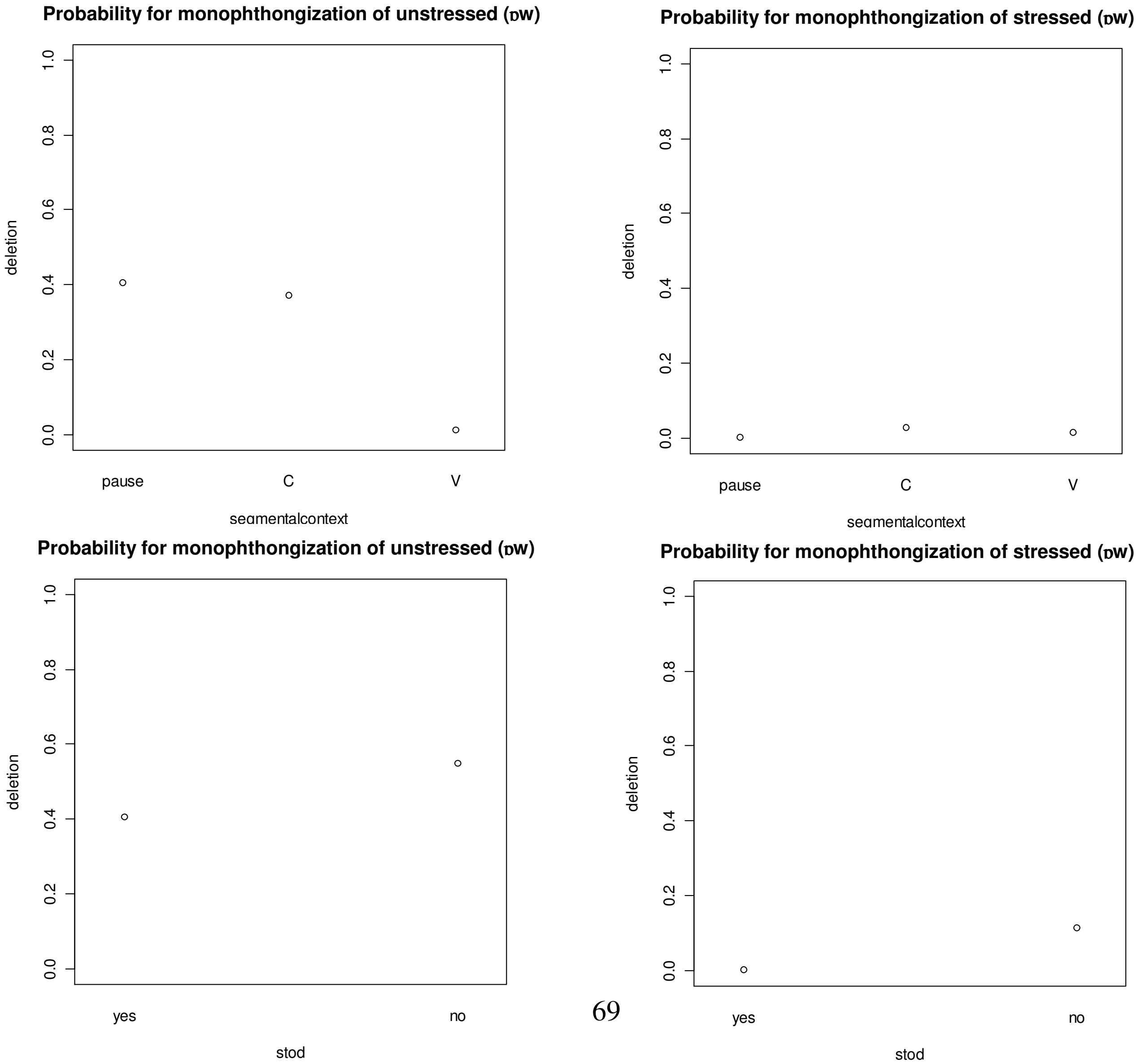
Table 6.8 – coefficients of simple factors in the model of (ɒw) monophthongization

Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	-9.066980	1.687668	7.77e-08 ***
Local articulation rate	0.646639	0.115316	2.05e-08 ***
Phrase final	0.289169	0.584072	0.62054
Phrase initial	-1.406716	0.593999	0.01787 *
Phrase singleword	0.207937	0.705366	0.76815

Table 6.9 – coefficients of the interactions in the model of (ɒw) monophthongization			
Factor	**Estimate**	**Std. Error**	**Pr(>|z|)**
unstressed:following C	-2.238586	0.974948	0.02167 *
unstressed:following V	-5.318113	1.191408	8.05e-06 ***
unstressed:no stød	-3.459604	1.171478	0.00315 **

Table 6.9 shows the significant interactions in the model of (ɒw) monophthongization. The interactions may be interpreted as a general inhibitory effect of stress, as can be seen from the illustrations of the interactions in the figure below where the log odds have been transformed into probabilities.

Figure 6.2 – Interactions in the model of (ɒw) *monophthongization. Top: the effect of segmental context by stress. Bottom: the effect of stød by stress. Left: unstressed tokens. Right: stressed tokens*



The interaction with stress has similar effects for both segmental context and stød. Note that for the unstressed tokens with stød, the stress is reduced (either because of compounding or phrase formation), not canonically unstressed, i.e. they are not lexically specified for null stress, because then they could not carry stød (see section “The Factors”). Stressed tokens are very much less likely to be monophthongized than unstressed tokens *irrespective* of the following segment and whether or not the token has stød, as can be seen from the two panels in the right column of the figure. For unstressed tokens, the situation is different. Monophthongization is more likely, but the probability is affected by the segmental context and stød.

Monophthongization is more likely in pre-pausal position and before consonants than before vowels. The effect of pre-pausal position is almost exclusively constituted by tokens of the word ‘og: 20 of the 21 tokens that are monophthongized before a pause are realization of ‘og’ as [ɒ]. The effect of a following consonant is perhaps more surprising, given that very few of the word forms that exhibit monophthongization of (ɒw) contain an underlying consonant immediately following the diphthong – this is only true of ‘okay’ [ˈɒwˈkʰɛj] and ‘sjovt’ [ɕɒwˀd̥] *fun*. However, a following consonant most often occurs after a token of ‘og’, meaning that the following consonant belongs to the following word. But consonants can also occur word internally after (ɒw) when an underlying schwa is assimilated to a following sonorant or deleted entirely with concomitant loss of the syllable, which happens in the word forms ‘oven’ [ˈɒwən] *above* and ‘sovende’ [ˈsɒwənə] *sleeping*, and frequently in ‘over‘ [ˈɒwˀɐ] *over*, when the morpheme occurs in any form of the compound ‘jernbaneoverskæring’ [ˈjæɐ̯nb̥æːnəɒwɐsg̊æːˀeŋ] *railway crossing*. This suggests that the effect of a following consonant is perhaps better interpreted as an effect of overall reduction of the word form, since at least for word internal tokens, loss of a following schwa is a prerequisite for monophthongization of (ɒw). I return to this interpretation of the role of the following consonant below in the section “Further analysis of the following segment”.

The interaction of stress and stød is more straightforward: stressed syllables with stød are never monophthongized, whereas unstressed syllables with stød can be monophthongized. This difference concerns the word ‘over’, which may retain stød even when stress is lost due to unit accentuation. That is, ‘over’ is more likely to be realized [ɒˀɐ] in unstressed position than in stressed position, where it is always realized as [ˈɒwˀɐ]. For stødless tokens, monophthongization is more likely in unstressed syllables over all.

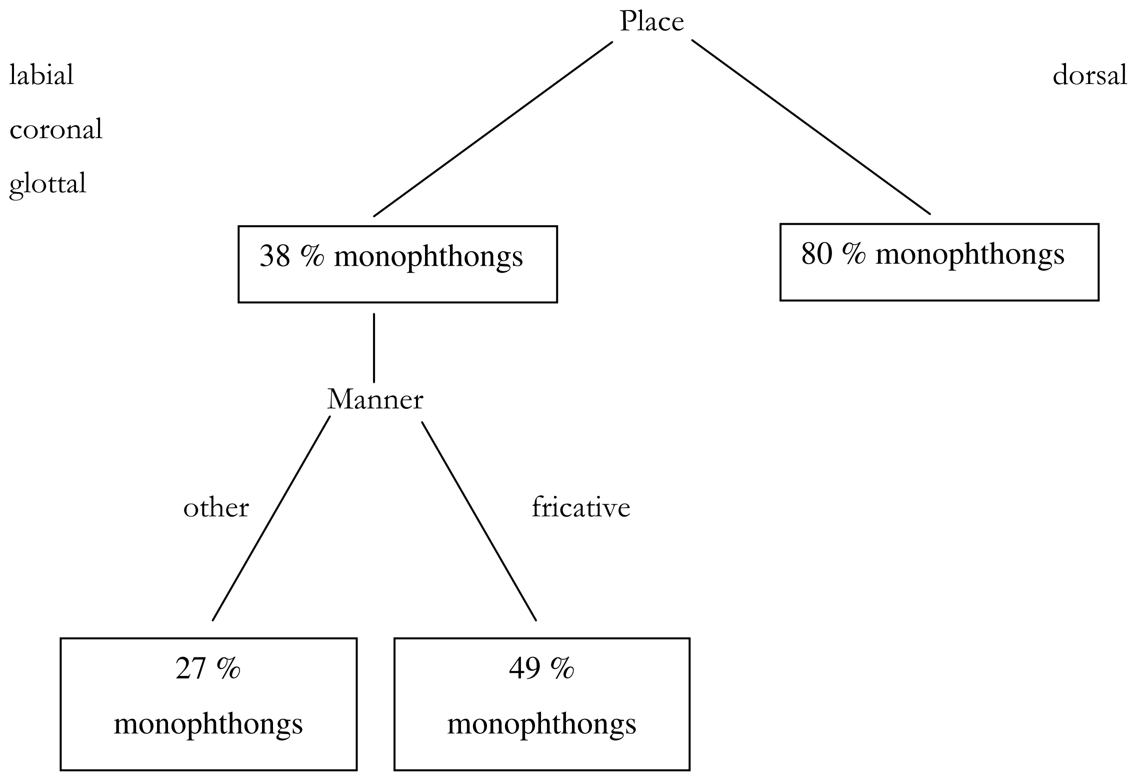
Note that the influence from the type of the following segment for unstressed tokens suggests an interpretation of the effect in terms of the *phonetic* syllabification of (ɒw). While the underlying [w] is always in coda position in the phonological syllable in contemporary Copenhagen Danish, the phonetic syllable allows for [w] in onset position, due to resyllabification of intervocalic consonants in phonetic syllabification, such that they constitute onsets in the following phonetic syllable even when the nucleus is not a full vowel (Grønnum (1999) where what I am calling phonetic syllables here, are referred to as "speaker" or "cognitive" syllables) as long as illegal consonant clusters are not formed, i.e. sequences of consonants which do not occur in word initial position. Therefore, when (ɒw) is followed by a consonant in the distinct word form, as in 'ovre' [ˈɒwʁɐ], or a pause, i.e. word finally as in 'og, dog, grov' [ɒw d̥ɒw g̊ʁɒwˀ] the [w] must be assigned to coda position in the preceding phonetic syllable, since consonant clusters with initial [w] are not allowed in word initial position. But when [w] is followed by a vowel in the distinct word form, as in 'over' [ˈɒwˀɐ] and 'sovende' [ˈsɒwənə], it is either ambisyllabic or resyllabified to onset position in the following syllable. Under these principles for phonetic syllabification, the effects of both position in the word and the following segment may be reinterpreted as the result of one simple factor: deletion of the [w] is more likely in the coda than in the onset of the phonetic syllable, since as shown for unstressed tokens in the upper left panel of figure 6.2 above, monophthongization is more likely before consonants and pauses. The influence from the type of the following segment is conditioned by the stress of the syllable containing (ɒw): regardless of the following type of segment, (ɒw) is very rarely monophthongized in stressed syllables. Therefore, a test of the role of position in the phonetic syllable should be carried out separately for stressed and unstressed tokens. A simple chi square test of the difference between the number of monophthongized and non-reduced unstressed tokens in onset and coda position reveals no statistically significant difference between the two conditions. 41 of 107 tokens, or 38 %, are monophthongized, when the [w] of (ɒw) is in the onset of the following phonetic syllable, and 37 of 97 tokens, or 38 %, are monophthongized when the [w] remains in coda position. However, for stressed tokens, there is a significant difference ($p < 0.001$): in 33 of 207 or 16 % of tokens where [w] is resyllabified to onset position, monophthongization occurs, and in 60 of 177 or 34 % of tokens where the [w] remains in the coda, monophthongization occurs. This calls for a re-examination of stress as a factor, since all of the tokens where [w] is resyllabified to onset naturally occur in the post-tonic syllable, and should therefore be grouped with the tokens which were already classified as unstressed, i.e. (107 + 97 + 207 =) 411 tokens of [w] in unstressed syllables of which (41 + 37 + 33 =) 111 or 27 % of tokens are deleted compared to the 177 tokens in stressed syllables, where 60 or 33 % of tokens are deleted. This difference is not significant ($p = 0.09$), but phonetic resyllabification results in a greater proportion of deleted [w] in stressed syllables than in unstressed syllables, contrary to the expectation that stress inhibits monophthongization. With all tokens in coda position in the phonological syllable, the distribution of tokens with respect to stress is different: 78 of 204 or 38 % of tokens in unstressed syllables are deleted, whereas 93 of 384 or 24 % of tokens are deleted in stressed syllables, and this difference is significant ($p < 0.001$). This shows that the classification with respect to position in the phonological syllable which was the basis for the classification used in the modeling procedure described above, is the best way of representing the affiliation of [w] in (ɒw) also with respect to the classification of tokens as stressed or unstressed.

6.4.1 Further analysis of the following segment

Given that a following consonant increases the probability for monophthongization of (ɒw), it is relevant to investigate whether this probability is the same for all types of consonants, which would lend further support to the interpretation of the effect as one of position of the underlying [w] in the coda of the phonetic syllable, or whether some consonants favor monophthongization more than others. To explore this, three separate models were fit to the subset of tokens followed by consonants. One model contained information about the place of articulation of the following consonant, another model contained information about the manner of articulation of the following consonant, and the third model contained information about both place and manner of articulation of the following consonant. Manner of articulation only emerged as significant in the model that also contained information about the place of articulation of the following consonant, suggesting that place is a stronger predictor than manner, and that manner is only relevant for some places of articulation. To visualize the effects of place and manner, a CART analysis was conducted incorporating only information about the place and manner of the following consonant (i.e. position in the word, local articulation rate, stød and stress were not included). This analysis produces a decision tree which can be used to interpret the effects of the two factors and their interaction. The decision tree is shown below.
The decision tree is read from top to bottom. The CART analysis attempts an optimal division of the dataset on the basis of the factors included in the analysis. In the current analysis the procedure first selects which of the two factors place or manner of articulation of the following consonant is best for accurately predicting monophthongization of (ɒw). The basis for the decision is which split provides

the fewest contradictions of the generalization both regarding monophthongization and non-reduction of the diphthong (see further Baayen (2008), pp. 148-154). The analysis then goes on to consider whether the other factor provides relevant information for further improving the prediction of monophthongization. The CART analysis in this way provides a description of the dataset with respect to the process studied, in the form of the dependent variable, and the association between this process and the factors of interest, here place[8] and manner of articulation of the following consonant.

Place and manner of articulation of a following consonant and its influence on monophthongization of (ɒw)



The top (or root) node of the tree shows that place is a better predictor than manner, which is what is to be expected on the basis of the regression models. It further shows that monophthongization is most likely to occur when the following consonant has a dorsal place of articulation: when followed by a dorsal, as in 'okay', monophthongization occurs in 80 % tokens, but only in 38 % of tokens when followed by a labial, coronal or glottal. The CART analysis next shows that manner of articulation is relevant for the non-dorsal consonants, but not for the dorsals. For the non-dorsals,

[8] Classified according to the active articulator as described in the chapter "The Factors".

monophthongization is more likely before fricatives, where it occurs 49% of the time, than before approximants, sonorants and stops. That is, of the 103 tokens that occur before non-dorsals, monophthongization is most likely when the consonant is a fricative, meaning that monophthongization in this context is more likely in ‘jernbaneoverskæring’, when the syllable [ɐ] intervening between (ɒw) and [s] is omitted, than in ‘oven’ or ‘ovre’[9].

It is unclear, however, whether this result indicating that a following fricative favors monophthongization should be interpreted as a cause of the monophthongization. Looking at the raw data, the following interpretation of the association between monophthongization and a following [s] appears more plausible: when monophthongization occurs, the [ɒ] and the [ɐ] are juxtaposed, which may result in a fusion of the two phones that is manifested as a lengthening of the [ɒ]. The majority of tokens where this process is possible, i.e. where (ɒw) in distinct pronunciation is followed by [ɐ], as in the word ‘over’ [ˈɒwˀɐ] *over* and compounds with ‘over’ such as ‘jernbaneoverskæring’ contain long [ɒː] when (ɒw) has been monophthongized: there are 271 tokens of word forms containing ‘over’, and (ɒw) is monophthongized in 60 of them (i.e. 22 % of tokens). In 35 of these tokens the phonetic realization of the vowel in (ɒw) is long [ɒː], and 26 of these tokens occur in the compound ‘jernbaneoverskæring(en)’. In fact, further inspection of the dataset reveals that lengthening of the vowel commonly co-occurs with monophthongization of tokens that are followed by schwa in the distinct pronunciation, e.g. it is also found in ‘oven’ and ‘sovende’. As the CART analysis of the manner and place of articulation of the following consonant showed, this phenomenon is more frequent in ‘jernbaneoverskæring(en)’ than in ‘oven’ and ‘ovre’. However, the lengthening of the vowel cannot always in these cases be ascribed to a fusion with the following schwa, since the schwa more often assimilates to the following sonorant, i.e. ‘oven’ is realized as [ɒːn̩] not [ɒːn] or [ɒːɒn]. The [ɒ] is never long when the [w] is retained, i.e. when monophthongization has not applied. Since information about the vowel in (ɒw) has not been coded in the dataset, this correlation was not detected in the modeling of the process. However, the inspection of the data strongly suggests that the duration of the preceding vowel may be important, so a factor was devised to encode the quantity of the vowel in (ɒw) as given in the phonetic transcriptions in the DanPASS Dialogues. Adding this factor to the model and fitting it to

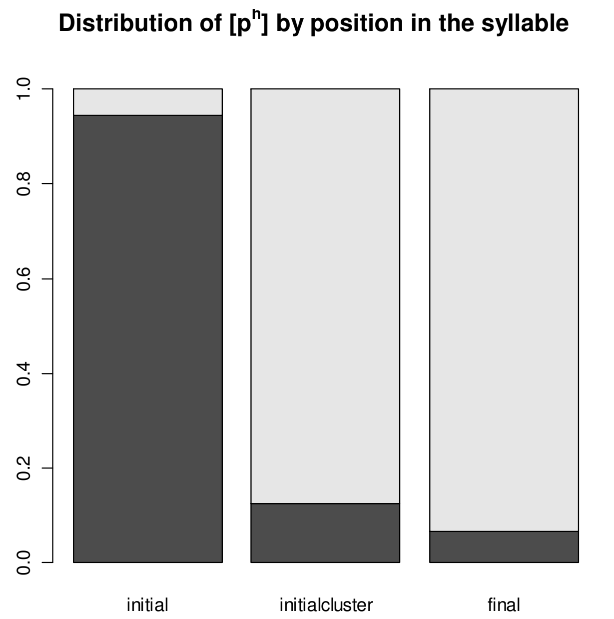
9 The CART analysis actually suggested a further split of the dataset by reintroducing the factor place for fricatives. However, the error rate was so high that this relationship has been omitted.

the dataset for (ɒw) reveals a strong significant main effect: monophthongization is more likely when the vowel is realized as long than when it is realized as short (estimate = 6.93792, $p < 0.0001$). However, the factors that emerged as significant in the initial model all remain significant. Adding the factor of duration of the vowel in (ɒw) to the CART analysis for studying the influence from the place and manner of the following consonant, however, shows that the quantity of the preceding vowel best predicts the monophthongization, i.e. the deletion of the [w] in (ɒw): monophthongization occurs in 46 out of 47 tokens where the vowel is realized as long. In the remaining 168 tokens of (ɒw) followed by a consonant, monophthongization is more likely before dorsals, than before other consonants: monophthongization occurs in 38 of the 48 pre-dorsal tokens. Manner of articulation, however, no longer emerges in the decision tree, indicating that the effect of fricatives emerging in the first CART analysis is an artifact of lengthening of the vowel in (ɒw) being more common in compounds with 'over' like 'jernbaneoverskæring'. The more detailed analysis of the dataset for (ɒw) shows that the lengthening of the vowel is an important factor in monophthongization, with the process being more likely when the vowel is lengthened in the surface manifestation than when it is not. The following segment still plays a significant role when the variable occurs in unstressed syllables. Here, monophthongization is more common before dorsal consonants and before pauses than when the variable occurs in pre-vocalic position.

In summary, there are no social factors that show an influence on monophthongization of (ɒw): men and women and young and old alike are equally likely to delete the non-syllabic segment. However, some factors inhibit monophthongization: stressed tokens are almost never monophthongized, and monophthongization is less likely in words occurring in initial position in the prosodic phrase. Increased articulation rate promotes monophthongization and the process is more common when the vowel is lengthened and before a dorsal consonant.

6.5 |p| *IN SPONTANEOUS SPEECH*

For the morphophoneme |p|, all tokens have been extracted from the DanPASS dialogues. As mentioned in the introduction to this chapter, the allophony of underlying plosives is sensitive to position in the syllable, with obligatory loss of aspiration in syllable final position in distinct speech. Preliminary inspection of the data confirms that the regularities observed in distinct speech also hold in connected speech, as expected. The DanPASS dialogues have not been segmented at the level of phonological syllables, so I coded word internal tokens for position in the phonological syllable, in order to carry out analyses at the level of the phonological syllable. The graph below shows the distribution of the allophone [pʰ] versus other variants by position in the phonological syllable. There are three possible positions: absolute initial position, occurrence in a syllable initial cluster, and final position. Position in a cluster only concerns tokens of |p| that are not in absolute syllable initial position, i.e. |p| in 'spiser' and 'sprang' are both coded as occurring in an initial cluster, whereas |p| in 'kompromis' is coded as syllable initial even though it shares the onset with a following [ʁ]. For final position, there is no distinction between immediately postvocalic tokens in absolute syllable final position and in a consonant cluster in the coda. This is because the expected realization of |p| in coda position is [b̥] regardless of whether it is in absolute syllable final position or part of a complex coda.



Light grey portions of the columns give the proportion of tokens that are *not* realized as [pʰ]. Clearly the distribution is skewed with respect to position in the phonological syllable, as expected. While tokens may be realized differently from [pʰ] in initial position, this happens much more frequently in onset clusters and syllable codas. Therefore, the tokens in the different positions must be treated as separate phonetic variables: (p) in absolute syllable onset and (b) in consonant clusters and syllable coda. The phonetic variables are given in regular brackets, following the conventions of the sociophonetic literature. (p) is defined as tokens of |p| which are expected to be realized as [pʰ] in distinct, casual speech, and (b) is defined as tokens of |p| which are expected to be realized as [b̥] in distinct, casual speech[10]. Note that not all tokens of [b̥] constitute tokens of (b) as defined here: some tokens of syllable final [b̥] are |b| underlyingly (if they are also realized as [b̥] in derivatives like 'plombe, plombere' or if they alternate with [w] in distinct, casual speech, like 'skib' *ship* (see Grønnum (2005), p. 320). The variable (p) is

[10] With the exception of tokens in onset consonant clusters: only 16 tokens are attested, 2 of which are realized as [pʰ], the remaining 14 as [b̥]. Obviously, the representation of cluster internal |p| is much too sparse to be analyzed statistically and the tokens have therefore been excluded from the analyses.

identical to the phoneme /p/, which only occurs in absolute initial position in the phonological syllable, but (b) is not identical to the phoneme /b/, since /b/ may occur in both syllable initial and syllable final position, and (b) is, here, restricted to tokens in coda position. In the DanPASS dialogues there are 3 tokens of |b| in syllable final position, and they are all realized as [b̥]. These three tokens are not part of the set of tokens of (b).

6.6 Variation of (b)

(b) exhibits the greatest amount of variation of the two bilabial plosive variables. There are a total of 539 tokens of (b), 408 of which are realized as the canonical allophone [b̥], with the remaining 131 tokens being distributed among 5 variants. The distribution of variants of (b) is given in table 6.10.

Table 6.10 – Distribution of variants of (b)						
Variant	b̥	β	pʰ	deleted	ɸ	f
Number	408	71	37	15	7	1
Percentage	76 %	13 %	7 %	3 %	.9 %	.1 %

In keeping with the canonical analysis of |p|, the standard variant of the variable (b) is [b̥] and the remaining variants can be classified as being the result of 3 different processes: strengthening in the case of [pʰ], reduction in the case of [β ɸ f], and deletion. In the examination of which factors that can be seen to influence the reduction of (b), it is of course necessary to exclude the tokens that are affected by strengthening, i.e. where (b) is realized as an aspirated bilabial plosive. Therefore, a subset of the data was created that included only the (539-37 =) 502 tokens where (b) was realized as either [b̥ β ɸ f] or deleted. There are 45 word forms in the corpus which contain (b), but only 15 of these contain reduced (b). They are given in table 6.11 with their rates of reduction (the full dataset is given in Appendix A).

Table 6.11 – word forms containing reduced (b) (n > or = 10)			
word form	**English**	**sum**	**% reduced**
op	up	131	4
klippe	cliff	74	31
oppe	up at	72	31
klippehave	rock garden	28	18
kalkstensklipper	limestone cliffs	18	28
simpelthen	simply	18	78
klippehaven	the rock garden	17	12
klippeskred	lavine	16	44
jep	yup	12	33

Only one word, ‘simpelthen’ [ˈsemˀb̥əld̥ˈhɛnˀ] *simply*, has reduced variants as the dominant realization of (b) – the four word forms that show categorical reduction of (b) all occur only once, making it impossible to tell whether (b) is always deleted in these words. Note that the majority of word forms with reduced (b) consists of inflections of and compounds with the lexeme [ˈkʰleb̥ə] viz. ‘klippe, klipper, klippehaven, klippehavens, klippeskred, kalkstensklipperne’. Grouping all of these word forms yields a total of 170 tokens, 46 of which have reduced (b), a reduction rate of 27 % for the lexeme [ˈkʰleb̥ə]. It seems unlikely, then, that reduction should be associated with this lexeme, but together with the relatively high reduction rate of (b) in ‘oppe’ it is possible that intervocalic position may emerge as significant in the statistical modelling of (b) reduction. Given the relative scarcity of deletion of (b), I will model reduction and deletion as a single process, rather than attempt to distinguish between the two.

6.7 A MODEL OF THE REDUCTION OF (b)

A mixed effects model with the fixed factors token stress, the preceding and following segment, word boundary, local articulation rate, position in the prosodic phrase, age and gender, and with the random factors subject and word form was fit to the subset. The goodness of fit characteristics for the mixed model are: C = 0.91 and Dxy = 0.83 indicating an excellent fit. The table of coefficients for significant factors is shown below. No interactions emerged as significant, i.e. not even the interaction of the preceding and following segment, in spite of the tendencies observed upon initial inspection of the dataset. The coefficients for the factors that emerged as significant are shown in table 6.12.

Table 6.12 – coefficients for the model of (b) reduction

	Estimate	Std. Error	Pr(>|z|)
(Intercept)	-6.27705	2.16662	0.00377 **
Stød: no	3.71525	1.41359	0.00858 **
Segment before Vowel	-1.93669	0.98545	0.04938 *
Segment after pause	1.86006	0.78411	0.01768 *
Segment after Vowel	1.24704	0.61713	0.04331 *

Reduction of (b) is most likely in syllables without stød, which is unsurprising, since none of the word forms that exhibit reduction of (b) contain stød in the syllable in which (b) occurs. Reduction is more likely after consonants than after vowels, e.g. reduction occurs more often in ‘simpelthen’ than in ‘oppe, klippeskred’, and it is more likely before vowels and pauses than before consonants. The effect of the following segment reflects the increased tendency for (b) to be reduced when a following schwa is realized rather than elided. E.g. when the schwa in ‘oppe’ [ˈʌb̥ə] *up at* is realized, (b) is more likely to be reduced than when the schwa is deleted and the next word begins with a consonant. This influence from the following segment may also be interpreted in the light of phonetic syllabification, as was done for monophthongization of (ɒw) above. When (b) is followed by a vowel phonetically, it may be resyllabified either as onset in the following syllable, or, when intervocalic, as ambisyllabic, constituting both the coda of the preceding syllable and the onset of the following syllable (cf. Basbøll (2005), p. 257). Assuming such phonetic resyllabification of the phonological string, it would appear that ambisyllabic tokens of (b), as in 'oppe', or onset tokens of (b), as in 'simpelthen', are more likely to be reduced, than tokens of (b) in coda position in the phonetic syllable. This is supported by a chi-square test on the dataset for (b) when resyllabification has been carried out: 27 % in onset are reduced but only 3 % in coda, and this difference is significant ($p < 0.0001$). The difference is even supported when all tokens of expected [b̥] in distinct pronunciation are added, i.e. when including tokens of /b/ in onset of phonological syllables and classifying them as (part of) onsets in the phonetic syllable (of which there are 1292). For this expanded dataset, 3 % of tokens are reduced in coda position in the phonetic syllable, but 7 % are reduced in onset position, a difference that is significant with $p = 0.037$, although this is in large part due to the overwhelming amount of tokens in onset position: there are 1612, 121 of which are reduced, whereas there are only 181 in coda, 9 of which are reduced. As for phonological syllable position-effects, dividing the dataset for all tokens of distinct [b̥] into a distribution on the basis of position in the phonological syllable, a significant difference is also found, but this time much larger than for position in the phonetic syllable and in the expected direction: 2 % in the onset of the phonological syllable are reduced, and 18 % in the coda are reduced. Note that tokens of /b/ in the onset of the phonological syllable are only very rarely reduced: 33 tokens of the total of 1292 tokens, i.e. 2 % , are reduced (most commonly to [β]), most commonly in pre-tonic syllables in words like 'begynder' [b̥eˈg̊ønˀɐ] *begin (pres.)* and word internally in compounds like 'granitbruddet' [g̊ʁɑˈnid̥b̥ʁuðˀəd̥] *the granite quarry* and 'safaribilen' [saˈfɑːʁib̥iːˀlən] *the safari car*. But the process is so rare that no factors emerge as significant when a mixed-effects model is fit to the dataset.

The consequences for stress in the dataset for (b) following resyllabification suggests an effect of stress, at least when the factor is examined on its own. In unstressed tokens there are 27 % reduced tokens in onset (10 of 37) and 20 % in coda (1 of 5), but the difference is not significant, (according to a Fisher's exact test, which was used due to the very small number of tokens in coda position). For stressed tokens, the same pattern is found, this time significant ($p < 0.0001$). 27 % of tokens in onset, or 74 of 278, are reduced, whereas only 5 % or 9 of 181 in coda are reduced. Since all tokens are followed by a (canonically) unstressed syllable, it is possible to reanalyze the effect of stress in the dataset for (b) by adding all tokens in stressed syllables that are resyllabified to the onset of the following syllable, 278, to the set of unstressed tokens, (37 + 5 =) 42, i.e. 320 tokens in unstressed position of which (74+10+1 =) 85 or 27 % are reduced. Compared to the reduction rate for the remaining stressed tokens (i.e. the stressed tokens where (b) is not resyllabified), 9 of 181 or 5 % are reduced, and the difference between stressed and unstressed tokens remains significant. Taken together, analysis in terms of position in the phonetic syllable shows reduction of (b) to be much more likely in onset position than in coda position, and more importantly, to be more likely in unstressed than in stressed syllables. Incorporating the revised factor of stress into the model for (b) reduction further supports the use of affiliation with phonetic syllables in the study of the role of stress. Whereas stress did not emerge as significant in the initial model, it does emerge after resyllabification with, as expected, reduction of (b) being less likely in stressed syllables than in unstressed ones (coefficient for stressed tokens = -3.3804, p = 0.00219). Furthermore, phonetic syllabification for the dataset of all tokens of distinct [b̥], that is in any position where [b̥] is expected in the word and syllable, also reveals the expected effect of stress: 2% are reduced in stressed syllables, 11 % are reduced in unstressed syllables.

Position in the phonetic syllable itself cannot be a relevant principle for (b) reduction, however, since the probabilities for reduction before pauses and phonetic vowels are not statistically significantly different (p = 0.20 for the difference between vowels and pauses). Pre-pausal tokens of (b) are necessarily in coda position (since there is nothing in the following pause with which the (b) can form a syllable), and therefore the relevant distinction must be the nature of the following segment, not the syllabic affiliation of the (b). Phonetic resyllabification thus enhances the effect of stress, but the evidence suggests an indirect effect of phonetic resyllabification: when a tokens can be resyllabified to form part of the following unstressed syllable, the probability for reduction increases.

There is no effect of position in the word or local articulation rate, nor is there an effect of gender or age. The overall results from the modeling correspond well to the observations that can be made from the table of reduction rates for word forms, since the word 'simpelthen' [ˈsemˀb̥əld̥ˈhɛnˀ] is the only word for which reduction is more likely than non-reduction, and here the (b) occurs after a consonant, before a vowel, and the syllable containing the (b) is stødless. Closer inspection of the variation in the realizations of the word 'simpelthen' reveals that the reduction of (b) in this word is part of a more extensive process of overall reduction of the word, namely loss of the second syllable. This level of detail cannot be captured in the modeling of the process, since deletion and reduction have been grouped together due to the relatively low instance of complete deletion of (b), cf. table 6.10 above which shows that there are only 15 instances of complete deletion of (b) in the corpus. Furthermore, the coding of the segmental context does not contain information regarding the relationship between the transcribed realization of the adjacent segments and their expected manifestation in distinct speech. This means that the modeling cannot take into account that a token of a variable occurs in a reduced context – it only takes into account the reduction of the variant itself. However, examining the 18 occurrences of 'simpelthen' in the DanPASS Dialogues reveals that (b) is omitted or realized as [m] in 14 of the tokens, i.e. nearly all of the instances of deletion attested[11]. In 2 of the 14 tokens, a syllabic [m̩] is all that surfaces as a manifestation of the second syllable, and since in these two tokens the preceding syllable ends in an [m] (carrying the stød), the syllabic [m̩] can be interpreted as a variant of (b) that has been assimilated to the preceding nasal. In 4 of the 14 occurrences, the second syllable is realized as a syllabic [l̩], but in the remaining 8 tokens, the word 'simpelthen' has been reduced to a disyllabic word in the phonetic realization without any trace in the transcriptions of the segmental sequence [əl] that would be expected in distinct pronunciation, nor of the (b). Taken together, the data for 'simpelthen' reveals that what has been operationalized as deletion of (b) covers two distinct processes: the (rare) assimilation of (b) to a preceding nasal resulting in an [m̩] in the overt manifestation, and omission not just of the (b) but of the entire second syllable of the word.

6.8 VARIATION OF (p)

As the initial inspection of the data showed, syllable initial position is where |p| is most resistant to reduction. But apparently, a small proportion of tokens are in fact reduced in syllable initial position. What is of interest here is of course which variants other than [pʰ] occur in this position and which factors influence their occurrence. In particular, it is interesting to find out whether there is also an effect of the following segment on (p) as was found for (b). Seven variants other than [pʰ] are attested in syllable initial position, their number and proportion shown in table 6.13 below.

Table 6.13 – Distribution of variants of (p)							
Variant	b̥	β	ɸ	f	ʙ[12]	p̞[13]	h
Number	57	17	11	3	2	7	1
Percentage	58 %	17 %	11 %	3 %	2 %	7 %	1 %

Only 5.5 % or 98 of 1770 tokens of the variable are realized as something other than the expected variant [pʰ]. As is apparent from the table above, the overwhelming majority of reduced tokens of (p)

[11] The remaining 15th instance of (b) deletion occurs in the word 'jep' [jɛb̥] *yup*

[12] A bilabial trill

[13] The diacritic for lowering indicates incomplete closure not accompanied by frication

are realizations as [b̥], i.e. the most common process to apply is loss of aspiration. However, reduction to a fricative is also fairly common given that the variants [β, ɸ] and [f] together constitute 31 % of the non-standard tokens. There are no cases of (p) being deleted.

Table 6.14 – word forms with reduced (p) (n > or = 10)			
Word form	**English**	**n**	**% reduced**
på	on	655	8
passerer	passes (vb)	87	2
passeret	passed (past ptcpl)	73	5
byggelegepladsen	the playground	35	3
startpunktet	the starting point	33	3
parkerede	parked	28	7
private	private (pl)	27	4
totempælen	the totem pole	20	10
præcis	exact	18	11
løveparken	the lion park	16	6
tidspunkt	point in time	15	33
byggelegeplads	playground	13	8

There are 97 different word forms containing (p) in the DanPASS Dialogues, but only 16 of them contain reduced variants of (p), some of which are very infrequent in the sample. The word forms with reduced (p) occurring 10 times or more in the corpus are shown in table 6.14 on the left together with number of occurrences and proportion of reduced variants. The fourth column which gives the proportion of reduced variants for each word form clearly shows that there are no word forms in which reduction of (p) is dominant (the two word forms in the dataset showing categorical reduction of (p), 'parkerer' [pɑˈkʰeːˀʁɐ] *park (vb.pres.)* and 'puh' [pʰuː] *phew*, each occur only once). Three word forms in the dataset show relatively high proportions of reduction in comparison to the rest of the word forms, namely 'papiret, papirets, tidspunkt' [pʰaˈpʰiːˀʁəð, pʰaˈpʰiːˀʁəðs, ˈtˢiðspʰɔŋd̥] *the paper, the paper's, point in time*, with the reduction rates 33 %, 20 % and 33 % respectively (see Appendix A for a list of the full dataset together with the reduction rates). Note that in 'papiret' and 'papirets' the reduction concerns the word initial (p) – the word medial (p) in these word forms is never reduced in the corpus, nor is it reduced in compounds with 'papir-' as in e.g. 'papirkanten'.

6.9 A MODEL OF THE REDUCTION OF (p)

A mixed effects model with token stress, position in the word, preceding and following segment, local articulation rate, position in the prosodic phrase, gender and age was fit to the subset of tokens of (p). The random factors were speaker and word form. A check for collinearity in a pure fixed effects model reveals a slight risk of collinearity for position in the word, but since the variance inflation factors were not disconcertingly high (they were between 3 and 4), I choose to proceed by fitting the mixed effect model to the tokens of (p). The significant factors are given in table 6.15 below.

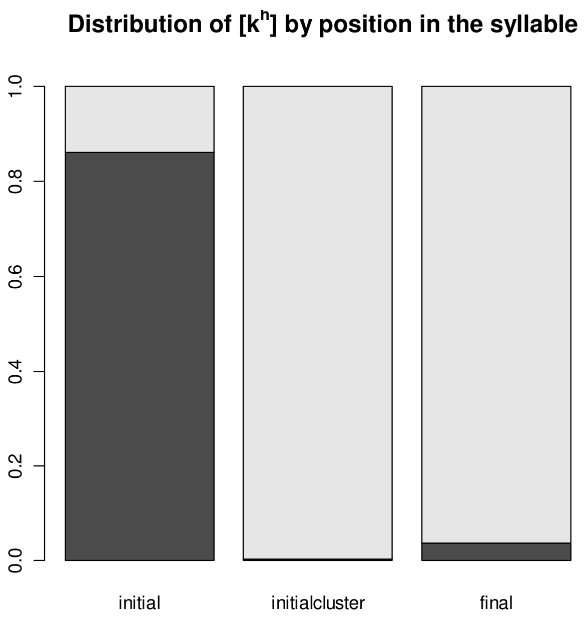
Table 6.15 - coefficients for the model of (p) reduction			
Factor	Estimate	Std. Error	Pr(>|z|)
(intercept)	-5.787261	1.386865	3.01e-05 ***
stødless	-0.876344	0.331429	0.00819 **
segment before: vowel	0.564613	0.262138	0.03125 *
local articulation rate	0.302224	0.093094	0.00117 **
word internal	-1.348434	0.504788	0.00756 **

The goodness of fit characteristics of the model are C = 0.85, Dxy = 0.72 indicating a good fit. There were no significant interactions of any factors. The table of coefficients shows that reduction of (p) is most likely in word initial position, after a vowel and in syllables with stød. In addition, the probability for reduction increases with articulation rate. The effect of stød is somewhat surprising, given that stød is a property of syllable rhymes (cf. Grønnum & Basbøll (2007)) and therefore would not be expected to influence consonants in the onset. The effect observed here is also surprising in the sense that tokens *with* stød are more likely to contain reduced (p) than tokens without it. Closer inspection of the dataset reveals that the effect mainly concerns the word 'på' [pʰɔːˀ] *on*, which may lose stød when unstressed and does so in the 95 % of occurrences in the corpus. But the loss of stød is not accompanied by an increased tendency for reduction of (p), and therefore the tendency for reduction of (p) is greater in tokens with stød than in tokens without it. It is important to recall here, that none of the contexts mentioned favor reduction of (p). The effects only show that given a reduced token of (p) it is more likely that it occurs in word initial position than in word internal position, that it is preceded by a vowel rather than a consonant, that the syllable has (retained its) stød, and that the local articulation rate is relatively high.

Neither age nor gender emerge as statistically significant, nor does stress, position in the prosodic phrase or the following segment. There are no discernible effects of different types of preceding consonants and vowels, but the difference between the two types of segment suggest that (p) is more prone to reduction in environments where the vocal tract is relatively open prior to articulation of (p).

6.10 |k| IN SPONTANEOUS SPEECH

This section presents the statistical analyses of the variation attested for the morphophoneme |k| in the DanPASS Dialogues. As with the morphophoneme |p|, there are restrictions on the allophony of |k| relative to position in the phonological syllable: in the canonical phonological analysis, only one allophone occurs in syllable initial position, namely [kʰ]. In syllable final position, [g̊] is the standard allophone of |k|, with the possibility of [kʰ] occurring in pre-pausal position. The morphophoneme |k|, then, encompasses two phonemes: the defectively distributed /k/, which only occurs syllable initially, and the evenly distributed /g/, which in syllable final position is |k|, but in syllable initial position is |g|. As with |p|, it is necessary to examine to what extent the regularities of the canonical phonological analysis, based on distinct pronunciation, generalize to running speech. If the restrictions apply, as we would expect, the analysis should be split into two separate variables: (k) covering all tokens of |k| in absolute initial position in the phonological syllable, with the expected realization [k^h] in distinct speech, and (g) for tokens of |k| with the expected realization [g̊]. In order to investigate the distribution of [k^h], the tokens have been syllabified at the phonological level, again with 3 positions being coded: initial for tokens in absolute syllable initial position, like 'kan' [k^{h}anˀ] *can*, initial clusters for tokens in a consonantal cluster, except where the |k| occurs as the first segment of the cluster, i.e. 'skarpt, skov' [sg̊ɑːb̥d̥ sg̊ɒwˀ] *sharply, forest* are classified as clusters, whereas 'krævede' [ˈk^{h}ʁæːwəðə] *demanded* is not. All tokens in syllable final position are coded as final, irrespective of whether they are part of a consonant cluster, as in 'faktisk' [ˈfɑg̊t^{s}isg̊] *actually*, or not, as in 'stik' [sd̥eg̊] *due*, since in distinct pronunciation, the allophony is not expected to be different[14].



The graph on the left shows the distribution with respect to position in the syllable of the variant [k^h] of all 5749 tokens of |k| in the DanPASS Dialogues. The light grey portions of the columns give the proportion of tokens where |k| is realized differently from [k^h]. Clearly, as with |p|, the limitations on the allophony of |k| with respect to position in the phonological syllable also generally hold for running speech, even if variants other than [k^h] do occur in syllable initial position, and conversely, [k^h] does occur in syllable final position. Therefore, the analyses will be conducted for the variables (k) and (g) separately. (k) is defined as tokens of |k| where the expected realization in distinct, casual speech is [k^h], and (g) is defined as tokens of |k| where the expected realization is [g̊]. That is, (k) covers all tokens of |k| in absolute syllable initial position, and hence also all tokens of the phoneme /k/, which does not occur elsewhere. (g) covers all tokens of |k| in syllable

[14] All phones may be aspirated pre-pausally, but here I choose to classify that as a characteristic of hyperdistinct speech.

final position as well as in consonant clusters, and hence also all syllable final tokens of the phoneme [g̊]. Since the variation is apparently greater in syllable final position, I begin by modeling reduction in the tokens of (g).

6.11 VARIATION OF (g)

There are 2418 tokens of (g) in the DanPASS Dialogues, but the final subset of tokens of (g) only contains 2344 tokens for reasons detailed below. The proportion of the different variants of (g) are given in the table below, together with total number of occurrence. The numbers are based on the dataset for (g). As the table shows, the majority of tokens are realized as [g̊], the standard variant of (g). There are 7 other variants including complete deletion.

Table 6.16 – Distribution of variants of (g)

Variant	g̊	ɣ	x	deleted	kʰ	ŋ	j	ç
Number	1233	757	117	45	89	83	13	6
Percentage	53 %	32 %	5 %	2 %	4 %	3 %	.7 %	.3 %

The majority of the reduced tokens, 37 %, constitute cases of frication, i.e. where (g) is realized as either [ɣ] or [x]. A very small proportion, 2 %, of tokens are deleted completely.

6.11.1 Reducing the dataset

Note that 74 tokens that could be classified as deletion have been excluded from the dataset. These tokens are all from the word ‘punkt’ and compounds with ‘punkt’, e.g. ‘tidspunkt, startpunkt’, and were originally included in the dataset, since the morphophonological representation given to this lexeme in the corpus is |pɔnkt|. However, in all tokens, the underlying |k| is deleted, which begs the question of whether the lexeme ought not to be represented as |pɔngt|, in which case the deletion is to be expected (see Grønnum (2005) p. 308 ff. for details). In other words, these tokens would appear to participate in a regular phonological process, rather than a process of deletion in the realization of the surface form. Inspection of the dataset reveals that of all of the 131 tokens containing the lexeme ‘punkt’, *none* contained the sequence [ŋˀg] in the realizations of the coda: all codas were realized as either [ŋˀd] or [ŋˀ] (when the final plosive had been deleted). This indicates that even in distinct form, the simple lexeme and compounds containing ‘punkt’ are [phɔŋʔd̥] rather than the expected [phɔŋʔg̊d̥]. Furthermore, it strongly suggests that the morphophonological representation of the lexeme ‘punkt’ should be |pɔngt|

rather than |pɔnkt|, although all derivations with ‘punkt’ ought to be analyzed before such a change in the morphophonological representation is carried out, and since none occur in the DanPASS Dialogues, this cannot be settled using the present dataset. Given the categorical lack of [g̊] in any of the word forms with ‘punkt’ shown here, the tokens were removed from the dataset, except for the 57 compounds that contained (g) in another morpheme in the compund, i.e. all tokens of ‘udsigtspunkt’ [ˈuðseg̊d̥spʰɔŋˀd̥] and ‘udsigtspunktet’ [ˈuðseg̊d̥spʰɔŋˀd̥əð].

The tokens where (g) is realized as [ŋ] are all realizations of the negation ‘ikke’ [ˈeg̊ə] *not*. They are included as variants of (g) here, because it is not clear that the realization as [ŋ] is necessarily a process of assimilation to a following nasal, but rather that the word form [eŋ] is a particular form of ‘ikke’. The distribution of the phonetic form [eŋ] for ‘ikke’, suggests that this form is particularly common when ‘ikke’ is used as a tag. There are 84 tokens of [eŋ] in the DanPASS Dialogues, 81 of which are in phrase final position, and 79 of which are followed by a pause. This does not in itself prove that the function of [eŋ] is a tag, since I have made no analysis of the function of the utterances in the discourse where [eŋ] occurs, but it seems reasonable to interpret the distribution along these lines, and certainly indicates that an independent analysis of the functions of ‘ikke’ and the relationship to the different phonetic forms in spontaneous speech is warranted. Note also the very small proportion of tokens, 1 % in all, evincing a process of palatalization. (g) is palatalized mainly in the word ‘rigtigt’ [ˈʁæg̊d̥id̥] *right*, but also in the words ‘seks’ [sɛg̊s] *six* and ‘ikke’. Clearly, they are all preceded by front vowels, suggesting an articulatory motivation for the palatal variants, but the scarcity of palatalized tokens precludes statistical modeling of the process on its own. Due to the relatively small proportion of tokens that show full deletion of (g), I will model the overall reduction of (g), i.e. including deleted tokens in the subset of tokens in syllable final position. There are 148 different word forms containing (g) in the DanPASS Dialogues, 86 of which contain reduced variants of (g). With such a multitude of word forms containing reduced variants of (g), the process clearly cannot be a word specific phenomenon. The reduction rates are given for each word form occurring at least 10 times in the corpus in table 6.17 below with an English gloss, and total number of occurrences (the full dataset is given in Appendix A).

Table 6.17 - word forms with (g) (n > or = 10)			
Word form	**English gloss**	**sum**	**% reduced**
ikke	not	587	54
ligger	lies (vb)	473	74
stik	due	217	6
nok	probably	74	18
stykke	piece	64	44
faktisk	actually	59	36
udsigtspunkt	vantage point	43	35
kirkegård	churchyard	39	62
byggelegepladsen	the playground	35	66
måske	perhaps	33	3
vagtpost	guard post	32	6
rigtigt	right	31	61
telefonboksen	the telephone booth	31	10
tak	thanks	28	14
kornmarker	fields	27	44
marker	fields	27	30
direkte	directly	26	8
pakhus	storage facility	26	27
kirkegården	the churchyard	25	44
kigger	looks (vb)	24	75
telefonboks	telephone booth	17	12
løveparken	the lionpark	16	13
ligge	lie (vb)	16	38
stråtækt	thatched	16	0
løvepark	lionpark	14	21
udsigtspunktet	the vantage point	14	14
byggelegeplads	playground	13	77
seks	six	13	15
gik	went	12	0
sikkert	certainly	11	91
stadigvæk	still	11	9
takker	thanx	11	9
ganske	quite	10	20
hvilken	which	10	30
ryggen	the back	10	10
teksten	the text	10	30

70 of the word forms containing (g) occur only once, and it is of course impossible to say anything general about such rarely occurring word forms. However, all tokens may of course contribute to the analysis of the influence of phonetic and extra-linguistic factors on reduction of (g), and therefore all are retained in the dataset (cf. Appendix A). Only a small set of word forms are more likely to contain reduced variants of (g) rather than the standard variant [g̊], in the sense that more than 50 % of the tokens of these word forms contain a reduced variant. The word forms prone to (g) reduction are: ‘ikke, ligger, kirkegård, byggelegepladsen, rigtigt, kigger, byggelegeplads, sikkert’ [ˈeg̊ə ˈleg̊ɐ ˈkʰiɐ̯g̊əg̊ɒːˀ ˈb̥yg̊əlɑjəpʰlasən ˈʁæg̊d̥id̥ kʰig̊ɐ ˈb̥yg̊əlɑjəpʰlas ˈseg̊ɐd̥]. For 7 of these 8 word forms, (g) is in intervocalic position, and some are likely to be unstressed due to unit accentuation, e.g. ‘ikke, ligger, kigger’. It is therefore probable that an effect of segmental context and stress will be found in the statistical modeling.

6.12 A MODEL OF THE REDUCTION OF (g)

The fixed effects factors in the mixed effects model of reduction of (g) are given in table below, together with their reference levels, where relevant. The random effects factors were subject and word form.

Table 6.18 – Fixed effect factors in the model of (g) reduction

Fixed effect	**Reference level**
token stress	stressed(yes)
local articulation rate	n.a. (continuous)
preceding segment	Consonant
following segment	Pause
position in prosodic phrase	non-marginal
word boundary	Yes
Decade	n.a. (continuous)
Gender	Female

A pure fixed effect model shows no risk of collinearity. The mixed effects model specified in the table on the left models the probability for reduction of (g) in unstressed, word internal syllables by men, while controlling for the effects of local articulation rate, the preceding and following segment, position in the prosodic phrase and age of speakers. The random effect subject controls for influence from individual speakers, and the random effect word form controls for the influence from individual words. All interactions were also tested.

6.13 CHARACTERISTICS OF THE MODEL

The goodness of fit characteristics of the model are: C = 0.86 and Dxy = 0.73, indicating a relatively good fit. Gender but not age emerges as significant: male speakers are slightly more likely to reduce (g) than female speakers (estimate = 0.54, p =0.04). The interaction of stress and position in the word emerged as significant as did an interaction of the preceding and the following segment. Local articulation rate is the only other factor to emerge as significant, and as with the remaining phonetic variables studied, it shows the expected effect of an increasing probability for reduction of (g) as articulation rate increases (estimate = 0.255, $p < 0.001$).

In order to interpret the interactions that emerge as significant in the model I will illustrate each of them in turn, starting with the interaction of stress and position in the word. The estimates have been converted into probabilities and these have been plotted for the factor stress in the two graphs below, where the graph on the left shows the probabilities for deletion by stress in word internal position, and the graph on the right shows the probabilities for deletion by stress in word final position.

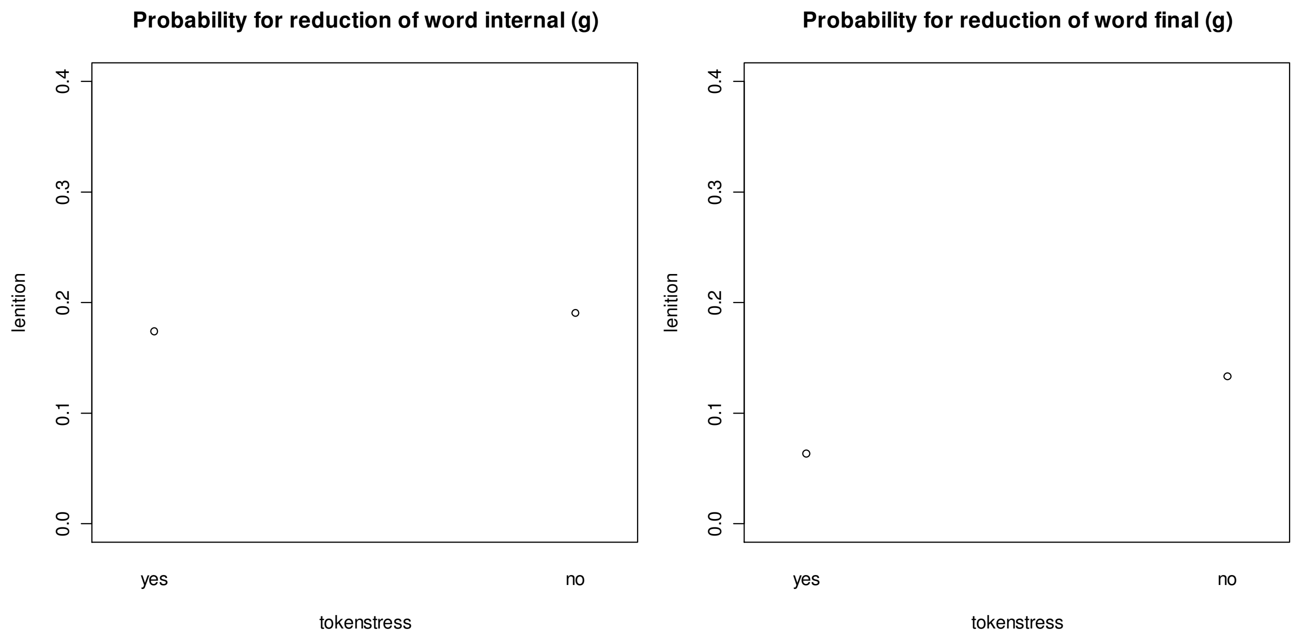


Figure 6.3 – The probabilities for reduction of (g) by stress in the two conditions word internal (on the left) and word final (on the right)

The graphs show that reduction of (g) is more likely in word internal position than in word final position. This effect of position in the word is so strong that there is no difference in the probability for reduction of (g) in stressed and unstressed syllables in word internal position: reduction is equally likely when words like 'ligger' and 'kigger' are realized with stress as when they lose stress due to unit accentuation. When (g) occurs in word final position, as in 'nok, tak, stik' [nʌg̊ tˢɑg̊ sd̥eg̊] *probably, thanks, due*, reduction is somewhat inhibited by stress, i.e. the probability for reduction is greater when the syllable is unstressed than when it is stressed. This effect can be exemplified by inspecting the rate of deletion of the word 'ikke' [ˈeg̊ə] *not*. When 'ikke' loses its final schwa, (g) occurs in word final position, and then it is more likely to be reduced in one of the 156 unstressed tokens of ‘ikke’, than in one of the 317 stressed tokens of ‘ikke’.

The interaction between preceding and following segment also emerged as significant. In order to analyze this effect the probabilities for reduction of (g) by the type of the following segment is shown in the graphs below for postvocalic (g) on the left and for post-consonantal (g) on the right. The interaction shows that reduction of (g) is generally less likely in post-consonantal position than in post-vocalic position, e.g. reduction is less likely in words like ‘hvilken’ [ˈvelg̊ən] *which* and ‘ganske’ [ˈg̊ansg̊ə] *quite*, than in words like ‘kigge’ [ˈk^{h}ig̊ə] *look* and ‘sikkert’ [ˈseg̊ərd] *certainly*. Furthermore, the interaction reveals that while the probability for (g) reduction remains the same in post-consonantal position irrespective of the type of the following segment, this is not true of post-vocalic (g). Reduction is much more likely for post-vocalic tokens when the following segment is a vowel, than when it is a consonant, and post-vocalic (g) is just as likely to be reduced in pre-pausal position as it is before a vowel. In other words, reduction of (g) is most likely in intervocalic position.

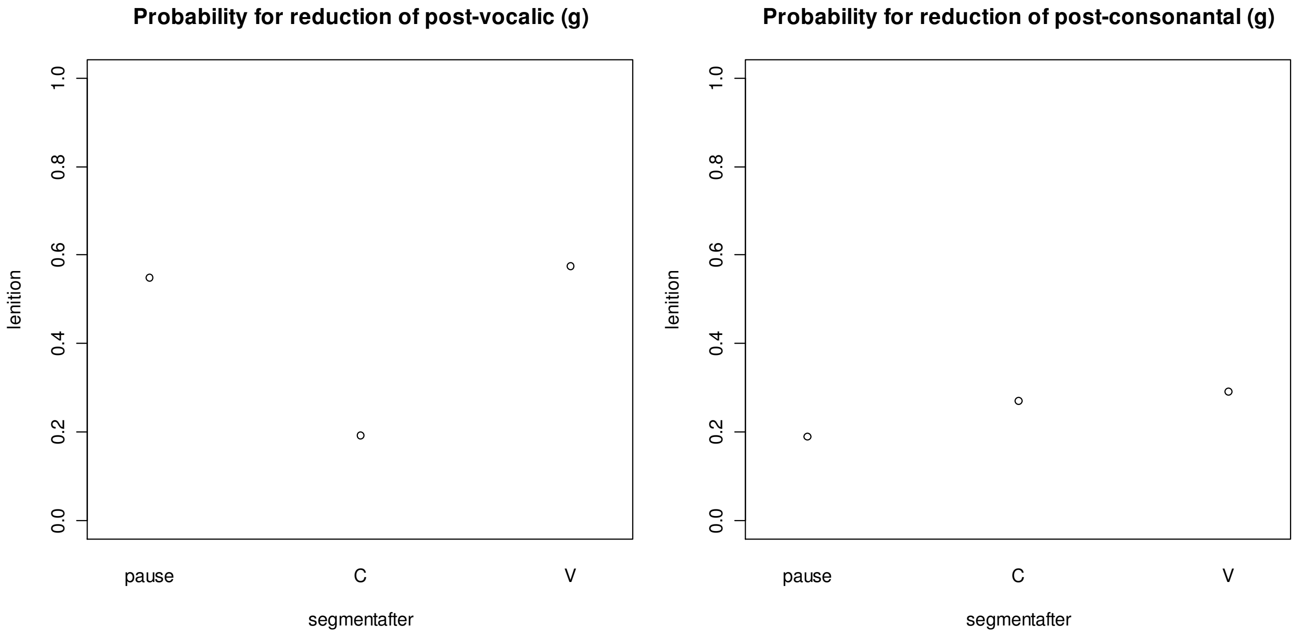


Figure 6.4 – The partial effect of the following segment in the two conditions post-vocalic (on the left) and post-consonantal (on the right)

So while the reduction of (g) is the same in ‘hvilken’ regardless of whether the schwa of the second syllable is assimilated to the following [n] or not, reduction of (g) is more likely in ‘ligger’ [ˈleg̊ɐ] *lie (vb.)* than in ‘direkte’ [ˈd̥iʁæg̊d̥ə] *direct(ly)*. The effect of the surrounding segments is not affected by the presence of a word boundary, meaning that reduction of (g) in ‘stik, nok, tak’ is more likely when the following word begins with a vowel than when it begins with a consonant, e.g. 'stik' [sd̥eg̊] *due* is more likely to be realized as [sd̥eɣ] in the word sequence 'stik øst' [sd̥eg̊ øsd̥] *due east* than in the word sequence 'stik nord' [sd̥eg̊ noɐ̯ˀ] *due north*. The similarity of the effect of the following segment for both positions in the word suggests a general effect of target undershoot in intervocalic environments. It should be noted that occurrence of (g) between two consonants is very rare: only 57 tokens are attested in the corpus, and while the word 'tænkte' [ˈtˢɛŋg̊d̥ə] *thought (past ptcpl)* is one of them, most of them arise from the assimilation of a schwa after the (g) to a following sonorant, such as in 'vinkel' [ˈveŋˀg̊əl] *angle* or 'hvilken' [ˈvelg̊ən] *which*.

As was done for (ɒw) and (b), it is possible to relate the findings for the effect of the segmental context on (g) reduction to phonetic resyllabification of the phonologically syllable final (g). Since reduction is generally least likely before a consonant, regardless of the nature of the segment preceding (g), it could be hypothesized that reduction is most likely for tokens of (g) that can form the onset of the following syllable. Following the criteria for phonetic resyllabification described previously, all tokens that are prevocalic in the distinct pronunciation of the word forms were ascribed to onset position in the (following) phonetic syllable. Word final as well as word internal pre-consonantal tokens (and tokens immediately followed by a morpheme boundary) remain in coda position in the phonetic syllable. This shifts 1534 tokens into onset position and leaves 807 in coda position. 874 of the tokens in onset position are reduced, whereas only 147 of the tokens in coda position are reduced. This difference is significant according to a chi-square test ($p < 0.0001$), but as with (b) it shows the unexpected pattern of reduction being more likely in onset than in coda – 57 % of onset tokens are reduced, whereas only 18 % of tokens in coda are reduced. Further support for the validity of phonological syllabification comes from an examination of the reduction of distinct [g̊] in onsets of the phonological syllable, i.e. the tokens of the phoneme /g/ that are not included in the dataset for (g), since (g) is defined as tokens of /g/ in the coda of the phonological syllable. For initial /g/, reduction is extremely rare: 99 of the 2321 tokens of syllable initial /g/ are reduced, a mere 4 %, less in fact than the reduction of /k/, which is reduced in 13 % of tokens (and is treated below in the modeling of the variable (k)). I will not go into detail with the reduction of syllable initial /g/, but I find it interesting to note that none of the factors explored for the variables that are the focus of the present investigation can be seen to have a statistically significant influence on the reduction of initial /g/ when a mixed-effects logistic regression model is fit to the dataset for syllable initial /g/. The reduction of syllable initial /g/ is most common in the phrase 'så går du' [ˈsʌ g̊ɒː d̥u] *then you go* in which the initial [d̥] of 'du' is also commonly reduced to a tap or even an alveolar approximant, [ɹ], i.e. the reduction is very common in a highly frequent phrase, and one that is particularly predictable given that the tokens come from recordings of map tasks. However, reduction to [ɣ] also occurs in content words like 'guldminen' [ˈg̊ulmiːnən] *the gold mine* and 'girafferne' [g̊iˈʁɑfɐnə] *the giraffes*, so the process is not particular to high frequent phrases where the word containing /g/ is predictable.

Naturally, the resyllabification of (g) has consequences for the analysis of stress, since all of the tokens that may be reassigned to the onset of the following phonetic syllable should be counted among the tokens in unstressed syllables. Recall that there was an interaction of stress and position in the word before resyllabification, with no statistically significant difference between the probability for reduction of (g) between stressed and unstressed syllables in word internal position, but a decreased probability for reduction of (g) in stressed syllables in word final position, i.e. in monosyllabic words with (g) in the coda. Disregarding position in the word for the present purpose of examining the validity of phonetic resyllabification for elucidating the contribution from stress on reduction of (g), a simple chi-square test shows that the difference between the proportion of reduced tokens in stressed and unstressed phonological syllables is small, 49 % are reduced in unstressed syllables, 42 % in stressed syllables, but statistically significant: $p = 0.003$. After resyllabification this difference between tokens in stressed and unstressed syllables is much stronger: 89 % of tokens in unstressed syllables are reduced, whereas only 11 % of tokens in stressed syllables are, and this difference is, of course, significant ($p < 0.0001$). This increased tendency for reduction of (g) after phonetic resyllabification warrants an update of the model in which the resyllabified tokens of (g) are added to the set of tokens occurring in unstressed phonetic syllables. Fitting this model to the dataset for (g) shows that the interaction of stress and position in the word is no longer significant. Stress emerges as a tendency ($p = 0.069$), whereas position in the word emerges as a main effect: reduction is more likely in word internal position than in word final position. The interaction between the preceding and the following segment remains significant, with intervocalic position still being the context in which (g) is most likely to be reduced, and the effect of local articulation rate also remains significant.

As with (b), resyllabification in itself cannot be interpreted as a direct effect, since post-consonantal (g) is not more likely to be reduced before a vowel than before a consonant, and indeed reduction of post-vocalic (g) is equally likely before vowels and pauses, e.g. (g) is equally likely to be reduced in both pre-pausal 'tak' [tˢɑg̊] *thanks* and pre-vocalic 'takker' [ˈtˢɑg̊ɐ] *thanx*. Instead, the effects of the segmental environment show that reduction of (g) is most likely when both of the neighboring segments require a relatively open vocal tract for articulation, which may then lead to undershoot of the target for (g).

6.13.1 Further analysis of the following segment

With such a high rate of reduction, it is possible to study the effects of different types of vowels and consonants statistically. I will start by exploring the role of the manner of articulation of the following consonant, even though this context is where reduction of (g) is least likely to occur. A model replacing the factor 'following segment' with manner of articulation of the following segment is fit to the subset of tokens followed by a consonant (1094 tokens, a mere 227 of which are reduced). The model shows that reduction is least likely before stops, and that there is no difference between the likelihood for reduction of (g) before approximants, sonorants and fricative. In the analysis of the possible effect of place of articulation of the following consonant, the stops are therefore excluded, reducing the subset further to 773 tokens, 171 of which are reduced. Fitting a model to this subset with the factor encoding the place of articulation of the following consonant reveals that reduction of (g) is least likely before dorsals and most likely before labials, coronals and glottals, with no difference between the latter three places of articulation. For following vowels, there is no effect of place of articulation, i.e. reduction of (g) is equally likely before front, back and centralized vowels.

It is interesting to note that the apparent inhibitory effect of following dorsals is clearest for sonorants: (g) is least likely to be reduced *before* velar sonorants, i.e. [ŋ̩]. This can be interpreted as a tendency for progressive assimilation of (g) on a following sonorant: rather than weakening the closure of the (g), the place of articulation spreads to the following sonorant. All of the following velar sonorants, except one, are word internal [n]'s which have become syllabic by assimilation of an intervening schwa, and then velarized to [ŋ] by the preceding variant of (g). Reduction, and even deletion, does occur in this context, but only rarely – 6 of the 49 tokens with a following syllabic [ŋ̩] are reduced or completely deleted. As with (b), one might wonder whether deletion of (g) is the result of a more extensive process of syllable deletion. Recall that deletion of (b) was almost exclusively attested in the word 'simpelthen' where the entire second syllable of the word was deleted (cf. above). Deletion of (g) does appear to be the result of a more extensive process of syllable deletion in a few word forms. When the (g) in 'kirkegård' [ˈkʰiɐ̯g̊əg̊ɒːˀ] is deleted, the entire second syllable is deleted, resulting in the overt form [ˈkʰiɐ̯g̊ɒːˀ][15].

[15] Note that the second [g̊] in the distinct form of 'kirkegård' is in syllable initial position, and therefore does not constitute a token of (g), but of the phoneme /g/, which, as described above is only very rarely reduced.

Similarly, the deletion of the medial (g) in the word 'faktisk' [ˈfɑg̊tˢisg̊] *actually* is occasionally the result of reduction of the disyllabic word to a monosyllable [ˈfɑjs] or [ˈfɑs]. Unlike the word 'simpelthen', in which deletion of the second syllable was predominant, this is not the case for neither the word 'kirkegård' nor the word 'faktisk'. Furthermore, in many of the words in which (g) is deleted, none of the adjacent segments are reduced, viz. the realization [ˈkʰɑtˢus] for 'kaktus' *cactus*, which in distinct pronunciation would be [ˈkʰɑg̊tˢus], or [ˈg̊ansə] for the word 'ganske' *quite* which would be [ˈg̊ansg̊ə] in distinct pronunciation. In other words, whereas deletion of (b) could be seen to be a consequence of the process of syllable loss, deletion of (g) is more clearly a process of segment deletion.

In summary, reduction of (g) is a relatively common process. In contradistinction to (b) it is affected by many articulatory processes: the tendency for reduction of (g) increases with articulation rate. It is most frequently reduced before and after vowels, and reduction occurs more frequently in word internal than in word final position.

6.14 VARIATION OF (k)

There are 2209 tokens of words containing (k) in the dataset. Table 6.19 below gives the proportion and total number of the variants of (k).

Table 6.19– Variants of (k)										
Variant	kʰ	ɣ	g̊	x	deleted	ɦ	h	ç	j	kj
Number	1903	118	103	79	1	1	1	1	1	1
proportion	86 %	5 %	4 %	4 %	1 %					

Clearly, reduction of (k) is more limited than reduction of (g). The main processes observed for (k) are loss of aspiration and reduction to a fricative, as with (p), but it is also clear that it does not make sense to distinguish between different types of reduction processes for (k). Therefore, all variants other than [kʰ] will be treated as reduced in the following analysis.

There are 130 word forms with (k) in the DanPASS Dialogues, the vast majority of which only occur with the standard variant [kʰ] for (k). Table 6.20 below shows the word forms that occur at least 10 times in the corpus and where reduction of (k) does occur, together with the total number of tokens and rate of reduction for each word form (the full dataset is given in Appendix A).

Table 6.20 – word forms with reduced (k) (n> or = 10)			
Word form	**English**	**sum**	**% reduced**
okay	okay	582	12
kan	can	252	25
kommer	come (pres)	234	6
klippe	cliff	74	1
kommet	come (past.ptcpl)	62	5
kolonihaver	allotments (gardens)	32	9
parkerede	parked	28	7
omkring	around	27	7
kortet	the map	25	8
kunne	could	23	22
kalkstensklipper	limestone cliffs	18	6
bjergkløften	the ravine	15	7

The table reveals that reduction is not prevalent in any word forms. Inspection of the data revealed that the word internal [kʰ] was categorically reduced in the word 'cirka' [ˈsiɐ̯kʰa] *approximately*, and for the second instance of (k) in all forms of the lexeme 'krokodille' [kʰʁokʰoˈd̥ilə] *crocodile* irrespective of inflection (the word initial (k) in 'krokodille' is never reduced). Even though these tokens of expected [kʰ] in distinct pronunciation were not only realized as [ɡ̊] but also sometimes reduced further to [ɣ], it was still decided to exclude the tokens from the dataset for the variable (k) on the grounds that categorical use of non-distinct variants of the variable in such relatively frequently occurring words can be interpreted as a sign that the realizations [ˈsiɐ̯kʰa] and [kʰʁokʰoˈd̥ilə] for 'cirka' and 'krokodille' constitute hyperdistinct pronunciations of these words. Excluding these word forms reduces the dataset to 2083 tokens of (k). No other word forms show reduction of (k) to be the predominant realization, but there appears to be a higher rate of deletion for word internal tokens of (k) in intervocalic position and between [ɐ̯] and a vowel, (just as the words with categorical reduction show), as can be seen from the relatively high reduction rates for words like 'forkert, fokus, sekund' [fʌˈkʰeɐ̯ˀd̥ ˈfoːkʰus seˈkʰɔnˀd̥] *wrong, focus, second* (cf. the full dataset in Appendix A). The modeling of (k) reduction will help to determine to what extent intervocalic position promotes reduction of word initial (k).

6.15 A MODEL OF REDUCTION OF (k)

A model with the fixed effects factors stress, position in the word, local articulation rate, preceding segment, following segment, position in the prosodic phrase, decade of birth and gender is fit to the tokens of (k). As always, subject and word form are included as random factors. The factors for preceding and following segment classify the adjacent segments as either consonants or vowels.

The goodness-of-fit characteristics of the model are: C = 0.93 and Dxy = 0.87, indicating an excellent fit. Neither of the extra-linguistic factors emerges as significant, nor do all of the linguistic ones. No interactions emerged as significant. The estimates are given for the significant factors in the table of coefficients below.

Table 6.21 – coefficients for significant factors in the model of (k) reduction				
Factor	Estimate	Std. Error	z value	Pr(>lzl)
(Intercept)	-12.89023	2.03219	-6.343	2.25e-10 ***
Preceding C	1.90175	0.80522	2.362	0.018187 *
Preceding V	2.82143	0.82143	3.435	0.000593 ***
Local articulation rate	0.25828	0.05510	4.687	2.77e-06 ***
Word internal	2.52801	0.93444	2.705	0.006823 **

The table of coefficients shows that the probability for reduction increases with local articulation rate and that word internal tokens are more likely to be reduced than tokens in absolute word initial position, e.g. reduction of (k) is more likely in 'okay' [ˈɒwˈkʰɛj] *okay* than in 'klippe' [ˈkʰlebə̥] *cliff*. Preceding consonants and vowels both increase the probability of reduction relative to preceding pauses. The difference between preceding consonants and vowels is also statistically significant (estimate = -1.22142, p =0.022), meaning that reduction of (k) is more likely after a vowel than after a consonant, e.g. it is more likely in 'parkerede' [pʰɑˈkʰeːˀʁɐðə] *parked* than in 'kalkstensklipper' [ˈkʰalg̊sd̥eːnskʰlebɐ̥] *limestone cliffs* and also in e.g. 'kommer' [ˈkʰʌmˀɐ] *comes* when this word is preceded by a word ending in a vowel, as in the sequences 'så kommer' *then comes* and 'du kommer' *you come* than when it is preceded by a word ending a consonants as in the sequence 'man kommer' *one comes*. Due to a high risk of collinearity, it is not possible to examine the details of the preceding segments and possible differences between different types of consonants and vowels.

In summary, (k) is more prone to reduction than (p), but reduction is still relatively rare. The pattern of factors that influence reduction of (k) and (g) are somewhat different, with (g) reduction being sensitive to both the preceding and the following segment, and also showing a slight effect of the gender of the speaker.

6.16 THE VARIATION OF UNDERLYING FINAL |t|

Word final |t| is canonically produced [d̥] in both isolated words and running speech, but may also be produced as [tˢ] in pre-pausal position. Phonologically, it is /d/, which can also occur in word and syllable initial position, whereas /t/ is restricted to word and syllable initial position. This analysis focuses on the morphophoneme |t| in word final position because it will allow inclusion of the factor of morphological category in addition to the linguistic and extra-linguistic factors explored in the analyses of the other consonants in the investigation of synchronic variation in contemporary Copenhagen Danish. In keeping with the representation of the other plosives I will refer to the variable as (d), but note that this variable is more restricted than the variables (b) and (g), since it only concerns word final tokens, and not syllable final word internal ones.

6.17 VARIANTS OF (d)

All 3385 tokens of (d) have been extracted from the DanPASS Dialogues. There are four different variants of the variable. They are given in table 6.22 below together with their distribution in absolute numbers and as a percentage of the total amount of tokens of the variable.

Table 6.22 – Variants of (d)				
Variant	d̥	tˢ	deleted	θ
Number	1772	950	623	40
%	52 %	28 %	19 %	1 %

A little more than half of the tokens are realized as the canonical standard variant [d̥]. The majority of the non-standard tokens are examples of strengthening, i.e. where (d) is realized as [tˢ]. Out of all of the tokens, only 20 % are reduced, with the majority of these constituting tokens of words where (d) has been deleted. In light of this distribution, only the process of deletion will be modeled in the statistical analyses, i.e. tokens realized as [θ] are included as non-elided.

There are 176 different word forms ending in (d) in the corpus, and (d) is deleted in 87 of them. Categorical deletion occurs in 24 word forms, but none of these occur more than once in the corpus, meaning that the proportion is unreliable. The table below gives the rate of deletion for the word forms in which (d) deletion occurs, but only for word forms that occur 10 or more times in the corpus (the full dataset is given in Appendix A).

Table 6.23 word forms with (d) deletion (n > or =10)

Word form	English	n	% deleted
et	a/an	429	14
øst	east	327	5
vest	west	270	5
lidt	little	251	3
godt	good	233	6
at	to/that	175	59
rundt	around	135	11
sydvest	south west	105	8
sydøst	south east	87	6
nordvest	north west	74	5
helt	all	71	51
langt	far	64	28
nordøst	north east	55	4
start	start	51	4
udsigtspunkt	vantage point	43	9
kort	short	42	2
forladt	abandoned	39	3
vagtpost	guard post	32	6
midt	middle	29	7
ret	quite	28	4
mit	my	27	4
rigtigt	right	26	73
dit	your	23	13

Table 6.23 - continued

Word form	English	n	% deleted
bjergkløft	ravine	22	5
først	first	22	27
set	seen	18	17
fint	fine	17	47
højt	high	17	6
hvert	each	17	94
nærmest	nearly	16	25
skarpt	sharply	16	56
stråtækt	thatched	16	25
punkt	point	15	40
tidspunkt	point in time	15	60
nyt	new	14	7
halvt	half	11	27
mest	most	11	27
sikkert	surely	11	18
snart	soon	11	18
afbrændt	scorched	10	10

6.17.1 Reducing the data set

As with the other phonetic variables, the dataset is heavily unbalanced with respect to word forms: a few word forms occur many times, but most word forms occur only a few times. 6 word forms show predominant deletion of (d), i.e. (d) is deleted in more than 50 % of tokens of these word forms, namely: 'at, helt, rigtigt, hvert, skarpt, tidspunkt' [ad̥ heːʔld̥ ˈʁæg̊d̥id̥ væɐ̯ʔd̥ sg̊ɑːb̥d̥ ˈtsiðsphɔŋʔd̥] to (infinitive marker), quite, every, sharp, point in time. The infinitive marker 'at' deserves special attention, before modelling can be carried out. As with the realizations of 'og' [ɒw] *and* as [ʌ], it is perhaps also more reasonable to only count as genuine (d) deletion the lack of (d) in tokens of 'at' that are realized with the vowel [a].The word 'at' may also be realized as [ʌ] – making it completely homophonous with the conjunction 'og' - and such tokens may be thought of as realizations of an alternate word form [ʌ] for 'at'. Accepting this premise, the dataset is reduced by 132 tokens, leading to

(3385 - 132) 3253 tokens in total. Of the remaining 175 tokens of 'at', (d) is deleted in 104, so the process is still fairly common even with the more conservative definition of the variable.

6.18 A MODEL OF (d) DELETION

The factor of position in the word, which was used to model the reduction of |p| and |k|, becomes irrelevant here – all tokens are by definition followed by a word boundary. Instead, it is possible to include the factor of morphological category, since (d) may constitute a separate, bound morpheme, namely the suffixes for adverbs, for the past participle and for neuter gender of adjectives and it may also be a part of the suffix for definite forms of nouns in neuter gender. All the tokens have therefore been coded for morphological category with a script using the part-of-speech tagging in the corpus and thereby classifying all tokens in the dataset of word final |t|occurring in adverbs, adjectives and past participles and nouns ending in |ət| as 'bound' and all other tokens as 'free'. All tokens of bound morphemes were also coded for word class, i.e. adjectives were classified as adjectives, adverbs as adverbs and definite singular suffixes in nouns as nouns, in order to investigate whether some morphological categories show greater likelihood of (d) deletion than others. Note that no past participles ending in (d) occur in the DanPASS Dialogues, and therefore only the three categories mentioned above were relevant. All tokens that have been classified as 'free' in the initial coding of morphological category were given the classification 'STEM'. A mixed-effects multiple logistic regression model was fit to the data with the factors given in table 6.24 together with their reference levels. The procedure models probability of deletion in unstressed syllables, in bound morphs by male speakers while controlling for segmental context, articulation rate, position in the prosodic phrase, and age. A pure fixed effects model was constructed in order to check for the risk of collinearity; there were no serious grounds for concern, although two levels of the factor segmental context had variance inflation factors slightly above 2. Both the factors 'subject' and 'word form' were added as random effects, even though one purpose of the study of (d) deletion is to explore differences between different morphological categories of (d),

Table 6.24 – Factors in the first model of (d) deletion

Dependent variable	**Reference level**
'deletion of (d)'	No
Fixed effect	**Reference level**
'token stress'	stressed (yes)
'preceding segment'	Vowel
'following segment'	Pause
'local articulation rate'	n.a. (continuous)
'morpheme'	STEM
'stød'	stød (yes)
'decade'	n.a. (continuous)
'gender'	Female

generally whether it is part of a stem or a suffix. But since such an effect might be a product of only a few (perhaps highly frequent) inflected word forms, it is desirable still to control for the random effect of word form. This should entail that any effect of morphological category that might emerge as significant in the model will be statistically more robust, just as is the case for effects of the phonetic environment.

6.19 Characteristics of the model of (d) deletion

The goodness of fit characteristics of the model are C = 0.90 and Dxy = 0.81 indicating a very good fit. The table of coefficients is given below. Age emerges as significant, and the effect is in the opposite direction of what is normally assumed: the coefficient is negative meaning that probability for deletion of (d) *decreases* with decade of birth. Local articulation rate has the expected effect, probability increases in faster speech, and there is also a simple effect of morphological category. The phonetic context is rather more complex as there are two significant interactions in the model, which are illustrated figure 6.5 and figure 6.6 below.

Table 6.25 - coefficients for significant factors in the model of (d) deletion

Factor	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-6.56258	0.84086	-7.805	5.97e-15 ***
Adjective	1.00400	0.29059	3.455	0.000550 ***
Adverb	-0.11881	0.26557	-0.447	0.654603
Noun (def. sg.)	0.90146	1.28101	0.704	0.481613
Local articulation rate	0.47803	0.05509	8.677	< 2e-16 ***
Age	-0.02450	0.01027	-2.385	0.017081 *
Unstressed:Preceding C	-1.63543	0.42801	-3.821	0.000133 ***
PrecedingC:Following C	0.91008	0.40873	2.227	0.025975 *
Preceding C:Following V	0.54669	0.61050	0.895	0.370533

6.19.1 The interaction of stress and the preceding segment

The model reveals two complex interactions with respect to the phonetic context in which (d) occurs. First of all, there is an interaction of stress with the type of the preceding segment. The interaction is illustrated in the graphs below. Note that the y-axes are compressed relative to previous graphs. This has been done in order to more clearly visualize the fairly slight difference between conditions. The interaction reveals that there is no difference between the probability for (d) deletion in unstressed syllables depending on the type of the preceding segment: deletion of (d) is equally likely in a word like 'nordligt', where it is preceded by a vowel, and in a word like 'nærmest', where it is preceded by a consonant.

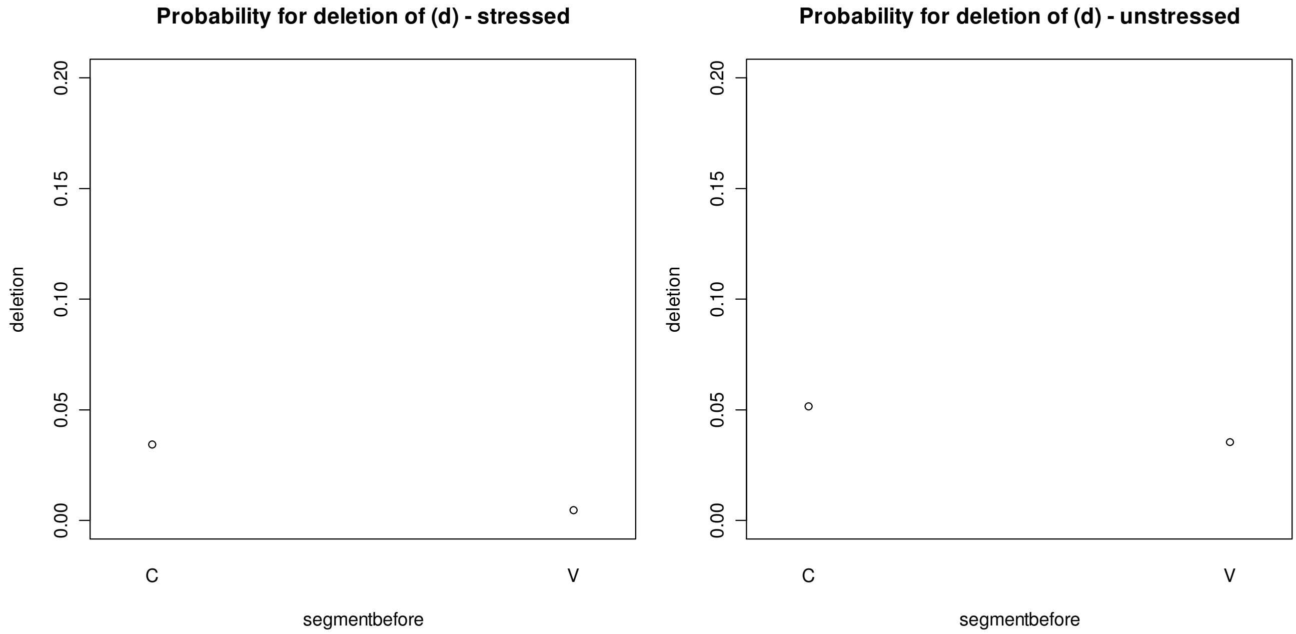
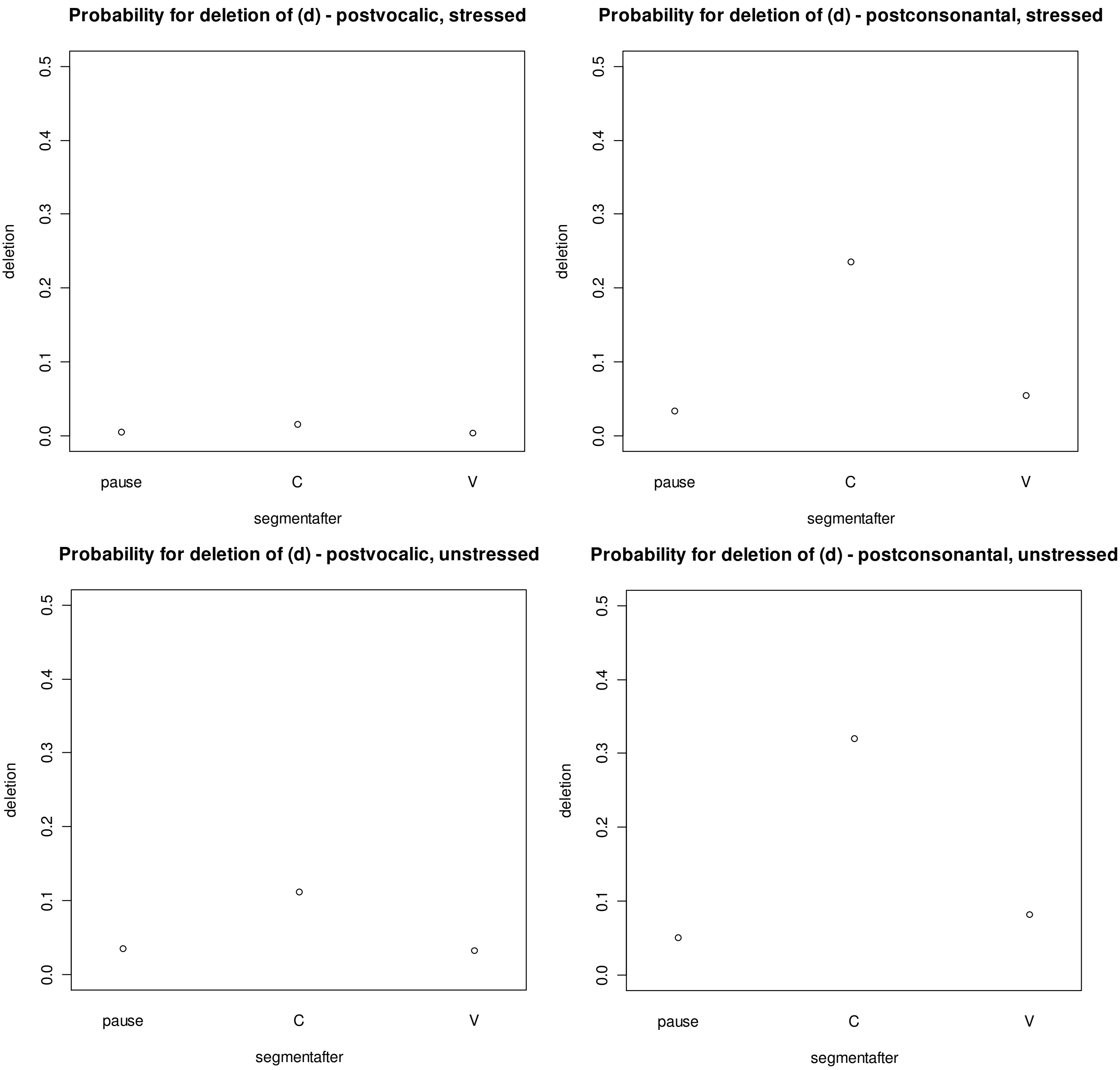


Figure 6.5 – The probability of deletion of (d) by the type of the preceding segment in the two conditions stressed (on the left) and unstressed (on the right)

For tokens in stressed syllables, however, there is a difference between preceding vowels and consonants: deletion of (d) is more likely in stressed syllables when the preceding segment is a consonant than when it is a vowel – in the latter context, deletion occurs in a mere 3 % of cases (25 of 723 tokens), whereas in post-consonantal position in stressed syllables, (d) deletion occurs in 14 % of cases (226 out of 1624 tokens). So (d) deletion is more likely in a word like 'helt', than in a word like 'godt'. This interaction suggests an overall tendency for segmental simplification of unstressed syllables irrespective of the segmental structure of the syllable, whereas for stressed syllables the increased probability for (d) deletion after consonants suggests a tendency for consonant cluster simplification in the syllable coda, whereas (d) is preserved when it is the only consonant in the coda. The interpretation of the effect of types of preceding segment must however also be viewed in the light of the type of the following segment, since there is also an interaction of the preceding and the following segment. While there is no interaction of stress with the following segment, the analysis of the interaction between surrounding segments is analyzed separately for stressed and unstressed syllables, given the effect of stress just shown above. The figure below illustrates the interactions of the preceding and following segment in the two conditions of stress.

Figure 6.6 – Partial effect of the following segment in each of four conditions: right: postvocalic position; left:post-consonantal position; top: stressed syllables; bottom: unstressed syllables.



The interaction of the surrounding segments further supports the interpretation of the influence of the phonetic factors on (d) deletion as a case of consonant cluster simplification, and suggests that this effect applies irrespective of an intervening word boundary since deletion of word final (d) is generally more likely when the onset of the following word begins with a consonant, whereby a consonant cluster can arise, than when the word final (d) occurs intervocalically. The only exception to this pattern is when the (d) occurs immediately after a vowel in a stressed syllable, in which case there is no difference between the probability for deletion of (d) with respect to the type of the following segment, as can be seen in the top left panel in the figure above, e.g. deletion of the (d) in stressed tokens of 'lidt' is equally likely – or, rather, unlikely – when it is followed by a word like 'op' and when it is followed by a word like 'mere' and even when it is followed by a pause. This fits well with the effect of stress seen in the interaction between stress and the preceding segment alone: the probability for (d) deletion was significantly smaller in stressed syllables when the (d) occurred immediately after a vowel than when it occurred after a consonant. This effect apparently persists irrespective of the nature of the following segment. However, in postvocalic position in *un*stressed syllables, deletion is significantly more likely if the following segment is a consonant rather than a vowel or a pause (bottom left panel,), e.g. 'rigtigt'[ˈʁæg̊d̥id̥] is more likely to be realized [ˈʁæg̊d̥i] when followed by 'ja' than when followed by 'og'. Since the probability for (d) deletion is the same in unstressed syllables irrespective of the type of the preceding segment, as was shown above, it would appear that the influence from a following consonant can now arise, i.e. the tendency for consonant cluster simplification, even across word boundaries, is no longer inhibited by the effect of stress. This pattern recurs in post-consonantal position irrespective of the stress of the syllable containing (d): when both neighboring segments are consonants, deletion is significantly more likely than when the following segment is a vowel or a pause. In this case then, a word like 'helt' is more likely to be realized as [heːˀl] when it is followed by a word like 'derop' than when it is followed by a word like 'op' (or when it is followed by a pause). Furthermore, while the effect of a following consonant can be seen even when the token of (d) is preceded by an unstressed vowel, the probability is greater when the (d) is preceded by a consonant, so the effect of a following consonant is greater for a word like 'nærmest' than for a word like 'rigtigt'.

Taken together, the influence from the segmental context can therefore be interpreted as a process of consonant cluster simplification, which may be inhibited by a preceding stressed vowel, but can be said to intensify the greater the number of (potential) consonants in the sequence

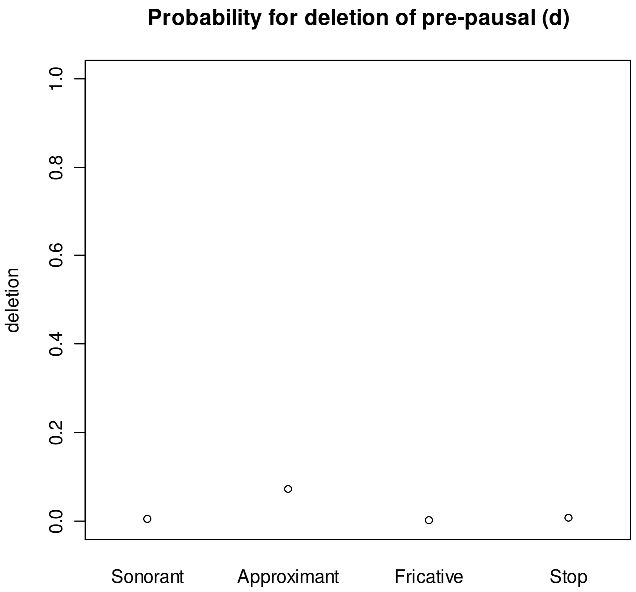
6.19.2 Further analysis of the segmental context

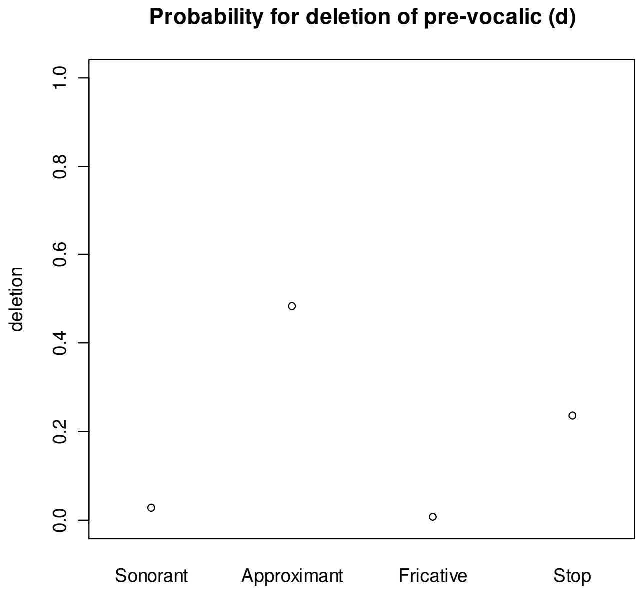
Deletion of a variable similar to (d) has been extensively studied in American English. These studies of so-called word final [t,d] deletion have shown that both place and manner of a preceding consonant are important factors in deletion of wordfinal (d). Guy (1980), Labov (1968) and Wolfram (1969) found that obstruents promoted [t,d] deletion, but sonorants inhibited it (according to Raymond et al (2006), p. 83), and both Fasold (1972) and Zue & Laferriere (1979) found that homorganic consonants favored deletion. Raymond et al (2006), in a study of word *internal* [t,d] deletion, found similar effects of manner for tokens in coda position, but found that deletion was more likely after non-homorganic obstruents (Raymond et al (2006), pp. 83-84) (Raymond et al (2006) also found quite different patterns for the preceding consonant for tokens of [t,d] in onset position, but since the present study only concerns word final, and hence coda, (d), these results are not discussed). The initial model of (d) deletion in stressed syllables in contemporary Copenhagen Danish follows the overall pattern of studies of American English in that a preceding consonant promotes deletion relative to a preceding vowel, but it remains to be seen how manner and place of articulation of the preceding consonant affects (d) deletion in contemporary Copenhagen Danish. For the following consonant, the studies referenced above, except Raymond et al (2006), found deletion to be less likely when followed by a sonorant, i.e. obstruents promote deletion of word final [t,d] in American English. Raymond et al (2006) on the other hand, found no clear connection between deletion of [t,d] in coda position and the sonority of the following segment – in fact, they found no clear correlation with any phonetic feature apart from an increased tendency for deletion when the following segment is a consonant rather than a vowel (Raymond et al. (2006), pp. 84 – 87). Clearly, studies on deletion of syllable and/or word final coronal stops in American English indicate an intriguing interplay of aspects of the segmental context. In the following, I shall attempt to provide an analysis of the role of manner and place of articulation of adjacent consonants in the deletion of (d) in contemporary Copenhagen Danish.

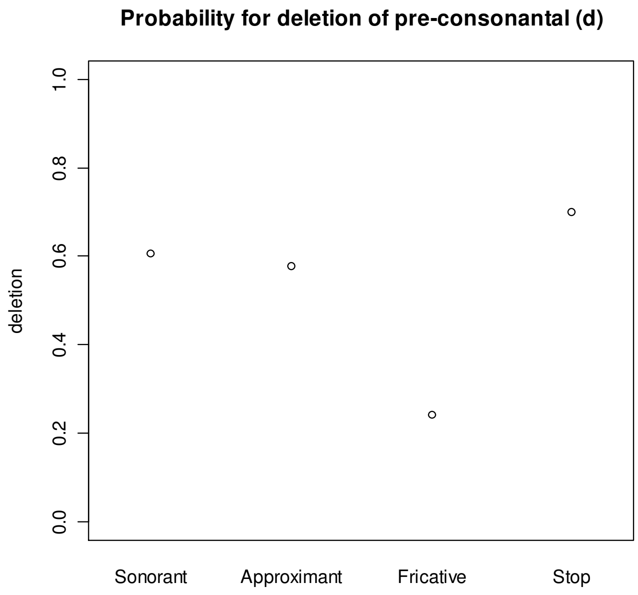
To investigate the influence from manner and place of articulation of the adjacent consonants, new models were fit to two subsets of the data: tokens preceded by a consonant and tokens followed by a consonant. Since there is no significant difference between the influence from a preceding consonant in stressed and unstressed syllables (this effect only applied to vowels, as shown above), tokens in both stressed and unstressed syllables were included in the subset of post-consonantal tokens, e.g. both stressed 'helt' and unstressed 'nærmest' were included. However, since pre-consonantal tokens were less likely to delete after a stressed vowel than after an unstressed vowel or a consonant, the analysis of the manner and place of articulation of the consonant following (d) is restricted to the subset of pre-consonantal tokens that do not occur immediately after a stressed vowel. The models were similar to the initial model, with the exception that the factor encoding the overall type of the neighboring segment was replaced by factors encoding the manner or place of articulation of the neighboring segment.

6.19.3 Manner of articulation of the preceding consonant

Rather than including an interaction between the manner of articulation of the preceding consonant with the type of segment after the (d) in order to take into account the interaction described above, the model for the manner of articulation of the preceding consonant was fit to three different subsets of post-consonantal (d), namely: pre-pausal (d), pre-vocalic (d) and pre-consonantal (d). Due to the interaction with the following type of segment it is to be expected that there will be differences across the three conditions. Therefore the partial effect of manner of articulation is presented separately for the three conditions. The reference level of the factor encoding manner of articulation was set to 'sonorant', in order to study whether sonorant can also be seen to inhibit deletion of (d), as was found for American English. If so, the remaining factor levels (approximant, fricative and stop) should all show positive coefficients, and hence higher levels of probabilities in the graphs below.





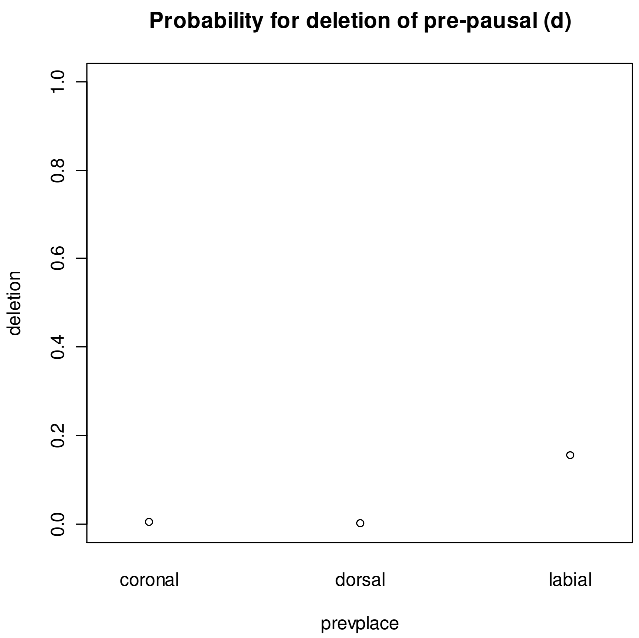


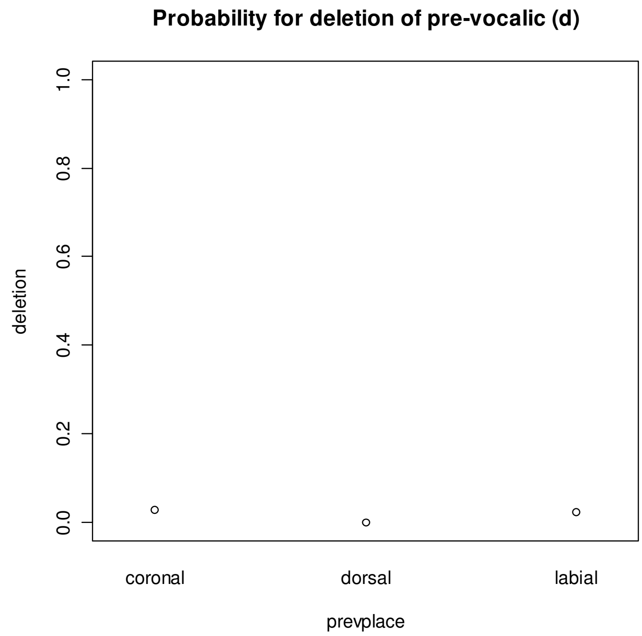
There is no effect of manner of articulation on pre-pausal tokens of post-consonantal (d), as shown in the upper left panel. None of the types of manner are significantly different from each other. For pre-vocalic (d) the pattern is reminiscent of that described for American

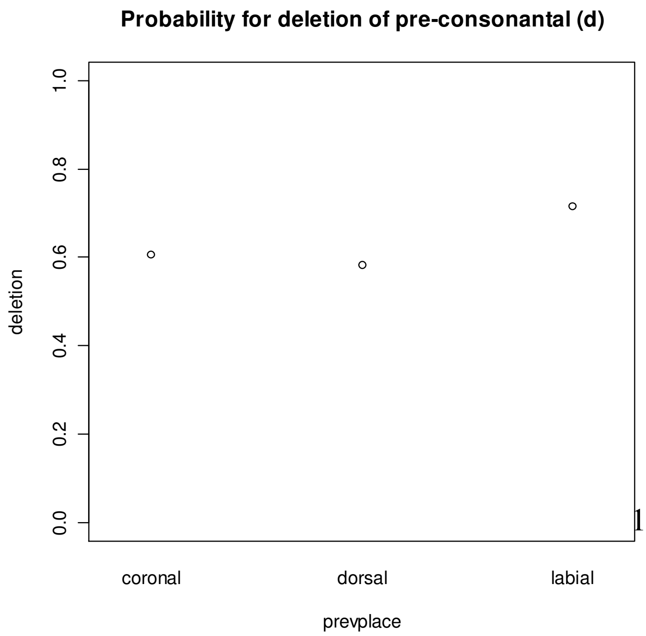
English: deletion is least likely after sonorants and fricatives and more likely after approximants and stops. However, for pre-consonantal (d), the pattern no longer resembles that found for American English, since deletion is now equally likely after sonorants, approximants and stops, and only fricatives are significantly different from sonorants. The general pattern, then, is that in contexts where (d) deletion does occur, it is least likely to occur after fricatives, e.g. deletion is less likely in words like ‘øst’, and ‘kløft’, than in words like ‘fint’, ‘hvert’ and ‘skarpt’.

6.19.4 Place of articulation of the preceding consonant

The analysis of the place of articulation of the preceding consonant was also carried out across the three conditions pre-pausal, pre-vocalic and pre-consonantal (d) seperately. The reference level for the factor encoding place of articulation was set to ‘coronal’ in order to study whether deletion of (d) is more likely after homorganic or after non-homorganic consonants.



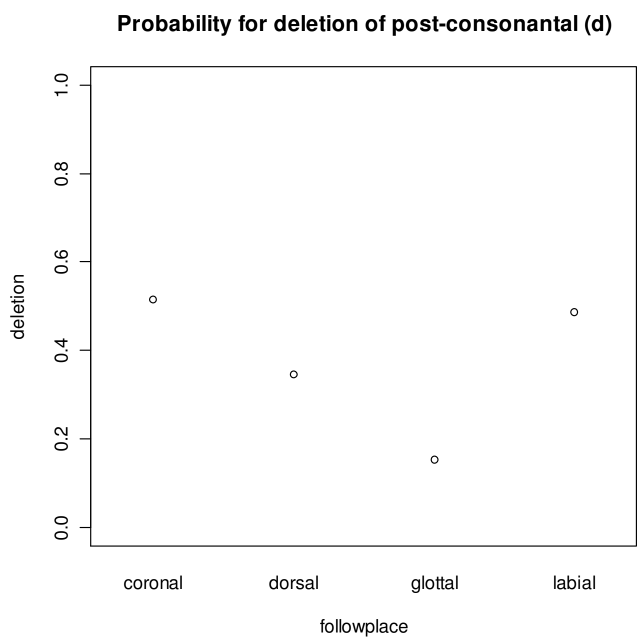




A significant effect of place of articulation of the preceding consonant only occurs in one condition, and this is, surprisingly, for pre-pausal tokens of (d). In this context, deletion is more likely after labials than after coronals and dorsals, i.e. more common in ‘skarpt’ [sg̊ɑːb̥d̥] *sharply*, than in ‘vest’ [vɛsd̥] *west* or ‘udsigtspunkt’ [ˈuðseg̊d̥spʰɔŋˀd̥] *vantage point*. None of the other contexts show a significant difference between the different places of articulation of preceding consonants. It appears then, that there is no effect of homorganicity of the preceding consonant on the deletion of (d) in contradistinction to the studies of a similar process in American English.

6.19.5 Manner and place of articulation of the following consonant

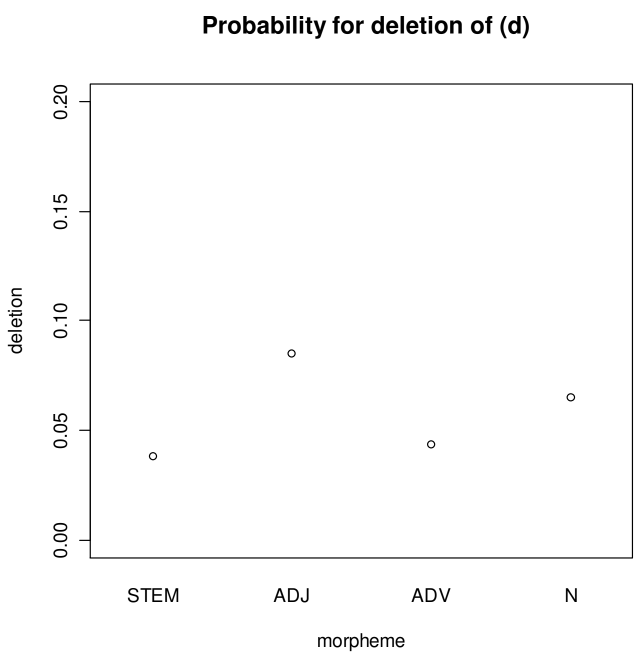
Recall that due to a difference between the influence of stressed and unstressed preceding vowel, the analysis of the manner and place of a following consonant occurring after a post-vocalic (d) must be carried out separately from post-consonantal (d).Furthermore, since a preceding stressed vowel almost entirely blocks the process of (d) deletion before a consonant, this analysis is restricted to the subset of tokens occurring after an unstressed vowel. Fitting a model with manner and place of the following consonant to the subset of pre-consonantal tokens of (d) occurring after unstressed vowels reveals no effect of either manner or place of articulation. In other words, deletion of (d) is equally likely before any type of consonant, which is to be expected since, as the bottom left panel in figure above shows, deletion is quite rare in this context. Deletion of (d) before a consonant is much more common when the (d) occurs after a consonant, i.e. in interconsonantal position. When the model is fit to the this subset of the data, there is still no effect of manner of articulation, i.e. deletion is equally likely before approximants, sonorants, fricatives and stops. But there is an effect of place of articulation of the following consonant, as illustrated in the graph below. According to the estimates for the factor encoding place of articulation, the only statistically significant differences are between glottals and labials and glottal and coronals. Dorsals are not statistically significantly different from any other place of articulation. Notice, however, that the effect of following glottals is inhibitory on the probability for (d) deletion, that is, deletion is more likely in a word like ‘langt’ [lɑŋˀd̥] *far*, when it is followed by ‘mod, ned, fra’ [moðˀ neðˀ frɑːˀ] *towards, down, from* than when it is followed by the word ‘hen’ [hɛnˀ] *towards*. The effect of place of articulation of the following consonant on deletion of inter-consonantal (d), does show that deletion



is quite likely before following coronals, however, since deletion is equally likely before dorsals and labials, it should not be interpreted as an increase due to the homorganicity of (d) and the following consonant. Rather, the analyses of articulatory aspects of a following consonant simply support the finding that there is an overall tendency for (d) deletion in inter-consonantal position, again supporting the interpretation of (d) deletion as a process of consonant cluster simplification.

6.19.6 Morphology

As mentioned above, there is an effect of morphological category of the tokens of (d). The partial effect is illustrated in the graph below. It should be noted that the only difference that can be observed to have a statistically significant effect is that between stems and adjectives – the categories adverbial and noun are not significantly different from either of these. In the case of nouns with a suffix this is clearly because there are too few tokens on which to base the analysis, since are only 6 tokens of nouns ending in (d) which is not part of the stem. The adverbial suffix is more richly represented in the data (731 tokens in all), and it does show a tendency to be less likely to be deleted than adjectival (d). The robust morphological effect, then, is for adjectival suffixes only, but as can also be seen from the graph below, the effect is rather small, and again the range of the y-axes has been compressed. It should be noted that the classification N for the factor 'morpheme' only refers to tokens of (d) that form part of the singular definite suffix for nouns, |ət|, and not to uninflected nouns ending in (d).



In summary, the largest effects on the probability for deletion of word final (d) are phonetic ones, with local articulation rate and stress exerting the largest influence. A range of effects are also seen from the surrounding segments, all of which suggest that the deletion of (d) is most likely to be a consequence of consonant cluster simplification even across the following word boundary. While a small effect of morphological category of the (d) token is also found, it seems too slight to warrant an interpretation as a consequence of linguistic processing.

6.20 General summary

The statistical modeling of the reduction of the phonetic variables studied in the DanPASS Dialogues show a variety of phonetic effects being involved in processes of reduction in spontaneous speech. While effects of local articulation rate and stress prevail, many of the processes seem to be only slightly affected by other phonetic and linguistic factors, although there is an indication that reduction of plosives is particularly sensitive to intervocalic segmental contexts. Very little effect is observed for the factor of position in the prosodic phrase, suggesting that effects found in laboratory speech of the relative strengthening of segments in the margins of phrases cannot be observed when the investigation is restricted to transcriptions of spontaneous speech. This is quite likely due to the relatively more fine grained correlates of segmental articulation in the studies of laboratory speech, e.g. duration of closure and larger linguopalatal contact, as compared to the binary variables employed in the present investigation, where annotation of degree of opening in reduced tokens of plosives in the transcriptions is lost, because all reduced variants have been grouped together.

It is worth noting that not all of the variables are equally likely to be reduced in running speech, and furthermore that whereas most of the plosive variables tended to show evidence of reduction, they were not very likely to be completely deleted. The exception to this is the variable (d), which shows an appreciable tendency for deletion, but very few tokens of mere reduction relative to the target articulation [d̥]. In the next section I will summarize the effects of the different phonetic factors that were included in the study of the phonetic variables in the DanPASS Dialogues.

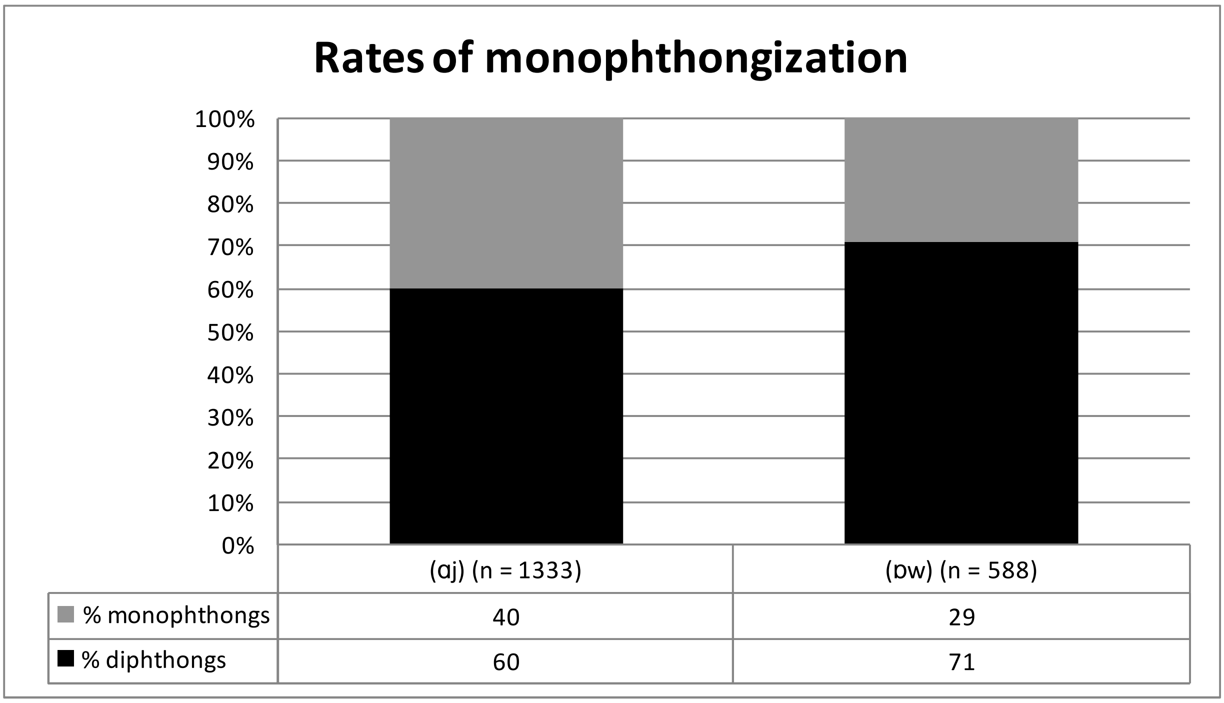
7. Phonetic effects in consonant reduction in contemporary Copenhagen Danish

This chapter collects the evidence on the influence from phonetic factors on the reduction of the consonants studied in the DanPASS Dialogues. The aim is to discuss the role of the various effects and to relate them to the process of speaking and to lexical representation. I begin by giving an overview of the patterns of reduction for the diphthongs and plosives studied and discuss the role of place of articulation of the target allophones and their position in the phonological syllable, where this is possible. Next, I discuss each of the phonetic factors that emerged as significant in the modeling of reduction and deletion. These analyses are based simply on the distributions of variants for each variable, and the factors are analyzed separately. As has been shown in the modeling of the variables, not all factors are significant for every variable when all factors are taken into account simultaneously, and some factors interact with each other. I will therefore refer back to the multivariate analyses in the interpretation of the results for each individual factor.

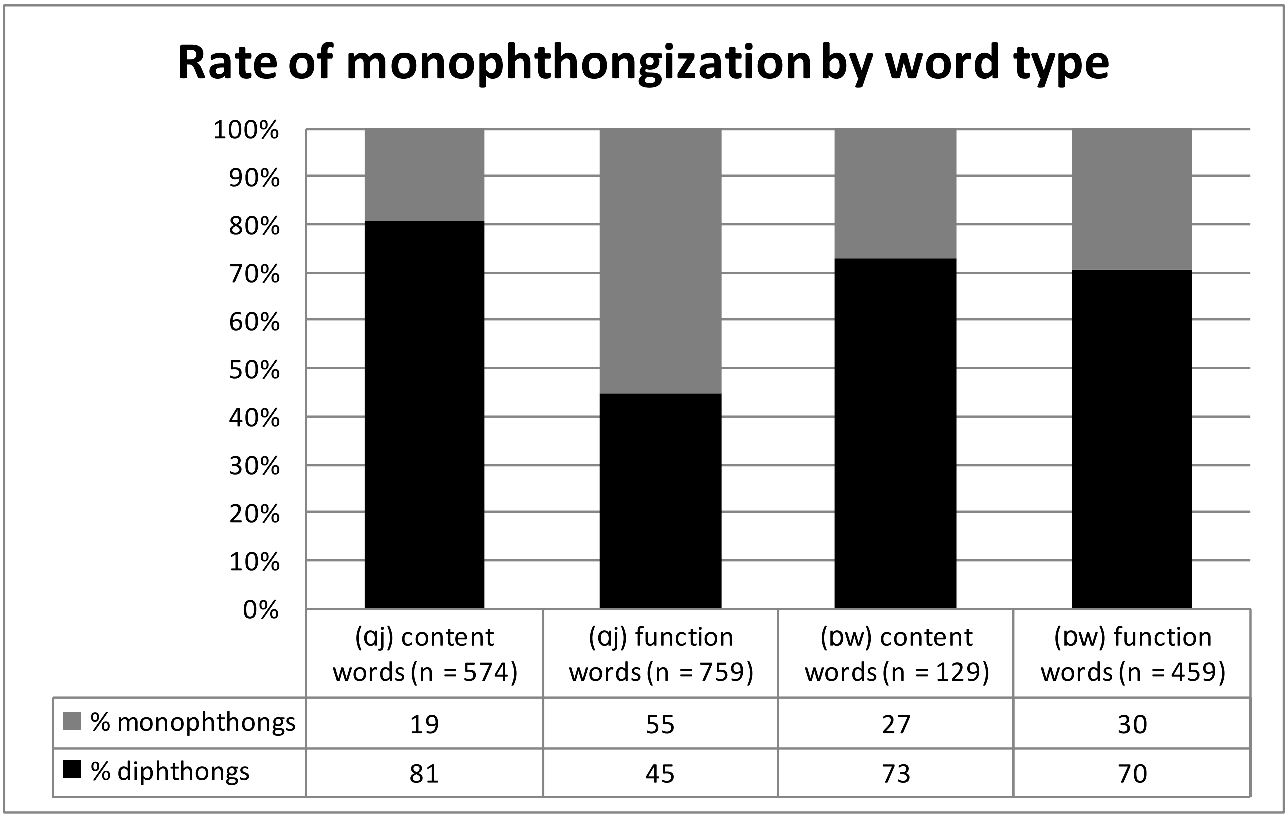
7.1 General patterns of reduction

7.1.1 Monophthongization

The rates of deletion for the two diphthongs studied, [ɑj] and [ɒw] , are shown in the graph below.



The datasets used here are the same as those described in the statistical modelling of the two variables. The difference is statistically significant (p < 0.001), with the rate of monophthongization being higher for (ɑj) than for (ɒw). This is undoubtedly due to the high proportion of tokens of (ɑj) occurring in the frequently monophthongized word 'jeg' [jɑj] *I*. Inspection of the full datasets for the two diphthongs suggests that function words in particular show a large span of variability: the two most frequently monophthongized words in each set, 'jeg' for (ɑj) and 'og' for (ɒw), appear in 20 and 15 different forms respectively[16]. This increased variability together with the characteristics that function words tend to be unstressed in running speech and are generally much more frequent than content words suggests that function words containing (ɑj) or (ɒw) may skew the probability for monophthongization. Therefore I have analyzed the distribution of monophthongs and diphthongs in content and function words separately, as shown in the graph below. Function words are restricted to: pronouns, conjunctions, interjections (like 'hej' [hɑj] *hi* and 'nej' [nɑjˀ] *no*) and true adverbs like 'derovre' [ˈdeɐ̯ˀˈɒwʁɐ] *over there*.



[16] This count is restricted to tokens where the word was clearly segmented from neighboring words. Stressed and unstressed tokens are treated as different forms.

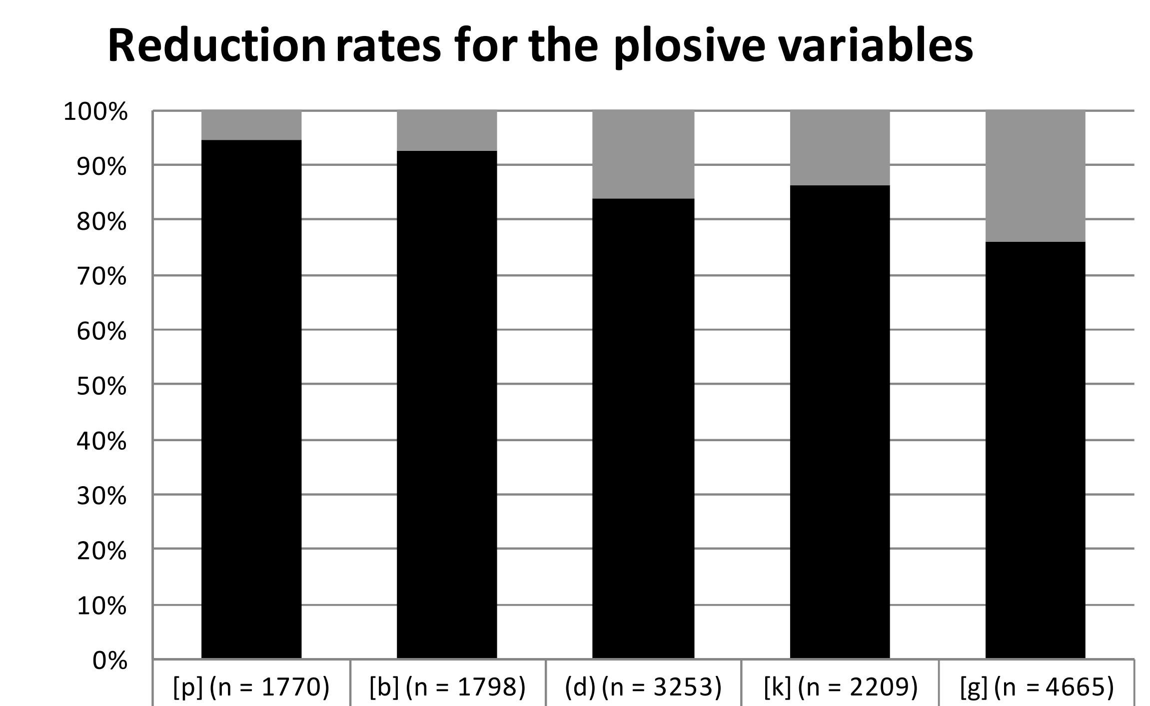
The graph shows an effect of word type for (ɑj) but not for (ɒw) – the difference for the two word classes is significant for (ɑj) ($p < 0.001$), but not for (ɒw) ($p = 0.5$). The multitude of monophthongs in function words with (ɑj) is accounted for by the great tendency for monophthongization of the word 'jeg' [jɑj] *I*. For (ɒw) no such effect is found, indicating that the lexeme [ɒw] for the word 'og' is not as inclined to monophthongization as the more frequent lexeme [ʌw] for 'og', which is predominantly monophthongized to [ʌ] (see chapter 6). For the words 'jeg' and 'og', then, the normal pronunciation in running speech can therefore be said not to include falling diphthongs at all, and it is possible that the monophthongization of these words are in some sense different from the monophthongization processes in content words like 'lejr' [ˈlɑjˀɐ] *camp* and 'skov' [sg̊ɒwˀ] *forest*. I return to this discussion in the chapter on phonetic reduction and lexical representation.

7.1.2 WEAKENING OF PLOSIVES – DIFFERENCES BETWEEN THREE PLACES OF ARTICULATION

The 5 plosives studied in the DanPASS dialogues are representative of the three places of articulation that are contrastive in Danish plosives: labial: (p) and (b); alveolar: word final (d); and velar: (k) and (g). The variables do not represent all tokens of labial, alveolar and velar plosives, since the variables (b), (d) and (g) have only been studied in syllable final position. In effect, the detailed studies of the DanPASS dialogues reveal the reduction processes of the morphophonemes |p t k|. Recall from the discussion of the variables (b) and (g), that analysis of the phonemes /b/ and /g/ in initial position in the phonological syllable has also been carried out, but showed no sign of effects which were strong enough to emerge as statistically significant. This means that it is possible to examine the tendency for reduction of all tokens of expected [b̥] and [g̊] in distinct pronunciation while also investigating the relationship between rates of reduction and representation at three levels of abstraction: distinct phones, phonemes and morphophonemes. I begin at the level of distinct phones by examining the rates of reduction with respect to the three contrastive places of articulation.

The proportion of reduced tokens is displayed in the graph below. The graph shows the proportion of variants that are realized as the target plosive in black and the proportion of reduced realizations in grey. Each column represents a distinct phone. For the labial and velar plosives, the aspirated and unaspirated phones are given in pairs, with the leftmost column in a pair representing tokens of the aspirated phone, i.e. the variables (p) or (k), and the rightmost column representing tokens of the unaspirated phone, i.e. the variables (b) and (g) together with the tokens of syllable initial /b/ and /g/.

For the alveolar plosive, only tokens of the variable (d) are included. As in the analyses of each variable, the variants have been divided into two major groups: tokens realized as the target allophone according to distinct pronunciation, here termed "distinct", and tokens realized as phonetically reduced relative to the target allophone in distinct pronunciation, here termed "weakened". Included in the set of standard variants are aspirated and affricated realizations of the word final variable (d) and the syllable final variables (b) and (g), i.e. realizations of (b), (d) and (g) as [pʰ tˢ kʰ] respectively[17]. These allophones are not strictly speaking target realizations, but rather constitute strengthened variants of the underlying phonemes. But since the object of the present section is to examine the tendencies for reduction, strengthened variants have been included in the "distinct" category. Similarly, the category "weakened" includes both reductions and deletions.



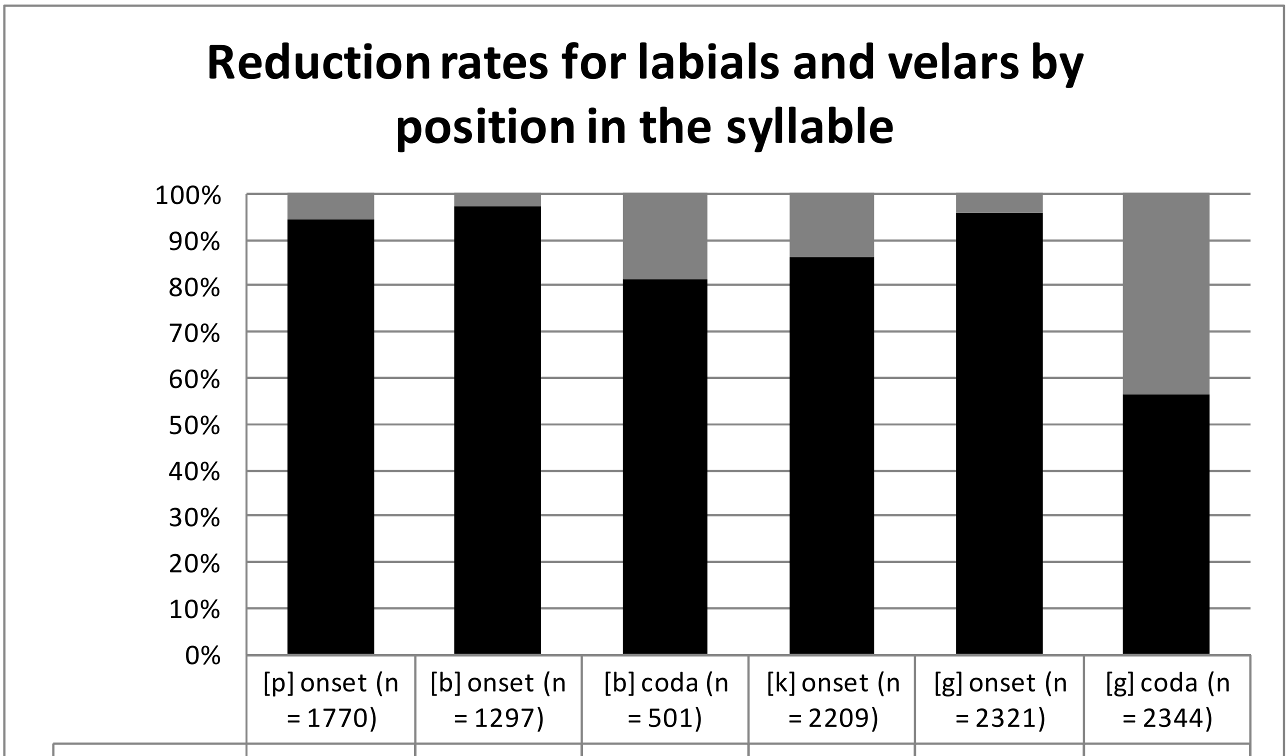
	[p] (n = 1770)	[b] (n = 1798)	(d) (n = 3253)	[k] (n = 2209)	[g] (n = 4665)
% weakened	6	7	16	14	24
% distinct	94	93	84	86	76

The graph shows that rate of reduction appears to increase the further back into the vocal tract the target place of articulation is located, suggesting a hierarchy of reduction rate based on place of

[17] The aspirated variant of (b) occurs exclusively in word final position, whereas the aspirated variant of (g), although it occurs predominantly in word final position, is also attested in a few tokens word internally, namely in the words 'virkelig' [ˈviɐ̯ɡ̊əli] *really* and 'virkeligheden' [ˈviɐ̯ɡ̊əliheːˀðen] *reality* which may be realized as [ˈviɐ̯kʰəli] and [ˈviɐ̯kʰeliheːˀðən] respectively.

articulation. Chi square tests reveal that the differences are significant (velar-alveolar $p < 0.0001$, velar-labial $p < 0.0001$, alveolar-labial $p = < 0.0001$).

The graph also shows that the tendency for reduction is higher for unaspirated [b̥] and [g̊] than for aspirated [pʰ] and [kʰ], however the difference is quite small for the labials with a difference of only one percentage point, whereas for the velars there is a difference of 10 percentage points. This might be expected given the proposed universal hierarchy of strength of consonants based on manner of articulation (cf. Vennemann (1988)), with the addition, as observed in Lavoie (2001, p. 44), that aspiration strengthens plosives, leading to a ranking of aspirated stops as stronger than unaspirated stops and hence less prone to reduction. Recall, however, that the proportions given for [b̥] and [g̊] masks the difference between position in the phonological syllable, since all tokens of distinct [b̥] and [g̊] are represented here. To study possible generalizations over the domain of the phonological syllable, the sets have been split up according to onset and coda position for the variable, and the distributions given in the graph below.



	[p] onset (n = 1770)	[b] onset (n = 1297)	[b] coda (n = 501)	[k] onset (n = 2209)	[g] onset (n = 2321)	[g] coda (n = 2344)
% weakened	6	3	19	14	4	44
% distinct	94	97	81	86	96	56

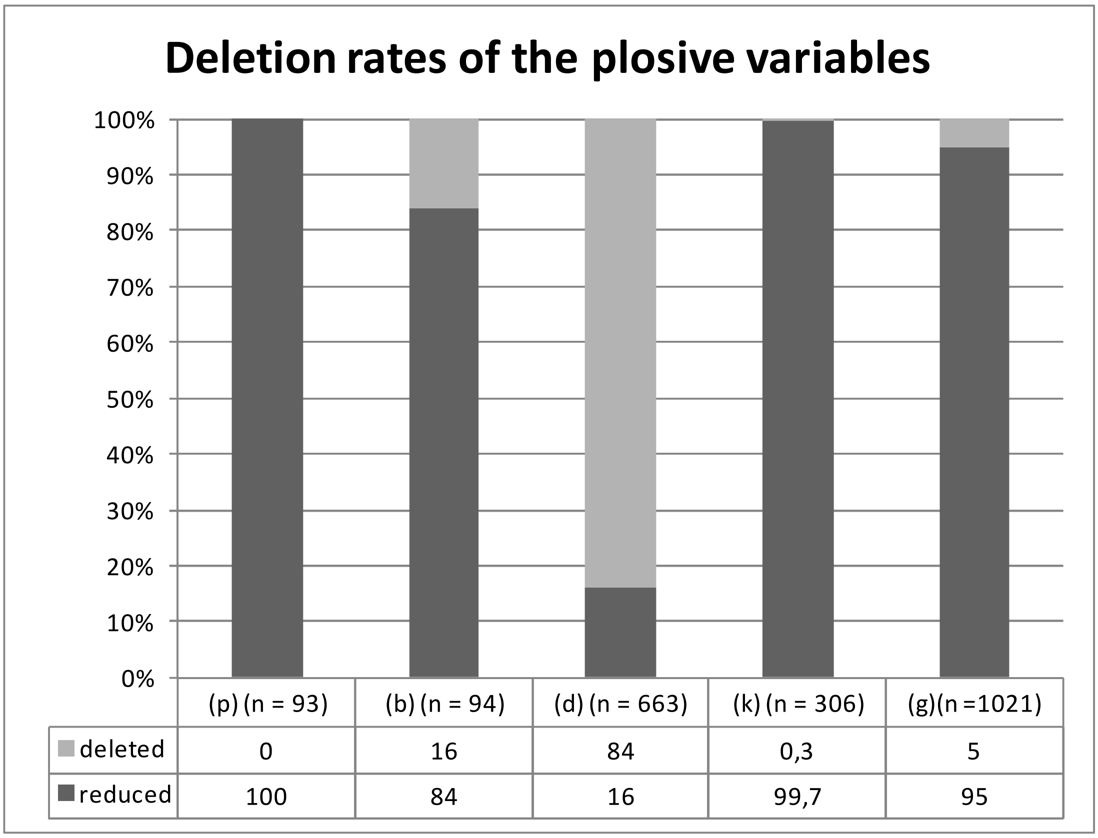
Comparing the tendencies for weakening in onsets suggests the exact opposite pattern of what was observed above: the rate of weakening is higher for the aspirated distinct phones than for the unaspirated distinct phones, although again the difference for the labials is quite small. For both pairs the difference is significant ($p < 0.001$ for both labials and velars). As is also apparent, the tendency for reduction in the coda of the phonological syllable is greater for both labials and velars (and this difference is also significant with $p < 0.001$ in all comparisons). In terms of distinct phones then, there is an overall difference between [pʰ], [kʰ] and [b̥], [g̊], but the difference is further modified by positional conditions, such that unaspirated distinct phones in onset position are least likely to be weakened, whereas aspirated phones in onset are somewhat more likely to be weakened, and phones in the coda are most likely to be weakened (but there is no contrast of aspiration here, since aspirated phones do not occur in the coda in distinct pronunciation). In terms of phonological representation, /p/ and /k/ are more likely to be weakened than /b/ and /g/ in onset position, with the further modification that there is a syllable position-effect for the two phonemes that are not defectively distributed: /b/ and /g/ are more likely to be weakened in the coda than in the onset. Abstracting further, to the level of morphophonemes, it can be seen that the pattern is opposite from the one observed at the phonological level, since the morphophonemes |b| and |g| in onset position are less likely to be weakened than the morphophonemes |p| and |k| in onset position. Furthermore, there is a syllable position-effect for |p| and |k|, which are more likely to be reduced in codas than in onsets. The same can be said for |b| and |g|, although this effect cannot be observed in the present dataset. This is because the processes of reduction for |b| and |g| in syllable final position have been phonologized: |b| may be realized as [w] in the coda, and |g| must be realized as either [w] or [j] in the coda in accordance with the preceding vowel: [w] after back vowels, [j] after front vowels and |l|. However, it is important to remember here that in the morphophonological analysis of Danish (cf. Basbøll (2005) and Grønnum (2005)), syllable final tokens of distinct [b̥] and [g̊] are identified with the morphophonemes |p| and |k|, and the difference between the phonological and morphophonological levels of representation described here is thus a consequence of the principles of analysis. There are exceptions to the identification of coda [b̥] with |p|, since [b̥] does not always alternate with [pʰ] in this position, e.g. the [b̥] in ['pʰlɔmb̥ə] *filling* remains a [b̥] in the derived word form [pʰlɔm'b̥eːˀɐ] *to seal.* Words with syllable final |b| that are realized as [b̥] in distinct pronunciation are quite rare, and none occur in the DanPASS Dialogues. The distinct phones [b̥] and [g̊] in coda position are almost never reduced to semi-vowels (with the exception of distinct [g̊] being reduced to [j] in 13 tokens), but their reduced variants bear some affinity to the realizations of |b| and |g|. From a morphophonological perspective, then, the processes of reduction of |p| and |k| in the coda of the phonological syllable observed for spontaneous speech are similar to the historical lenition processes that have affected |b| and |g|. For |k| in particular the reduction of tokens in coda is quite advanced.

The general results may be summarized as follows:

Hierarchy of reduction based on place of articulation: velars > alveolars > labials

Hierarchy of reduction based on position in the phonological syllable: coda > onset

Recall that the distributions above are based on weakening, i.e. no distinction was made between reduction and deletion. Below, I examine this distinction between types of weakening processes. Since the syllable initial variants of /b/ and /g/ are almost never deleted[18], they have been left out of the comparisons of tendencies for reduction and deletion. Reductions include the loss of aspiration in (p) and (k) and realizations as fricatives or approximants, which is possible for all 5 plosive variables, whereas deletion only includes tokens where the target allophone is not realized at all according to the phonetic transcriptions in the corpus. The questions of interest here is whether hierarchies of deletion may be established with reference to place of articulation and position in the phonological syllable, and if they match the ones found for overall reduction.



[18] Each phoneme is deleted exactly once in syllable initial position.

There is an effect of position in the syllable, since (b) is more likely to be deleted than (p), and (g) is more likely to be deleted than (k), and for both pairs the difference is significant according to a chi square-test ($p < 0.0001$ in both pairs of variables). The statistical significance is unsurprising since deletion of (p) never occurs, whereas 15 of the 94 tokens of (b) are deleted, and for the velars deletion occurs only once for the syllable initial variable (k), but 126 times out of the possible 1021 for the syllable final variable (g). Deletion of plosives in syllable initial position appears to be almost non-existent, leading to a similar hierarchy for deletion with respect to position in the syllable as the one that was observed for overall weakening: deletion is more likely in codas than in onsets. Recall again that there is necessarily a confound between aspiration and position of the morphophoneme in the phonological syllable for the data presented in the graph. However, given the equally low rate of deletion for syllable initial distinct [b̥] and [g̊], the effect of position in the syllable can be said to be clear at any level of representation.

However, the hierarchy for deletion based on place of articulation is different than it was for overall reduction, since - as the graph clearly shows - deletion of alveolar stops is far more likely than for labial and velar stops. The differences are of course statistically significant ($p < 0.0001$), both between alveolars and velars and between alveolars and labials, whereas the difference between labials and velars is not statistically significant ($p = 0.31$).

Taken together, the results for propensity for reduction and deletion of plosive variables in the DanPASS Dialogues show that the relation between place of articulation of the target allophone of the variable and different types of weakening is complex: reduction and deletion together is more likely for velars than for alvoelars and labials and more likely for alveolars than for labials, whereas deletion on its own is more likely for alveolars than for velars and labials.

7.1.3 THE EFFECT OF POSITION IN THE SYLLABLE

The difference between the two positions in the phonological syllable has long been acknowledged in work on neutralization, where contrasts that can be found in syllable initial position do not occur in syllable final position (cf. devoicing in German (Trubetzkoy (1962)). In fact, this is a possible interpretation of the differences in distribution between the phonemes /p t k/ and /b d g/ in standard Danish. Both [pʰ tˢ kʰ] and [b̥ d̥ g̊] occur in initial position in the phonological syllable, but only [b̥ d̥ g̊] occur in syllable final position. In other words, the contrast between aspirated[19] /p t k/ and unaspirated /b d g/ is neutralized in syllable final position (cf. Grønnum (2005)).

[19] Affricated in the case of /t/.

The effect is apparently mirrored in patterns of reduction, as has also been shown for spontaneous speech in American English, where the range of variation has been found to be greater in codas than in onsets (Greenberg (1999)) and where the probability for reduction and deletion has also been found to be greater in codas than in onsets (Raymond et al. (2006)). And the data analyzed here suggests that the phonological syllable may be a relevant domain for such processes. I.e. if we accept the level of abstraction that is constituted by the morphophonological representation and identify syllable final occurrences of the phonemes /b d g/ according to the principles of either Grønnum (2005) or Basbøll (2005), we can describe the reduction effect with reference to the phonological syllable. But what might account for this effect? Two suggestions will be discussed here, and suggestions for further, experimental, work given.

Articulatory studies, mainly within the framework of Articulatory Phonology (see Kochetov (2006) for a review) have found different patterns of gestural organization in onsets and codas. According to Browman & Goldstein (1995) consonant gestures in onset position tend to be synchronous, whereas consonant gestures in coda position are sequential. Evidence for the difference in gestural organization has, according to Kochetov (2006), been shown in Krakow (1989) where the tendency for vowels to nasalize before coda consonants in English but not before onset consonants is be explained by the difference in gestural organization, since lowering of the velum occurs before formation of the alveolar constriction in codas. This leads to nasalization of the preceding vowel and explains why it does not apply to vowels before onset nasals, since for them, constriction and lowering of the velum are synchronous, and therefore, constriction is completed before passage to the nasal cavity can affect the realization of the vowel. Furthermore, the magnitude of gestures, that is the degree of constriction, also differs in the two positions, such that the magnitude of gestures is greater in onsets than in codas, meaning simply that target, i.e. complete closure, is more often reached for plosives in onset than in codas. These magnitude effects have also been found for oral stops and arise again from the organization principles in codas: vowel articulation continues through the formation of the stop constriction leading to target undershoot. The patterns of reduction for stop consonants in contemporary Copenhagen Danish found here suggest that the magnitude effects vary with place of articulation.

Speakers apparently tend to make a greater effort to clearly articulate segments in onsets than in codas, and in this sense the constituents of the syllable may be interpreted as psychologically real units and relevant domains in the speech production process. Note that the patterns observed at the level of phonetic syllables as discussed for the variables (b) and (g) run counter to the pattern for phonological syllables, since reduction is more likely in the onset than in the coda of the phonetic syllable.

Furthermore, tokens in the onset of phonetic syllables are more likely to be reduced in word internal position than in word initial position. This difference could be related to stress. Word initial tokens are generally more likely to be in the onset of a stressed phonetic syllable than word internal tokens, since word internal plosives in all of the tokens in the datasets studied here form a phonetic syllable with an unstressed non-full vowel, either schwa or [ɐ] (which is phonologically /ər/, cf. Basbøll (2005)). This means that word internal plosives in the coda of the phonological syllable are also most often followed by an unstressed syllable. However, tokens in coda position, i.e. the variables (b) and (g), are not less likely to occur in stressed syllables than tokens in onset position, i.e. the variables (p) and (k): in fact, tokens of (b) and (g) are more likely to occur in the coda of stressed phonological syllables than in the coda of unstressed syllables, whereas tokens of (p) and (k) are more likely to occur in unstressed phonological syllables than are (b) and (g). For the bilabials this difference is particularly striking, since 92 % of tokens of (b) occur in (the coda of) a stressed syllable, whereas only 16 % of tokens of (p) occur in (the onset of) a stressed syllable. This large proportion of tokens of (p) in unstressed syllable onsets is due to the many tokens occurring in the word 'på' *on* which is normally realized as unstressed in running speech, i.e. [p^hɔ] with concomitant shortening of the vowel and loss of stød. This shows that the effect of position in the syllable is not just an effect of the stress of the phonological syllable in which the variable occurs. However, as will be discussed in the sections on the role of stress and position in the word below, there is a clear connection between reduction and (lack of) stress.

What might be the reason for this increased tendency of speakers to reach target articulation in onsets of the phonological syllable? Preference for clarity and contrast maintenance in one position over the other provides segmental cues for segmentation of the speech stream to listeners. If segments are less likely to reduce in onsets than in codas, and if they are less likely to reduce at the beginning of words than word internally, then articulations that reach target are likely to occur at the beginnings of syllables and words, thus giving the listener a cue to where in the speech stream new words begin, and where a lexical search ought best to be initiated in the speech recognition process. Such listener oriented motivations for syllable position-effects is further supported in the Cohort Model of speech perception (Marslen-Wilson (1984), Taft & Hambly (1986)). In the Cohort Model, words are recognized by reducing the number of possible word matches on-line as more and more phonemes are identified in the analysis of the incoming speech stream. When the assembled string matches only one word, recognition may be completed, since there is only one candidate left in the cohort of words that were activated at the beginning. Taft & Hambly (1986, p. 260), give the example of the English word 'crocodile', which can be recognized as soon as the (percept of) [d] is encountered, because no other words in the lexicon begin with [ˈkɹɒkəd], but it cannot happen earlier, because the string [ˈkɹɒkə] also partially matches the word 'crockery', i.e. the cohort still contains more than one candidate. So the [d] constitutes the Uniqueness Point for the word 'crocodile'. Clearly, the Cohort Model best describes the recognition of words spoken in isolation, but the notion of the Uniqueness Point may be used to interpret the greater likelihood of reduction in codas as a speaker strategy that capitalizes on assumptions of the needs of the listener. However, as can be seen from the analysis of the effect of position in the word below, this interpretation does not work so well for words in running speech. The probability for segment reduction is generally higher word internally than at word boundaries, and this is also true for the consonantal variables that occur in the coda of the phonological syllable. Since a word's Uniqueness Point is generally more likely to occur late in the word, one would assume that syllable final consonants like the variable (g), would be more likely to be reduced in word marginal, i.e. final, position than word internal position, since the communicative usefulness of maintaining the contrast at an earlier point in the word is higher than at a point where it is more likely that only one candidate remains in the cohort. Naturally, a thorough investigation of the influence on speech production of speaker assumptions about listener demands for clarity based on the lexical representation of words being produced would require an investigation in terms of the actual Uniqueness Points of each word form containing the variables. But the evidence found here does not suggest that this is a viable motivation for the effect of position in the phonological syllable. As will be discussed below, the effect of position in the phonological syllable can be derived from the factors stress and segmental environment, meaning that the generalizations that can be made with respect to structural levels of representation follow from the probabilistic tendencies from general phonetic features. Before discussing the effects of the phonetic factors, however, I want to briefly consider the motivation for the differences observed between the three places of articulation for the plosives.

7.1.4 THE EFFECT OF PLACE OF ARTICULATION

Why might velar plosives be more likely to be reduced than labials and alveolars? A possible motivation can be found in the different potentials for confusability with other phonemes and hence potentially other words for the three different places of articulation in the phonology of Danish. When a velar plosive is realized as a fricative, there are no other contrastive segments that it is likely to be confused with, since Danish does not have a contrast involving velar fricatives. If a labial plosive is reduced to a fricative, it could potentially be confused with the contrastive labiodental [f] and similarly, a reduced alveolar plosive could potentially be confused with contrastive [s] or even [ð] . That is the potential for confusion may influence speaker behavior such that reduction is avoided when the potential for confusability is high. Further evidence to test this hypothesis could be found by studying the tendency for reduction in words that form minimal pairs in the manner contrast. If speakers are sensitive to the potential for confusion, they should be less likely to reduce plosives to fricatives in words like 'kappe' [ˈkʰɑb̥ə] *cloak*, where reduction of the [b̥] to a fricative would make it more similar to the word 'kaffe' [ˈkʰɑfə] *coffee*, but more likely to do so in a word like 'lappe' [ˈlɑb̥ə] *patch (vb.)*, since no word *'laffe' [ˈlɑfə] exists, that the reduced word form might be confused with. It is of course likely that contextual information provided by the utterance in which the word occurs will override the effect, but this must be left for future investigations to resolve.

7.2 GENERAL EFFECTS OF THE PHONETIC FACTORS

7.2.1 STRESS

Studies of phonetic reduction have generally found that reduction is more likely to occur in unstressed than in stressed syllables (cf. Zue & Laferriere (1979), Raymond et al (2006)) and conversely that segments in stressed syllables are less variable than in unstressed syllables (Greenberg et al. (2002)). In the following, I summarize the effects of stress on the phonetic variables studied in the DanPASS Dialogues and discuss the role of stress in the speech production process.

7.2.2 DELETION AND STRESS

Monophthongization of both (ɑj) and (ɒw) and deletion of word final (d) show an association with stress: all three processes are more likely to occur in unstressed syllables than in stressed syllables. Recall that there is a difference in the types of processes modeled for these two sets of variables. Table 7.1 gives the average rate of deletion tokens of the variables (ɑj), (ɒw) and (d) in stressed and unstressed syllables. The fourth column gives the total number of tokens in which deletion occurred.

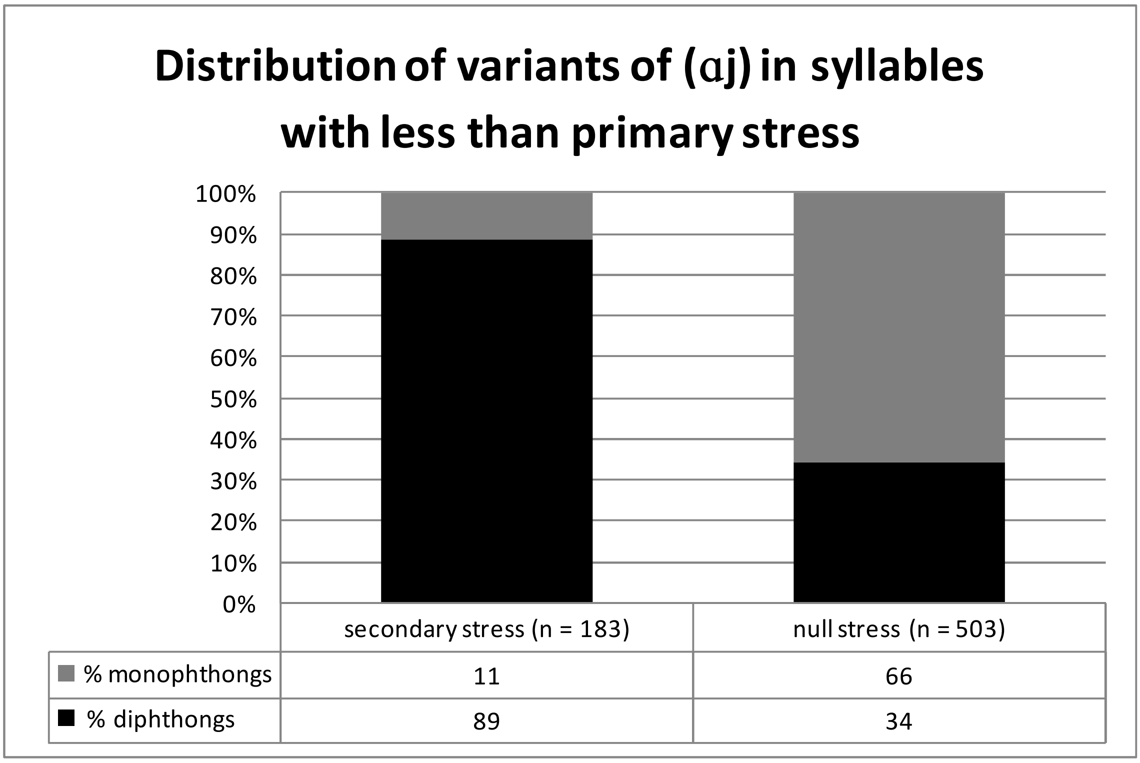
Table 7.1 – Rate of consonant deletion by stress

Variable	Stressed syllables	Unstressed syllables	Number of tokens deleted
(ɑj)	28 %	51 %	1333
(ɒw)	24 %	38 %	171
(d)	11 %	35 %	623

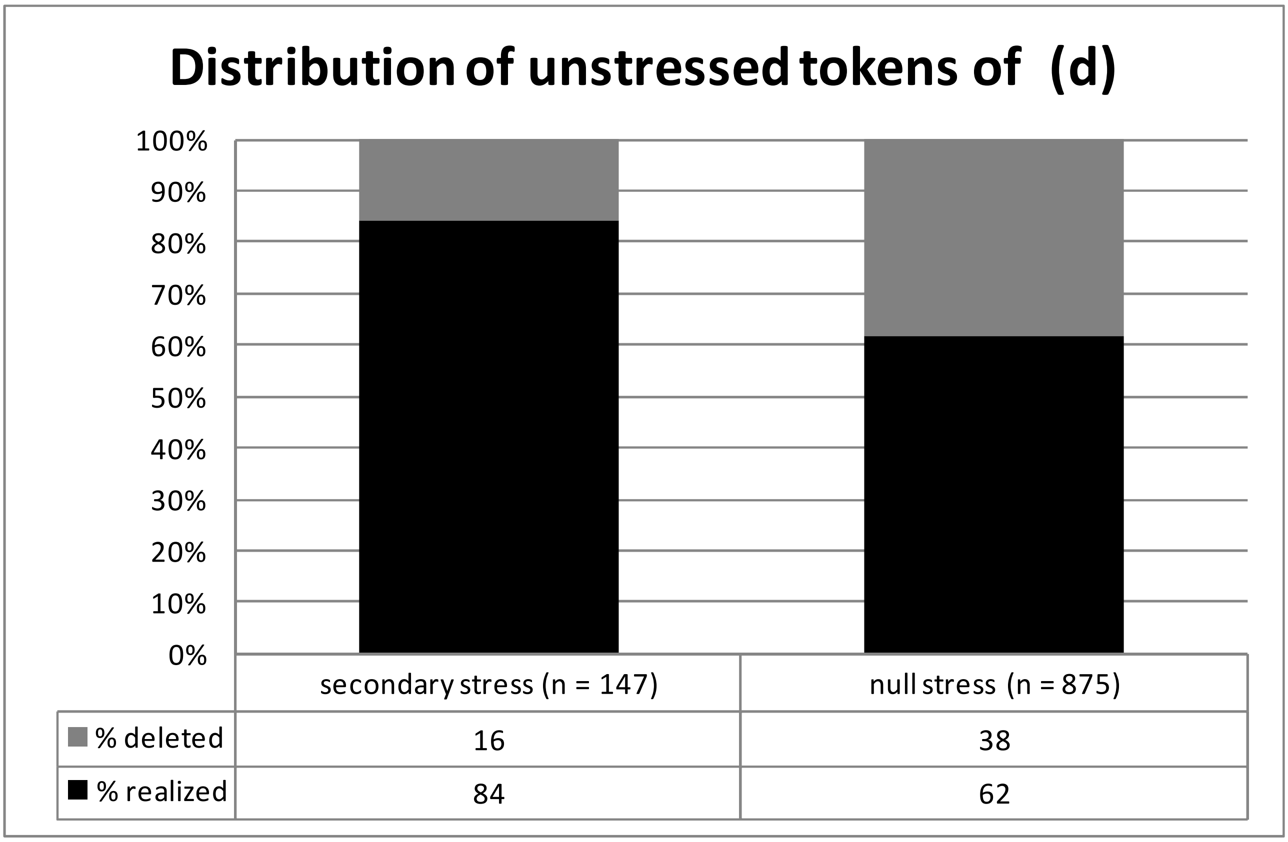
As the regression models have already indicated and as table 7.1 clearly shows, none of the three deletion processes show any absolute effects of stress: i.e. while monophthongization is more common in unstressed syllables, unstressed diphthongs are still attested in the corpus, i.e. lack of stress does not entail monophthongization indiscriminately. Conversely, (d) deletion is very rare in stressed syllables, but it does occur, so deletion of (d) is not necessarily blocked by stress. In order to better understand the connection between stress and segment deletion, the atypical tokens are explored further here, i.e. unstressed diphthongs and (d) deletion in stressed syllables.

Recall that stress here denotes perceived full stress, that is, stress noted by the transcribers and marked in the corpus. Stress was classified as a binary feature of syllables in the transcription, i.e. a syllable was either stressed or unstressed, but there were no intermediate degrees of stress (Grønnum (2009)). All tokens classified as stressed thus carry primary or full stress and are (presumably) all tonally marked for stress. However, some of the tokens which have not been marked for primary stress occur in syllables with secondary stress, whereas others occur in syllables with null stress. According to the principles of Basbøll (2005) unstressed (ɑj) in 'byggelegeplads' [ˈb̥yg̊əlɑjəpʰlas] *playground* has secondary stress, because it is the head of the second constituent of a compound, i.e. 'bygge' + 'legeplads', whereas (ɑj) in the pronoun 'mig' [mɑj] *me* under normal circumstances has null stress in running speech. For (ɒw), there is no difference between the rate of monophthongization in syllables with null stress and in syllables with secondary stress: in both conditions (ɒw) is monophthongized in approximately 40 % of tokens, and the difference is, of course, not statistically significant ($p = 0.76$). For (ɒw), then, there is no difference between the two levels of non-primary stress.

Preliminary inspection of the unstressed diphthongal tokens of (ɑj) shows that a great number of them occur in words where the diphthongs are in a position specified for secondary stress, e.g. in compounds like 'byggelegeplads, flodleje, halvvejs' [ˈb̥yg̊əˌlɑjəpʰlas, ˈfloðˌlɑjə, ˈhalˌvɑjˀs] *playground, riverbed, half way,* and (ɑj) also occurs in a series of pronouns – 'jeg, mig, dig, sig' – which are all normally unstressed in running speech. The proportions of monophthongs and diphthongs with respect to secondary stress and null stress are shown in the graph below.



The distribution reveals that there is a difference between which kind of non-primary stress that invites monophthongization and which one preserves diphthongs. The majority of monophthongal unstressed tokens consists of function words with null stress, in particular the pronoun 'jeg' [jɑj] *I*, whereas the majority of diphthongal unstressed tokens consists of content words with reduced stress due to compounding, i.e. secondary stress in subordinate constituents of complex words, such as in 'indianerlejren, byggelegepladsen'. Thus, monophthongization of (ɑj) is less likely to occur in syllables with secondary stress than in syllables with null stress, but it is still more likely to occur in unstressed syllables overall. This indicates that syllables with secondary stress tend to have the segmental properties of fully stressed syllables, but not the tonal contour, whereas syllables with null stress are segmentally reduced. This is further supported by an analysis of (d) deletion in unstressed tokens. Here too the tendency for deletion is weaker in syllables with secondary stress than in syllables with null stress, as the distribution shown in the graph below reveals. As with the semi-vowels of the diphthongs, (d) deletion is more likely to occur in completely unstressed syllables, than in syllables with secondary stress (and the difference is statistically significant according to a chi square test, $p < 0.001$). For both (ɑj) and (d), then, there is evidence that syllables with secondary stress tend to be segmentally richer than syllables with null stress.



Recall that a perhaps more surprising aspect is that stress does not completely block deletion of (d). As the model developed in the chapter 6 for (d) deletion showed, there was an interaction of stress and segmental environment: deletion in stressed syllables was most likely to occur in inter-consonantal position, and indeed this is borne out by the distribution of deleted tokens in stressed syllables after a consonant with respect to the following segment as shown in table 7.2 below.

The table shows that 73 % of deleted tokens of (d) in stressed syllables occur between consonants. This indicates that the inhibitory effect of stress on deletion of (d) can be overruled by the tendency for consonant cluster simplification as also discussed in the interpretation of the model of (d) deletion in chapter 6. Recall that there was an effect of the morphological category of the (d): deletion was more likely if the (d) constituted an adjectival or adverbial suffix, than if it was part of a stem. Inspection of the stressed tokens of (d) in inter-consonantal position reveals that the majority of the deleted tokens in stressed syllables are either adjectival or adverbial suffixes (45 % and 27 % respectively). In other

Table 7.2 – rate of deletion of (d) in stressed syllables with respect to the following segment

Following segment	% deleted	n
Vowel	18 %	40
Consonant	73 %	165
Pause	9 %	21

words, both segmental and morphological aspects conspire to induce deletion of (d) in the stressed tokens.

7.2.3 REDUCTION AND STRESS

Table 7.3 – Rate of consonant reduction by stress – phonological syllabification			
Variable	Stressed syllables	Unstressed syllables	Number of tokens
(p)	2%	6%	98
(b)	18%	26%	94
(k)	7%	25%	306
(g)	50%	58%	1022

When phonological principles of syllabification are used for classifying the occurrence of a token with respect to stress, the reduction rates for stressed and unstressed syllables are as given in table 7.3 above for the remaining plosive variables. Under this classification, only (g) showed an effect of stress in the regression models and this was only for word final tokens of (g), where reduction was more likely in unstressed syllables than in stressed syllable. Evaluating the effect of stress on the rate of reduction in stressed and unstressed syllables, on the basis of the distributions shown in table 7.3 above, by means of chi-square tests supports the lack of a difference for (b) in stressed and unstressed position (with $p = 0.14$), but for (p), (k) and (g) the differences are statistically significant. However, this significance for (p) and (k) only arises due to the fact that in this comparison, none of the other factors are controlled for. As the models of the variables showed, several other factors emerged, and apparently overrule the effect of stress suggested by the current treatment of stress on its own. Thus, for (p), reduction was more likely in word initial position than in word internal position and more likely after vowels than after consonants. The great majority of reduced tokens of (p) in unstressed syllables occur in word initial position, whereas there is no clear difference between occurrence relative to preceding vowels for the subset of unstressed tokens. For (k) both word internal position and occurrence after a vowel emerged as environments that favor reduction, and the great majority of reduced tokens of (k) in unstressed syllables occur in word internal position and after a vowel.

Table 7.4 – Rate of consonant reduction by stress – phonetic syllabification			
Variable	Stressed syllables	Unstressed syllables	Number of tokens
(b)	5%	27%	94
(g)	18%	53%	1022

The effect of stress was enhanced for both (b) and (g) when tokens were classified according to principles of phonetic syllabification and emerged as significant in the revised regression models, and unsurprisingly is also significant on its own. The distribution with respect to stress under the principles of phonetic syllabification is repeated in table 7.4. As already discussed, reduction of plosives does not show a clear pattern with respect to positions in the phonetic syllable: plosives are more likely to be reduced in word *internal* onsets in phonetic syllables than in word *initial* onsets. This can be seen when the data for /b/ are also included (recall from chapter 6, that none of the factors included in the present investigation could account for the reduction of /b/), i.e. tokens in the onset of the phonological syllable and hence also in the onset of the phonetic syllable, where a [b̥] is also expected in distinct pronunciation. Further restricting the examination to tokens in onset of unstressed syllables, we find 181 tokens of /b/ in word initial onsets in unstressed syllables. Here, the distinct [b̥] is deleted in 1 token, 'bliver' [b̥liwɐ] *gets (aux.vb.),* and reduced to [β] in 10 tokens, in the words 'bevæger, begynder, bliver, bananpalmen' [βeˈvɛːˀjɐ, βeˈgønˀɐ, βliwɐ βaˈnæːˀnpʰalmən] *moves, begins, becomes (aux.vb.) banana palm tree*, i.e. a reduction rate of 6 %. There are 531 tokens of word internal /b/ in unstressed syllables, 21 of which are reduced to [β]. Adding these to the set of 243 tokens of word internal (b) in unstressed onsets of which 72 are reduced, we get a reduction rate of 12 % for word internal distinct [b̥] in the onset of phonetic syllables, and this difference between position in the word is significant (p = 0.02). In other words, [b̥] is less likely to be reduced in onsets in word initial position than in onsets in word internal position, and this suggests that it is not the position in the phonetic syllable *per se* that is the important factor. However, the effect of stress is enhanced when the principles for phonetic syllabification are followed. In order to investigate whether position in the phonetic syllable or the stress of the phonetic syllable containing the (b) is the better predictor, a CART analysis was conducted with the two factors. In such an analysis, the stress of the syllable containing (b) emerges as the best predictor for reduction of (b), regardless of whether the token occurs in onset or coda of the phonetic syllable. It is of course impossible to test this prediction on the dataset for /b/, since all tokens are in onset position, but the fact that 32 of the 38 tokens of /b/ that are reduced occur in unstressed syllables supports the analysis of stress being the best predictor. However, this requires that the classification of stress for the token is based on the stress of the phonetic syllable. If we instead use the stress of the phonological syllable, we find that position in the phonetic syllable is a better predictor than stress. What might be the reason for this? I propose that the effect of phonetic syllabification may be reinterpreted as an effect of the segmental environment. In the dataset for (b), all word internal tokens which occur in the onset of a phonetic syllable are also intervocalic singletons (if we exclude the tokens occurring in the word 'simpelthen' [ˈsemˀb̥əld̥ˈhɛnˀ] *simply* in which the entire second syllable is often deleted as described in chapter 6).

Thus, the prediction that (b) in the onset of phonetic syllables is more likely to be reduced than (b) in the coda corresponds to the prediction that intervocalic (b) is more likely to be reduced than a (b) followed by a consonant. In word final position, (b) is only reduced in 20 out of 220 tokens, i.e. in 9% of tokens. 18 of these word final tokens occur in stressed syllables, and here pre-pausal position is the context in which reduction is most likely, occurring in 20 % of pre-pausal tokens, but there is no statistically significant difference for word final tokens followed by consonants and vowels – reduction occurs in 5 % and 9 % of cases for the two contexts respectively. In other words, distinct [b̥] is more likely to be reduced if the phonological syllable in which it occurs has reduced or null stress. If the syllable is stressed, (b) may still be reduced if it occurs in intervocalic position or, for word final tokens, in pre-pausal position. In this sense, phonetic syllabification is a good predictor for the reduction of distinct [b̥] in so far as it reflects the segmental environment.

The effect of stress is also enhanced for (g) under the principles of phonetic syllabification. Again, we may ask whether it is the position in the phonetic syllable *per se* or the stress of the syllable (irrespective of whether the (g) is in onset or coda) that best predicts reduction. A CART analysis of the factors stress and phonetic syllabification shows position in the phonetic syllable to be the best predictor, even when the classification of stress is done with reference to the phonetic syllable, i.e. when tokens of word internal (g) occur in the onset of an unstressed syllable. As with (b), when we examine the set of tokens that occur in the onset of both the phonetic and phonological syllable, we find a difference between word internal and word initial tokens: /g/ is reduced in 81 of the 2098 tokens in word initial position, i.e. a reduction rate of 4 %, whereas 18 of the 223 word internal tokens, or 8 % are reduced, and this difference for the dataset on /g/ alone is significant (p = 0.003). Continuing to compare reduction rates for tokens in onset of the phonetic syllable in word initial and word internal position, adding the word internal tokens in onset position from the dataset for (g), in which 873 tokens out of 1532 are reduced, we get a reduction rate of 51 % for word internal tokens, which is, of course, significantly different from the rate of 4 % in word initial position (p < 0.0001). So far, stress has not been controlled for. 55 of the 81 tokens of /g/ in word initial position that are reduced occur in unstressed syllables (the majority of which occur in the word ‘går’ [g̊ɒːˀ] *go* in the phrase ‘så går du’ [ˈsʌ g̊ɒ d̥u] *then you* go, as mentioned in chapter 6), whereas none of the reduced tokens of /g/ in word internal position occur in unstressed syllables – 17 of the 18 tokens occur in the word ‘igen’ [iˈg̊ɛn] *again* with stress. However, all of the word internal tokens of (g) are in the onset of an unstressed syllable and have a reduction rate of 59 %, which is of course also significantly different from the reduction rate for word initial /g/ in the onset of unstressed syllables ($p < 0.001$). Again, there is a difference between the rate of reduction for word initial and word internal tokens of distinct [g̊] in the onset of the phonetic syllable even when stress is controlled for, with reduction being much more likely in word internal position. As with (b), the majority of word internal tokens of (g) in the onset of the phonetic syllable occur in intervocalic position: only 3 % occur after a consonant in the words 'banket, ganske, hvilken, tænke' [ˈbɑŋg̊əð ˈg̊ansg̊ə ˈvelg̊ən ˈtˢɛŋg̊ə] *knocked, quite, which, think (inf.)* and none of these are reduced. This indicates that the effect of position in the phonetic syllable for (g) should also be reinterpreted as an effect of the segmental environment. When the phonological syllable in which (g) occurs is unstressed, reduction is more likely. If the syllable is stressed, reduction may still occur provided the (g) is intervocalic. This is further supported by the finding that word internal stressed intervocalic /g/ as in 'igen' is also reduced, and tokens in stressed syllables in word initial position almost all occur in intervocalic position as well (there are only 2 exceptions out of the 26 occurrences of reduced /g/ in word initial stressed syllables). Taken together, plosives have a greater tendency for reduction in unstressed phonological syllables, but may also be reduced in stressed syllables if both adjacent segments are vowels. The effect is stronger word internally than at word edges, but this may be due to the fact that word initial syllables tend to be stressed, and word initial tokens are more often preceded by a consonant than word internal tokens are. The results are discussed further below in the section on the role of the position in the word.

7.2.4 THE EFFECT OF STRESS AND ITS ROLE IN LEXICAL REPRESENTATION

The overall result of the analysis of the effect of stress on processes of phonetic weakening is that reduced stress increases the likelihood for deletion, but has only a marginal influence on reduction of plosives. Closer examination of the relation between stress and deletion revealed that complete deletion is most likely to occur in syllables with null stress, whereas syllables with secondary stress are more likely to retain segments, thereby providing a gradient segmental cue to the degree of stress.

In a single entry model of lexical representation where all morphological processes including the formation of compounds are carried out during construction of the phonetic plan for utterances (like that of Levelt (1989)), the effect of reducing stress must be post-lexical in the sense that all lexemes are specified for stress. The increased probability for deletion in unstressed syllables can then be conceptualized as a gradient rule allowing for deletion in unstressed syllables. However, the gradience is apparently tempered by the level at which stress reduction takes place, since deletion is less likely in syllables with secondary stress than in syllables with null stress. The ascription of secondary stress takes place in the formation of compounds, i.e. at a point in the derivation of word forms prior to the formation of the phonetic plan for utterances. Loss of stress for syntactic or pragmatic reasons, on the other hand, cannot take place before words have been combined into utterances, and therefore the effect of null stress applies here. In a single entry model of the lexicon, then, the probability for deletion as an effect of stress is regulated by the level at which degree of stress assignment takes place: the later in the process the stress is reduced, the higher the likelihood of deletion. The tendency for stress to inhibit reduction of plosives can be affected by the segmental environment of the token: when a plosive occurs in a stressed syllable, the tendency for reduction is much greater if the plosive is intervocalic. This suggests that the effect of stress may be adjusted by the segmental environment in the implementation of the phonetic plan in a single entry model of lexical representation.

I will return to the discussion of the role of stress in segmental reduction and the implications for lexical representation in the final chapter, since, as will become apparent in the chapter on frequency effects, stress may interact with extra-linguistic factors, and this phenomenon is difficult to account for in a single entry model. Next, I present an analysis of the factor stød, which is closely related to stress.

7.2.5 Stød

A decreased tendency for reduction in syllables with stød was found for the variables (ɑj), (ɒw) and (b).

This is surprising for (b) since non-sonorant consonants do not affect the phonological weight of syllables (according to Basbøll (2005), p. 287).

However, this association between stød and monophthongization is gradient, that is, the presence of stød does not block monophthongization entirely. This can be seen for example in the case of (ɑj). While monophthongization is rare in syllables with stød, it does occur. For example, word forms like 'indianerlejr' [en̥diˈæːˀnɐlɑjˀɐ] *Indian camp* and 'indianerlejren' [en̥diˈæːˀnɐlɑjˀʁɐn] *the Indian camp* are predominantly realized with a full glide in (ɑj), but monophthongized tokens also occur. There are 59 tokens of the two word forms in all, in which the (ɑj) has stød, but in 14 of these tokens, the glide is lost, while the stød is preserved. All of these tokens of (ɑj) have reduced stress, a condition which facilitates monophthongization according to the model for (ɑj). In fully stressed tokens with stød, monophthongization is also rare, but again does occur, most commonly in the frequent interjection 'nej' [nɑjˀ] *no*, but also sporadically in less frequent word forms like 'vej, vejen, lejr' [vɑjˀ ˈvɑjˀən ˈlɑjˀɐ] *road, the road, camp*. Therefore, stød cannot be said to automatically block monophthongization, since monophthongization *does* occur in syllables where the stød is maintained. Note that when monophthongization occurs in syllables with stød, it is accompanied by compensatory lengthening of the vowel [ɑ], meaning that sufficient voiced material remains in the manifestation of the syllable, so that stød can be realized.

The apparent effect of stød pertains to syllables with more than null stress only, since a syllable must be both segmentally heavy, i.e. have a long vowel or a short vowel and a sonorant in the syllable rhyme, and have more than null stress in order to have stød[20] (Basbøll (2005) where what I am calling segmental weight is analyzed in terms of morae). In the present investigation, only syllables with primary stress have been encoded as stressed, because only primary stress has been transcribed in the corpus. This means that what appears to be an effect of stød for unstressed tokens may instead reflect a difference between the propensity for monophthongization in syllables with seconday stress, i.e. reduced lexical stress in subordinate constituents of compounds, and syllables with null stress. In other words, the apparent effect of stød may be seen as an addition to the effect of stress. Since stressed tokens are also less likely to be reduced than unstressed tokens, the stød can be interpreted as a further inhibitory factor on the processes of reduction. Continuing the investigation of (ɑj), I will examine the reduction rates for tokens of the variable (ɑj) in syllables with primary stress with and without stød. Next, I will examine the reduction rates for syllables with secondary stress, again to see whether syllables with stød and without stød show a difference in rate of monophthongization. If there is a difference between syllables with stød and stødless syllables when the degree of stress is held constant, this will indicate that the segmental structure required for stød to occur, cf. above, tends to be maintained by speakers in the phonetic realization of the syllable in running speech.

There is a total of 642 tokens of (ɑj) with primary stress, 337 of which have no stød, and for these stødless tokens, monophthongization occurs in 47 %, primarily in stressed tokens of the words 'jeg, meget, nej' [jɑj ˈmɑjəð nɑj] *I, very, no*. Note that 'nej' normally has stød when it is stressed, but that it apparently may be realized as stødless [ˈnɑj] or [ˈnɑ]. For the 305 tokens with stød, only 6 % are monophthongized, most of them in the word 'nej'. This difference is, of course, statistically significant,

[20] The derivational suffixes '-ig, -lig, -ing, -ning' [i li eŋ neŋ] which are canonically unstressed may have stød when they occur in inflected word forms, e.g. 'tegning' [ˈtˢɑjneŋ] *drawing* without stød, but 'tegningen' [ˈtˢɑjneŋˀən] *the drawing* with stød. However, they do not pertain to the variables studied here.

and explains why the effect emerged in the model for (ɑj). The high rate of monophthongization in stødless stressed tokens of (ɑj) is mainly due to a few highly frequent words.

There are 179 tokens with secondary stress in the dataset for (ɑj), 69 of which, like 'byggelegepladsen' and 'flodleje' [ˈb̥yg̊əˌlɑjəpʰlasən ˈfloðˌlɑjə] *the playground, riverbed*, do not have stød, and 7 % of them are monophthongized. For the 110 tokens with stød, in words like 'indhegning, indianerlejren, halvvejs' [ˈenˌhɑjˀneŋ endiˈæːˀnɐˌlɑjˀɐ ˈhalˌvɑjˀs] *fence, the Indian camp, halfway*, monophthongization occurs in 15 %, but the difference is not statistically significant ($p = 0.14$). These rates for syllables with secondary stress can be contrasted with the rate of monophthongization for syllables with null stress (and therefore no stød), such as in 'jeg, drejer, tegne', where the rate of monophthongization is 59 %. In other words, the reduction rate for syllables with reduced lexical stress is closer to the reduction rate for syllables with primary stress and stød than the reduction rate for syllables with null stress. Since there is no (statistically significant) difference between the reduction rate for syllables with secondary stress with and without stød, the effect may be reinterpreted as an effect of the degree of stress, with syllables with secondary stress having a tendency to be as segmentally rich as syllables with primary stress even in running speech. However, there was a rather large difference between syllables with primary stress and stød and syllables with primary stress without stød. This is, as mentioned, in large part due to a few highly frequent function words. When the set of tokens with primary stress is restricted to content words, the rate of monophthongization for stødless syllables with primary stress drops drastically to 9 %, which is as low as the rate for syllables with secondary stress. When the set of stødless tokens with null stress is similarly restricted to content words, the rate of monophthongization only drops to 36 %. In other words, except in highly frequent items, the apparent effect of stød can be seen as a corollary of the degree of stress, an association that is not captured in the modeling of stress as a binary feature. Looking at all items, there is an association between stød and monophthongization which cannot be ascribed to differences between degrees of stress. This does not necessarily mean that the stød is the underlying cause for the increased tendency to preserve segments in syllables with stød. As was shown above, stød may be preserved in monophthongized tokens, in which case the [ɑ] is lengthened and thus the syllable retains sufficient voiced material in order for the stød to be realized. Recall from the description of stødless syllables with primary stress that the word 'nej' is sometimes realized without stød. There are 59 stødless tokens of 'nej' with primary stress, 19 of which are monophthongized, and 40 of which are diphthongal. If stød were the underlying cause for preservation of the glide, one would have expected all of these tokens of stressed, stødless 'nej' to be monophthongized, but clearly they are not. In other words, it is not because syllables have stød, that the glide is preserved, but when the glide is preserved, the stød is also likely to be realized. When a token is monophthongized, the stød is only realized if there is compensatory lengthening of the vowel. When the vowel is not lengthened, stød is also lost. Rather than stød being an inhibitory effect on monophthongization, the correlation between diphthongal tokens and stød is more likely to be a reflection of the fact that stød is more likely to be retained when monophthongization does not apply.

7.2.6 POSITION IN THE WORD

An effect of position in the word was seen for the variables (ɑj) (p) (g) and (k). Obviously, it is not possible to test for the influence of position in the word on the variable (d) since it was restricted to word final tokens of the morphophoneme |t|, and hence there are no word internal tokens to compare with. For the variables (ɒw) and (b), no effect was found for position in the word. The purpose of including position in the word as a factor was to study to what extent reduction processes are sensitive to word boundaries, with the hypothesis being that reduction would be more likely in word internal position, hence allowing for the interpretation that, if the hypothesis is confirmed, word boundaries can be said to have a tendency to inhibit reduction of segments. The fact that there is no effect for the two variables (ɒw) and (b) indicates that this effect is not general in the sense that apparently some segments are equally likely to be reduced irrespective of the presence of an adjacent word boundary, i.e. word boundaries do not necessarily inhibit reduction. However, for the variables where a statistically significant effect does emerge, it is of course necessary to examine whether the factor has the hypothesized effect. In the table below, the rates of reduction are given. Note that for (ɑj), tokens of the word 'jeg' are excluded, because they would skew the reduction rate for word marginal tokens.

Table 7.5 – rate of reduction by position in the word

Variable	% reduced word internally	% reduced word marginally	Number of tokens
Syllable final variables			
(ɑj)	35	22	114 (544)
(g)	53	30	1022 (2344)
Syllable initial variables			
(p)	2	7	130 (2324)
(k)	20	7	306 (2173)

The distribution is not the same for all 4 variables: for (ɑj), (g) and (k) there is a larger proportion of reduced segments word internally than word marginally, which confirms the hypothesis that an adjacent word boundary tends to inhibit segmental reduction. The effect is somewhat surprising for (k), since all word internal tokens occur before a full vowel, where reduction is not expected. The effect cannot be explained simply by referring to the stress of the syllable containing (k), since reduction is also very likely in the word 'okay' [ˈɒwˈkʰɛj] *okay* even when the word has stress on the second syllable. However, (k) is primarily reduced in stressed syllables in word internal position when it is preceded by a vowel, that is, when the initial syllable in 'okay' is monophthongized, (k) is reduced in 19 % of tokens, but when the first syllable is not monophthongized, reduction only occurs in 5 % of tokens. Reduced word initial tokens of (k) predominantly occur in unstressed tokens of the words 'kan' and 'kommer' [kʰanˀ ˈkʰʌmˀɐ] *can (aux.) comes* or in a pre-tonic syllable in the word 'kolonihaver' [kʰoloˈnihæːwɐ] *allotments (gardens)* which may be realized as [xa xʌmɐ xoloˈnihæːwɐ] respectively. When reduction does occur in a stressed word initial syllable, the preceding word always ends in a vowel as in the sequences 'par kilometer' *couple of kilometers* and 'du kommer' *you come* and 'og klippen' *and the cliff* which are realized as [pɑ ˈxiɾʊmɛɾə], [dʊ ˈɣʌmɐ] and [ʌ ˈg̊lebm̩] respectively. This suggests that the effect of position in the word for (k) may be derived from the conditions of stress and segmental environment, just as was found for (b) and (g) in the analysis of stress and position in the phonetic syllable above: reduction of (k) is more likely in unstressed syllables than in stressed syllables, but reduction may occur in stressed syllables if the (k) is intervocalic. The difference in reduction rate for (g) with respect to position in the word may also be reinterpreted in terms of the segmental environment. Stress alone cannot account for the difference, since word internal and word final tokens of (g) are equally likely to occur in stressed syllables – in both position 77 % of tokens occur in stressed syllables - but word internal (g) is much more likely to occur in intervocalic environments, as in 'ligger' [ˈleg̊ɐ] *lie (pres.)*, than word final (g), as in 'nok' [nʌg̊] *probably*: 56 % of word internal (g) occur in intervocalic position, whereas only 8 % of tokens of word final (g) occur in intervocalic position. As was found in the modelling of (g) reduction, intervocalic position is the segmental environment that favors reduction of (g). In this sense, the effect of position in the word may be derived from the fact that there is a difference in the typical segmental environment for word internal and word final tokens of (g): word internal tokens more often occur in a segmental context which invites reduction. Note, however, that intervocalic position cannot account for all reduced tokens of word final (g), since the reduction rate is much higher, namely 30 %, than the proportion of tokens occurring in intervocalic position. In other words, intervocalic position is not a necessary condition for reduction of (g). I return to the role of the segmental environment below.

For (p) the distribution is the opposite: the proportion of reduced tokens is greater in word marginal position than in word internal position. The explanation can be sought in the fact that (p) is more likely to occur in word initial position than word internally in the corpus, but note also that word initial tokens tend more often to occur after a vowel than word internal tokens do: 44 % of word initial tokens are post-vocalic, whereas only 13 % of word internal tokens are post-vocalic, and preceding vowels were seen to increase the likelihood of reduction of (p) (see section 6.9), which would explain the distribution shown in table 7.5 above. To further explore the effect, note that three words alone constitute the majority of words in which word initial tokens of (p) occur in post-vocalic position: 'på' [pʰɔːˀ] *on* 'passeret' [pʰaˈseːˀɐð] *passed* 'passerer' [pʰaˈseːˀɐ] *passes* constituting 46 %, 13 % and 11 % respectively of post-vocalic word initial tokens of (p). Note that these tokens also tend to occur in unstressed syllables: not only are tokens in 'passeret' and 'passererer' canonically unstressed, but reduced (p) in 'på' also always occurs in unstressed tokens of the word. The reduction rate is still very low for word initial (p), which explains why stress did not emerge as significant in the model. But again the factors stress and segmental environment conspire to make the effect of position in the word emerge as significant. Word internal tokens of (p) tend to occur after consonants, in words like 'østpå, udsigtspunkt, bananpalmen' [ˈøsd̥ˈpʰɔːˀ ˈuðseɡ̊d̥spʰɔŋˀd̥ b̥aˈnæːˀnpʰalmən] *eastward, vantage point, the banana palm tree*, and in fact in most of the 14 tokens of reduced word internal (p), the preceding segment is a consonant, in words like 'tidspunkt, totempælen' [ˈtˢiðspʰɔŋˀd̥ ˈtˢoːtˢɛmpʰɛːˀlən] *(point in) time, totem pole*, but the syllable tends to have reduced stress. The difference between rate of reduction for word internal and word initial (p) may thus be interpreted as a difference in the degree of stress on the syllable: word initial (p) is more likely to occur in syllables with null stress, whereas word internal tokens tend to occur in syllables with secondary stress.

The general results of position in the word show a tendency for word boundaries to inhibit reduction, supporting the hypothesis that the effect may serve as providing segmental cues for word edges and in this sense the word is a relevant domain in the speech production process. However, the effect can be seen to be derived from phonetic characteristics of the positions in the word, since word internal tokens more often occur in unstressed syllables and/or in intervocalic position. Furthermore, when word marginal tokens are reduced, they almost always occur either in unstressed syllables or in intervocalic contexts. The results suggest that the prosodic stress group may be the relevant domain for the reduction processes, rather than the word. The prosodic stress group is initiated by a stressed vowel and has a characteristic and qualitatively invariant contour with a rise to the first post-tonic syllable and a fall through any subsequent unstressed syllables, as established on the basis of acoustic analyses of intonation in Danish (cf. Grønnum (1992), p. 19). In other words, each vowel with primary stress initiates a new prosodic stress group. Thus, stress group boundaries do not necessarily coincide with word boundaries, since word initial unstressed syllables, as in the prefixes 'be-, for-' [b̥e, fʌ] in words like 'bevæge, forbi' [beˈvɛːˀjə fʌˈbiːˀ] *move, past*, constitute the final unstressed syllable in a stress group in spite of the fact that they constitute the initial syllables of the words. In connected speech, several word boundaries may occur within a stress group. For example, in the prosodic phrase 'så skal du igen mod syd' *then you have to go South again* as produced by speaker 002 in the DanPASS Dialogues has 6 words but only 3 stressed syllables viz: [ˈsʌ sg̊al d̥u iˈg̊ɛn moð ˈsyð] (with the segments represented here by distinct phones). Since the stress group begins with the onset of the stressed vowel, consonants in the onset of a stressed syllable belong to the preceding stress group. That is, the first prosodic stress group in the utterance above consists of the segmental string [ʌsg̊ald̥uig̊], the second stress group contains [ɛnmoðs] and the third and final stress group consists of [yð].

Reinterpretation in terms of the prosodic stress group would suggest a (probabilistic) rule stating that consonants in absolute final position in the stress group are less likely to be reduced relative to consonants which occur internally in the stress group. This inhibitory effect of absolute final position in the prosodic stress group may be weakened when consonants occur in intervocalic position, as was shown for the variables (k), (b) and (g). Using absolute final position in the prosodic stress group allows for the possibility to capture the effect of both word initial and word final position. For word initial tokens, the reinterpretation amounts to the generalization that consonants in the onset of stressed phonological syllables are less likely to be reduced than consonants in the onset of unstressed phonological syllables, with word initial syllables being more likely to be stressed than word internal ones. However, using the absolute final position in the stress group may also account for the effect that word final tokens are less likely to be reduced than word internal tokens, if word final tokens are more likely to occur before a stressed syllable than word internal tokens. As an example, I will re-examine the distributional pattern for reduction of (g). As can be seen for (g) in table 7.5 above, word final tokens are less likely to be reduced than word internal tokens. For this to be an effect derived from the position in the prosodic stress group, word final (g) should be less likely to be reduced when the following word begins with a stressed syllable, since that would place the word final token at the end of the prosodic stress group. This is in fact the case. Word final (g) is equally likely to occur before stressed and unstressed syllables, i.e. 50 % occur before a stressed syllable and 50 % before an unstressed syllable. But word final (g) is only reduced in 17 % of cases when the following syllable is stressed, but in 44 % of cases when the following syllable is unstressed. A CART analysis of the subset of word final tokens of (g) where the stress of the following syllable is included as a factor actually reveals the stress of the following syllable to be a better predictor than the segmental environment, which was argued above to be the important factor behind the difference in reduction rate for word internal and word final tokens. This supports the interpretation of the prosodic stress group as the domain for reduction processes. However, this principle fails to account for the findings that variables which necessarily occur in the coda of the phonological syllable, like (ɑj), (ɒw), (b), (g) and (d) are less likely to be reduced in stressed syllables than in unstressed syllables (cf. tables 7.1 and 7.3 above). When these consonants occur in stressed syllables, they are not (necessarily) in absolute final position in the prosodic stress group which is initiated by the preceding vowel, yet they are more likely to resist reduction than when they occur in unstressed syllables, e.g. the glide in 'drejer' [ˈd̥ʁɑjɐ] *turn (pres.)* always occurs internally in the prosodic stress group irrespective of whether the first syllable of the word is stressed or unstressed, yet monophthongization is more likely when this syllable is unstressed than when it is stressed. Furthermore, making absolute final position in the prosodic stress group the factor which underlies word position effects also cannot capture the fact that reduction is less likely for some variables in syllables with secondary stress than in syllables with null stress, which was found for the variables (ɑj) and (p). This can be seen from the fact that monophthongization of (ɑj) is less likely in 'byggelegepladsen' [ˈb̥yg̊əˌlɑjəpʰlasən] than in unstressed 'drejer' [d̥ʁɑjɐ]. In both word forms, the glide of the (ɑj) occurs internally in the prosodic stress group, but the probability for monophthongization is higher for the syllable with null stress than for the syllable with secondary stress. Similarly, the (p) in 'udsigtspunkt' [ˈuðseg̊d̥sˌpʰɔŋˀd̥] as well as the (p) in 'passeret' [pʰaˈseːˀʁɐð] both occur internally in the prosodic stress group, but again reduction is more likely under null stress than under secondary stress. While the patterning with respect to position in the prosodic stress group in running speech may serve to support the cue for position in the word, it conflicts with the general effects of stress that were shown above. Instead, the results indicate that the important factor is the stress of the phonological syllable containing the consonant, regardless of whether the consonant occurs in the onset or the coda, although consonants in the coda are generally more likely to be reduced than consonants in the onset (cf. above).In addition, reduction in stressed syllables is generally more likely when the token occurs in intervocalic position.

7.2.8 Local articulation rate

Local articulation rate was measured as the number of realized syllables pr. second in the prosodic phrase (cf. chapter 5 "The Factors") in which a token of a variable occurred. The local articulation rate emerged as a significant factor in the modeling of all of the phonetic variables in the DanPASS Dialogues except (b), and in all cases reduction or deletion was more likely the higher the local articulation rate. In order to provide a comprehensive image of the role of articulation rate, I first give the mean articulation rate of reduced or deleted tokens for each variable.

Table 7.6 below shows that mean local articulation rate is generally higher in phrases with weakened variants than in phrases with standard, i.e. non-weakened, variants. The significance of the differences in the means was tested using Wilcoxon's rank test, since Shapiro-Wilk tests for normality of distributions showed that articulation rates for the prosodic phrases were not normally distributed. The significances are given for each variable in the fourth column of table 7.6 below. It is to be expected that the differences will be statistically significant for all variables except (b), which was the only variable for which local articulation rate did not emerge as a significant factor in the regression modeling.

Table 7.6 – Comparison of mean articulation rate in phrases containing weakened and standard realizations of the phonetic variables

Variable	Weakened variants	Standard variants	Significance (Wilcoxon)
(ɑj)	6.13	5.29	$p < 0.0001$
(ɒw)	6.04	5.63	$p = 0.003$
(d)	5.97	5.11	$p < 0.0001$
(p)	6.19	5.67	$p = 0.0004$
(b)	5.62	5.54	$p = 0.44$
(k)	6.17	5.78	$p = 0.0001$
(g)	6.03	5.55	$p < 0.0001$

The summary confirms the tendencies found in the modeling, since the mean local articulation rate for weakened variants is higher than for standard variants for all of the variables, and the difference in means is statistically significant in every case, except for (b), as expected.

The effect of local articulation rate can be handled equally well in any model of the lexicon, since it is necessarily a property of the concrete articulation of an utterance. Whereas effects of stress and segmental context can be viewed as influencing the reduction processes at stages during the planning of an utterance, the effect of increased articulation rate is a factor that influences the timing of the articulators, leading to target undershoot during the actual articulation the planned utterance. As shown in Byrd & Tan (1996) increased articulation rate may lead to gestural overlap in the articulation process, which may eventually lead to masking of articulatory gestures, such that underlying segments are perceived as having been deleted. The data analyzed here further support the hypothesis that articulation rate may lead to target undershoot, a finding that has also been made in other studies of articulation (cf. Fosler-Lussier & Morgan (1999)). Local articulation rate is the most consistent factor in the datasets. It emerges as significant for all but one variable, and it has the same effect for all of the affected variables: the probability for reduction and deletion always increases with local articulation rate. Obviously, segmental reduction and segment deletion are both processes that tend to increase articulation rate, and it is therefore not possible to conclusively assert that local articulation rate is the cause rather than the result of the increased reduction observed at higher rates with the data that are used here. Drawing upon the more fine grained articulatory data of Byrd & Tan (1996), however, suggests that the results presented here may be interpreted as consequences of articulatory adjustment during the production of an utterance.

7.2.9 Position in the phrase

An effect of position in the phrase was only found for the diphthong variables (ɑj) and (ɒw). As mentioned previously, the lack of an effect for the plosive variables may be due to the coarseness of the representation of consonantal variation that is a necessary artifact of the way in which the variables can be represented in the logistic regression modelling. By reducing the degrees of segmental reduction that can be observed in the transcriptions of the consonants, spanning the gamut of deaspiration over frication to deletion of segments, to one single value, reduced, in a binary variable, information about the gradient nature of the variation is necessarily lost. In the studies of laboratory speech where effects for domain initial strengthening have been found (cf. Keating et al. (2003) for a summary of studies), measures including amount of linguopalatal contact and duration of the closure were used to examine the correlation between strength of articulation and position in prosodic domains, notably position in the prosodic phrase. These more fine-grained measures allow for a more nuanced exploration of the effect, whereas the method used here will lead to more conservative results, due to the coarseness of the measure of reduction employed. It is possible then that articulatory or even acoustic studies of the plosive variables in the DanPASS Dialogues may reveal an association between position in the prosodic phrase and gradient plosive reduction in spontaneously spoken contemporary Copenhagen Danish. A recent study of degree of aspiration in syllable initial plosives based on the reading of sentences (Arrild (unpublished)) has shown an effect at the level of the prosodic stress group. In that study, realizations of what corresponds to the variables (p) and (k) in the present investigation were found to be less aspirated in unstressed syllables than in stressed syllables. This indicates a hierarchical effect on the basis of position in the stress group, with plosives being stronger in initial position in the stress group than in medial position. However, no significant difference was found between pre-tonic and post-tonic syllables in Arrild (unpublished), so the formulation in terms of position in the stress group merely reflects a difference between stressed and unstressed syllables, since the prosodic stress group always begins with a stressed syllable (cf. above and Grønnum (1992)).

The effect of position in the prosodic phrase on the diphthongal variables (ɑj) and (ɒw) studied here is congruent with the interpretation of phrase initial strengthening described for four different languages (Keating et al (2003)). Monophthongization was less likely to occur in phrase initial position than in non-marginal position for (ɒw) irrespective of the stress of the token. For (ɑj), however, the situation is more complex. All tokens in single word prosodic phrases were unlikely to be monophthongized. This is largely due to the fact that tokens in single word prosodic phrases were almost all stressed. For tokens in longer prosodic phrases, no difference was found for any of the three positions in the phrase when the examination is restricted to stressed tokens. That is, although monophthongization does occur in stressed tokens, it is equally likely to occur in phrase intial, internal and final position for stressed tokens. For unstressed tokens, however, monophthongization was least likely in phrase initial position, when the likelihood for monophthongization of individual words was taken into account. Doing so revealed that the otherwise high rate of monophthongization (ɑj) in phrase initial position could be derived from the fact that most of the unstressed tokens occurring in this position in the prosodic phrase consisted of the words 'jeg' and 'nej'. As will be shown in the chapter on word frequency effects, these two highly frequent words generally tend to be monophthongized when they are unstressed. Thus, the high rate of monophthongization in phrase initial position may be reinterpreted as a consequence of the type of words which can occur as unstressed in this position, and when this is taken into account, phrase initial position emerges as the phrasal position where monophthongization of (ɑj) is least likely. For (ɒw) position in the prosodic phrase was a main effect, i.e. there was no interaction with any other factor, so nothing conditions the effect of position in the phrase, which shows a statistically significant decreased likelihood for monophthongization in phrase initial position compared to any other position in the prosodic phrase.

These tendencies obviously reflect properties of utterance phonology, since words must have been combined into utterances, that must be grouped into prosodic phrases before the inhibitory effect on monophthongization can be implemented in the articulatory plan. This effect of position in the phrase as well as the very general effect of increased local articulation rate described in the previous section both show that the tendency for reduction can be influenced by global properties of the utterance, and that therefore they are effects which arise at a point in the speech production process where not only the lexical item itself but also characteristics of the utterance in which it is to be produced are available.

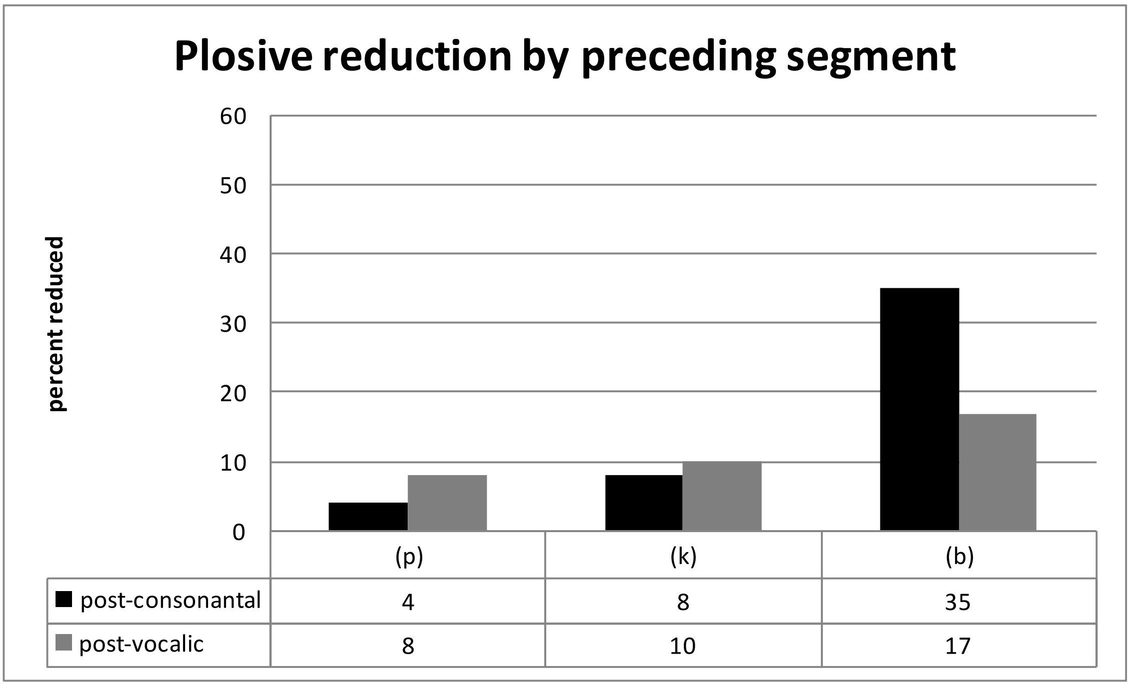
7.2.10 Segmental context in reduction and deletion – similarities and differences

The analyses above drew upon the segmental context to account for the occurrence of reduced variants of consonants in positions where they were least expected, primarily the occurrence of reduced segments in stressed syllable. In this section, I summarize the results for segmental context in general, and show that intervocalic position is not a necessary segmental condition for the reduction of consonants and that plosive deletion in particular is most likely when plosives are surrounded by consonants.

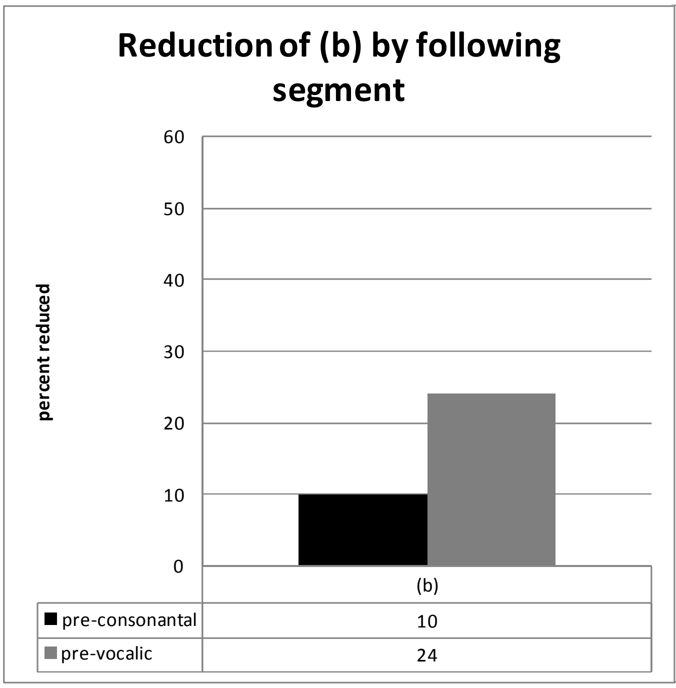
The role of the segmental context in the reduction of the selected consonants and their tendency to be reduced in spontaneous running speech in contemporary Copenhagen Danish was studied for all 7 variables. Investigation of the influence from neighboring segments was sometimes limited to very general properties, e.g. which major class did the adjacent segment belong to, whereas for some variables it was possible to investigate more nuanced effects, like the role of the manner or place of articulation of adjacent segments in the reduction processes modeled. As always, the tests for the role of the segmental contest were designed to reveal whether there were any (statistically robust) differences between the types of segmental contexts attested for each variable, and subsequently, in which kind of context reduction was most likely to occur. In some cases interactions between the preceding and following segment emerged as significant. This section will provide a summary of the effects found and examine the evidence for each variable in order to attempt to find similarities across the variables, in effect exploring the extent to which intervocalic position or homorganicity effects might be observed. Recall that the context is classified with respect to the *phonetic* environment, and not the (morpho)phonological representation of the adjacent segments. In this respect, the effects are all necessarily phonetic implementation effects which emerge during the assembly of the phonetic plan for words in utterances. In this sense they are different from processes of assimilation and dissimilation, which may be determined already at the level of phonological representation

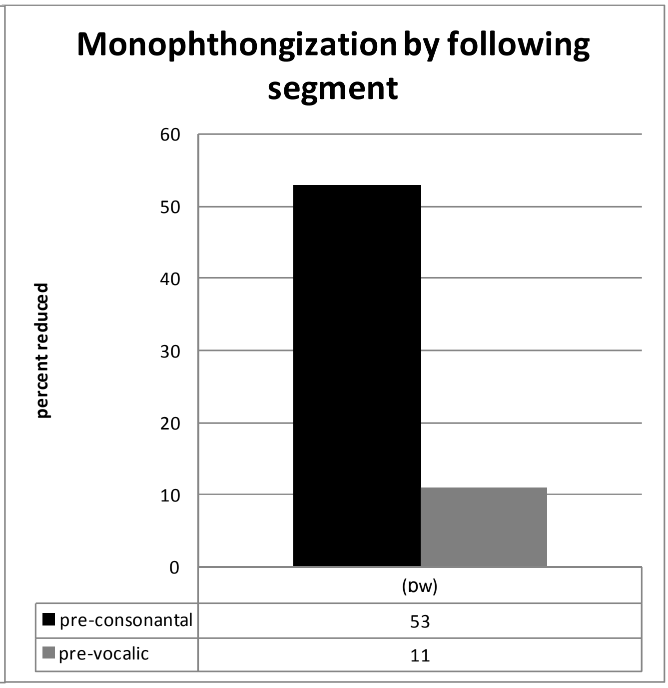
Segmental context effects emerged for (ɒw), (b), (p), (k), (g) and (d). For the variables (ɒw), (g) and (d) the effects were complex, including interactions with other factors and effects of place and manner of articulation.

I start by looking at the difference between preceding consonants and vowels for the variables where preceding segment mattered. Naturally this cannot be done for monophthongization since the semi-vowels that are being deleted in monophthongization are always preceded by a vowel, so there is nothing to compare to. However, for the plosives this has been done in the graph below.



This illustration can only provide a reliable reflection of the variables for which there was a main effect of the preceding segment, and therefore the distributions are not shown for the variables (g) and (d) where the preceding segment entered into an interaction with the following segment. More information is needed to visualize the effect. However, for the variables (p) and (k) the effect apparent from the graph is directly interpretable: reduction is most likely to occur after a vowel. The opposite holds for (b), where reduction is most likely to occur after a consonant. No interaction was found for (b) either, so the effect and the difference between this syllable final variable and the two syllable initial ones can be taken to be robust and reliable, i.e. for all three the illustrated effect is equally likely irrespective of the following segment. However, for (b) a main effect of the following segment also emerged. Therefore I present below the tendency for reduction or deletion of (b) according to the type of the following segment. In addition to (b), the variable (ɒw) was also affected by the type of the following consonant, but entered into an interaction with stress, such that the effect was only observable for unstressed syllables. The effect shown in the graph on the right below is therefore based only on the unstressed tokens of (ɒw). This allows for a comparison of the effect of a following consonant on the two variables where an effect was found. After this I turn to the effect of segmental environment on (d) for which the effect was most complex.



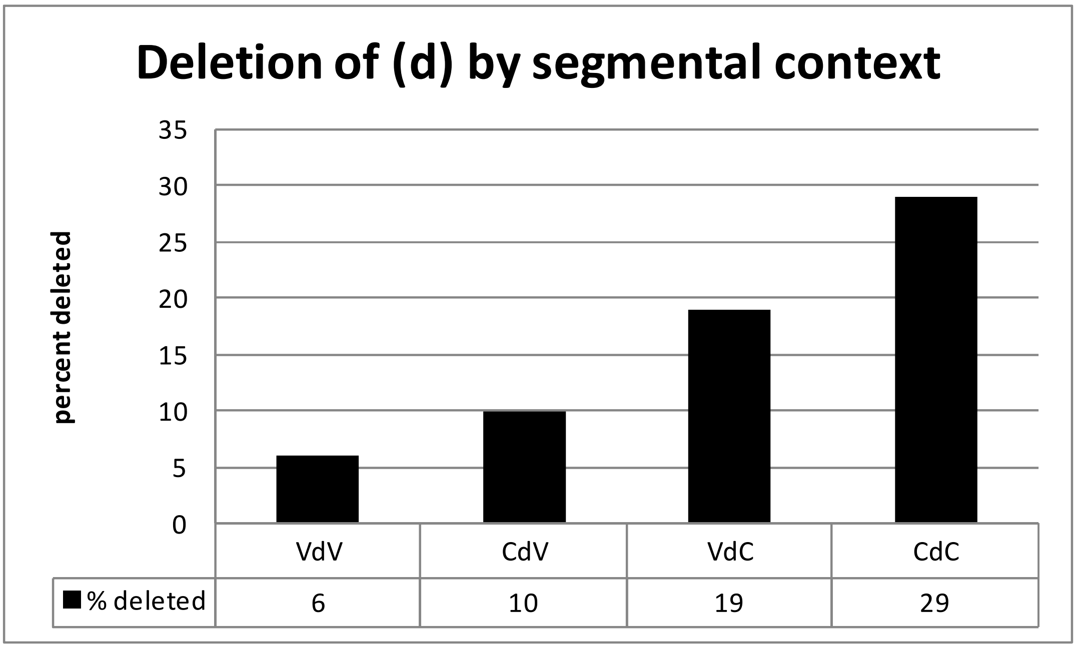


As is apparent from the graphs, the effect of a following segment is very different for the two variables. Reduction of (b) is more likely before vowels than before consonants, whereas monophthongization of (ɒw) is more likely before consonants than before vowels. For (b) since there is no interaction and therefore the effect of the preceding and following segment together is purely additive, we get a “cline” of probability for reduction of (b) with respect to the segmental context, which is represented here going from least likely to most likely environment for reduction:

VbC, VbV, CbC, CbV

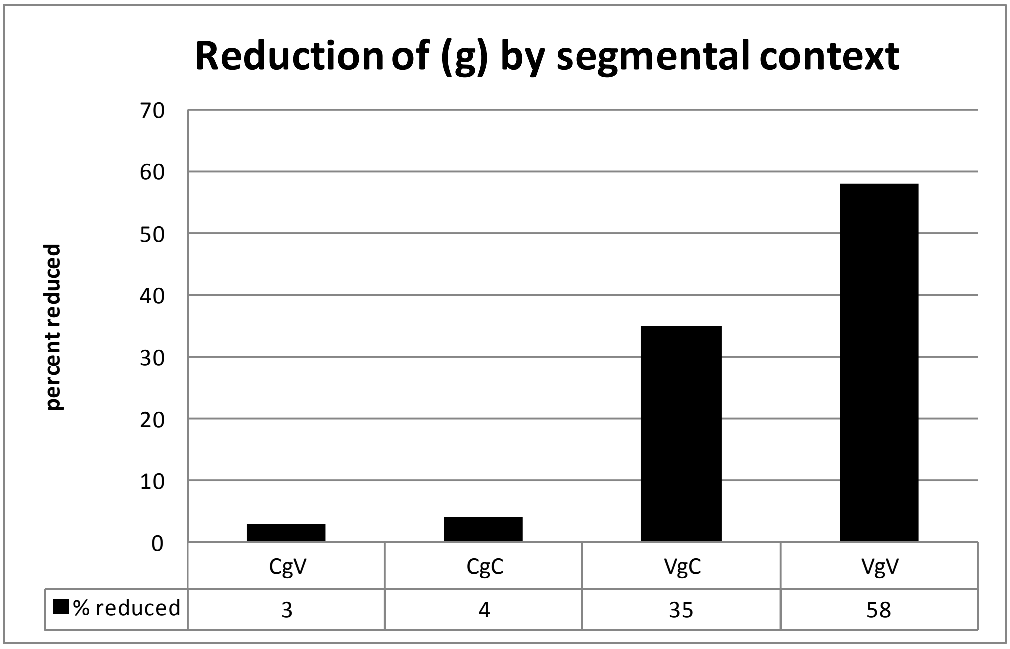
Note that in 13 of the 16 tokens of reduced (b) followed by consonants the consonant is either a fricative or an approximant. The three tokens of reduced (b) that are followed by a stop are all preceded by a vowel. This suggests that reduction of (b) is generally motivated by an adjacent segment having a relatively more open vocal tract for its target articulation than (b) does, and that the reduction can be interpreted as an extension of the vocal tract configuration for the neighboring segment that either continues into the articulation of the (b) or is anticipated during its articulation.

I conclude by illustrating the effect of segmental environment for (d) and (g), which both showed an interaction of the adjacent segments.



As already discussed, (d) deletion can be interpreted as a process of consonant cluster simplification obtaining across word boundaries: relative to the following segment, likelihood of deletion increases as soon as the class switches from vowel to consonant, and when both adjacent segments are consonants, deletion of (d) is most likely. Again, studies within the framework of articulatory phonology may be taken into consideration, since the increased tendency in pre- and inter-consonantal position suggests that what is perceived as a deleted (d) is a product of increased gestural overlap, with neighboring consonantal gestures masking the tongue tip gesture (see also Nolan (1986) for evidence of a residual tongue tip gesture in a sequence where the alveolar plosive is perceived as deleted). It is of course impossible to tell from transcriptions alone whether traces of the gesture for (d) remain in the signal, but the pattern suggests that this may very well be the case. For the contexts in which (d) deletion is rare, transitions are available to provide cues for the (d) even if the explosion of the closure may be weak. However, the fact that deletion of (d) is attested even in intervocalic contexts shows that complete deletion of the segment does occur, and is thus not always a consequence of gestural overlap.

For (g), intervocalic position was found to be the segmental environment in which reduction was most likely. The graph below shows the proportion of reduced tokens of (g) in each of the four possible segmental environments in the order of smallest to highest proportion of reduced tokens.



Comparing the distribution of reduced tokens of (g) to the distribution of deleted (d) highlights the difference between the type of segmental environment that invite reduction and the type that invites deletion. Whereas reduction of plosives appears to be motivated by demands for a less constricted vocal tract for the articulation of one or both of the adjacent segments, deletion, on the other hand, is most likely to occur in environments with a greater degree of constriction, reflecting a preference for simplification of consonant clusters in running speech.

7.3 General summary

The analyses of the phonetic factors summarized in this chapter emphasize that their effects are probabilistic rather than categorical in nature. Conditions may favor reduction, but reductions need not apply even in the most favorable conditions. Patterns of reduction are also probabilistic with respect to levels of structural description, and these effects were seen to be derived from the typical phonetic properties of the individual structural positions. The tendencies found in the analyses of the DanPASS Dialogues are reminiscent of the effects seen in studies of consonant reduction in other languages. This shows that reduction processes in middle class contemporary Copenhagen Danish spontaneous speech are influenced by general properties of articulation. However, as the analyses of effects of word frequency on reduction and on a change in progress will show, the phonetic effects may be modified by language external factors, and the implications for the representation of reduced forms in models of speech production will be discussed in the final chapter.

8. Usage based factors in synchronic variation

In order to study the role of language use in the patterns of reduction of the selected variables in contemporary Copenhagen Danish, two factors were added to the datasets and included in the models developed to assess the influence from phonetic factors and social background factors.

8.1 Word form frequency

An increasing number of studies show a tendency for reduction processes to affect high frequency word forms before low frequency word forms (Bybee (2002) provides a thorough overview). In recent years, these findings have led to models of speech processing which use word frequency as an explanatory factor in accounting for the distribution of reduced variants of phonemes, most notably Bybee (2002) and Pierrehumbert (2001) & (2002). In this usage-based framework, word forms contain reduced variants because they are of high frequency, allowing reduction processes to apply to them more often. The usage-based framework differs radically from models of lexical representation which assume one invariant lexical form representation for each lemma by positing a vast store of detailed phonetic representations of all encountered tokens of each word form in memory. Such memory traces, referred to as exemplars, are organized in acoustic and articulatory space and can be associated with categories at various levels of abstraction (see Hawkins & Smith (2001) for a similar proposal). In this way, the use of words becomes not just an aspect of performance or implementation, but a mechanism for continuous reshaping of the lexicon, and with it, a constant updating of the shape of the acoustic and articulatory space associated with particular phonemes.

To study the frequency effect in the variability of Danish consonants, word form frequency is included as a factor in the investigations of reduction. The studies of leniting sound changes summarized in Bybee (2002) all take frequency counts based on written language as the point of departure for measuring the role of language use in shaping changes in lexical representation by using the corpus quantified in Francis & Kucera (1982). In the present investigation, frequency counts from speech will be used instead. The frequencies were calculated on the basis of the LANCHART Corpus of sociolinguistic interviews. Only speech from informants was used for calculating the frequencies, because I did not want to bias the list with word forms that are used repeatedly by interviewers in the beginning of the interviews, where a fixed list of questions were used to gather background information and to provide a maximally formal situation. The counts were based on the phonetic annotations of the orthographical transcripts of the interviews, which means that they represent word-form counts on the basis of real speech. Incomplete words were excluded. A total of 51.979 different word forms occur in the corpus, with a total of just over 3 million tokens in all (3.088.350). Word forms are assigned relative frequencies by dividing the number of tokens of a given word form with the total number of tokens in the LANCHART Corpus. The relative frequency counts have been log transformed, because there is evidence that language users process frequency in a logarithmic manner, i.e. a frequency difference in the lower frequencies carries more weight than a frequency difference of the same magnitude in the higher frequencies (Mendoza-Denton, Hay & Jannedy (2003), p. 122). Because relative frequencies are always less than 1, the log transformed values result in negative numbers. The factor is represented as a continuous measure.

Note that not all of the words in which the phonetic variables occur in the DanPASS Dialogues also occur in the LANCHART Corpus. For example, many of the words naming locations and objects on the maps used in the map task, like 'bananpalme' [b̥aˈnæːˀnpʰalmə] *banana palm tree* and 'kalkstensklipper' [ˈkʰalg̊sd̥eːnskʰleb̥ɐ] *lime stone cliffs* do not occur in the LANCHART Corpus. While the frequencies in the corpus are assumed to adequately model the relative frequency of each word form in everyday discourse, it is not necessarily the case that word forms which are peculiar to the DanPASS Dialogues are of the lowest frequency possible in the LANCHART Corpus in everyday speech. Since it is impossible to obtain an accurate measure for their frequency in everyday running speech, such tokens are left out of the analyses of the influence of word form frequency on reduction in contemporary Copenhagen Danish. The word forms that have been excluded from these analyses are given for each variable in Appendix B.

8.2 REPETITION

Previous studies have found a repetition effect for segmental reduction. Words show an increased tendency to contain reduced variants of segments, and indeed tend to be shorter, when they are used again in a conversation relative to the first occurrence of the word. This effect has been found for both monologues and dialogues but not word lists, according to Pluymaekers et al (2005), who in their own study found a greater effect of repetition on reduction of suffixes than on reduction of stems. This latter finding may be interpreted as an indication that the repetition effect is a reflection of informational load. Since suffixes carry less information relative to stems, they can be more reduced in repetitions.

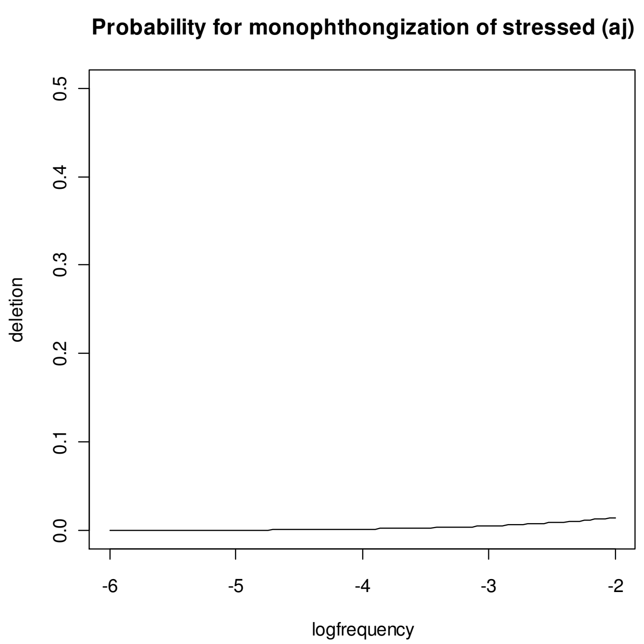
To investigate the possibility of an effect of repetition on consonant reduction in contemporary Copenhagen Danish, each token was coded for whether it occurred in the first mention of a word form in the particular recording session or a subsequent mention. For word forms that were used more than twice in a session, no distinction was made between subsequent mentions. A token was only classified as a subsequent mention if the same speaker had uttered the word previously in the session.

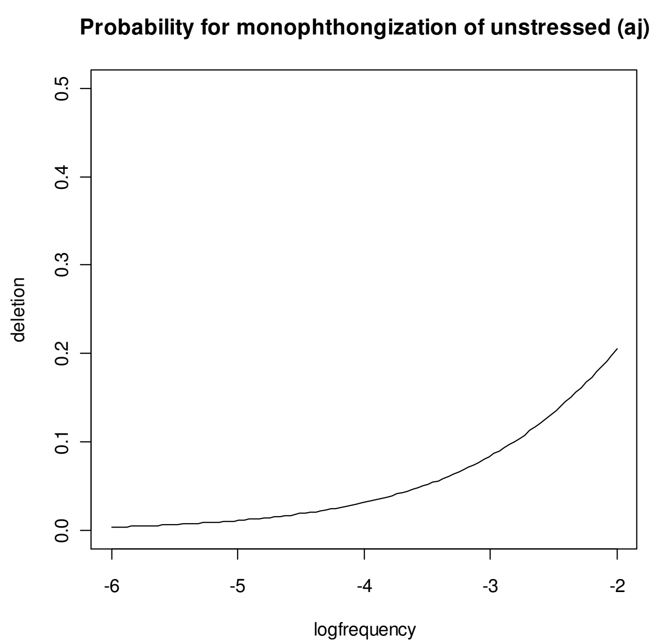
9. USAGE BASED EFFECTS IN CONTEMPORARY COPENHAGEN DANISH

The effects of phonetic factors on the reduction and deletion of selected consonants were described and analyzed in chapter 7. Subsequently, each phonetic variable analyzed in the DanPASS Dialogues was examined for word form frequency effects by adding the factor of word form log frequency to the models. Only 2 of the phonetic variables studied in the DanPASS Dialogues show a statistically significant effect of word form frequency: neither the [ɒw] diphthong, nor the bilabial and velar plosives show evidence of frequency effects on the probability for reduction when the phonetic factors found to be significant in the modeling of their reduction are taken into account. They thus all constitute examples of reduction processes that are not, in a synchronic sample of contemporary Copenhagen Danish, related to word frequency and hence do not support the hypothesis that reduction processes are, necessarily, usage based. I return to a discussion of these variables after a description of the frequency effects found for (ɑj) monophthongization and for deletion of (d).

9.1 WORD FREQUENCY AND (ɑj) MONOPHTHONGIZATION

The influence from word form frequency on monophthongization of (ɑj) was investigated by adding the factor logfrequency of word forms to the final model. All significant factors remained significant at $p < 0.002$. Word form frequency is also significant, but the effect is complex because of an interaction with stress. The effect of word frequency is illustrated in the two graphs below, where the graph on the left shows the effect for stressed tokens and the graph on the right shows the effect for unstressed tokens.





The graphs show that word frequency has no effect on the probability for (ɑj) monophthongization when the diphthong occurs in stressed syllables, but that there is some effect of frequency for unstressed tokens. This does not mean that monophthongization never occurs in stressed items, as the distribution given in table 9.1 below also shows.

Table 9.1 – Distribution of monophthongized [ɑj] in stressed syllables by word form frequency					
Word frequency	-6	-5	-4	-3	-2
Number monophthongized	1	2	5	3	166
% monophthongized	0.5 %	3 %	11 %	3 %	40 %

The high-frequency category does show a marked increase in monophthongization, but it only consists of the words 'jeg, meget, nej' [jɑj ˈmɑjəð nɑjˀ] *I, very, no.* Having included the word form as a random effect allows the mixed model to take into account the fact that some word forms might favor monophthongization over and above the average of the frequency category to which it belongs. Thus, a frequency effect can only emerge if it is stronger than the inter-word variation. Because there is a large amount of variation within the category of high-frequency words, with 'jeg' and 'meget' being almost categorically monophthongized, while monophthongization is very rare in stressed 'nej', frequency does not emerge as significant for stressed syllables. In other words, the increased likelihood for monophthongization in high frequency words apparent in table 9.1 above is an effect of specific words, but not of words of high frequency in general. Not taking this variability into account would have resulted in frequency being a significant factor for both stressed and unstressed items.

For unstressed tokens of (ɑj) however, the effect of frequency is significant, as shown in the graph above. The rate of monophthongization for each frequency category of unstressed tokens is given in table 9.2 below.

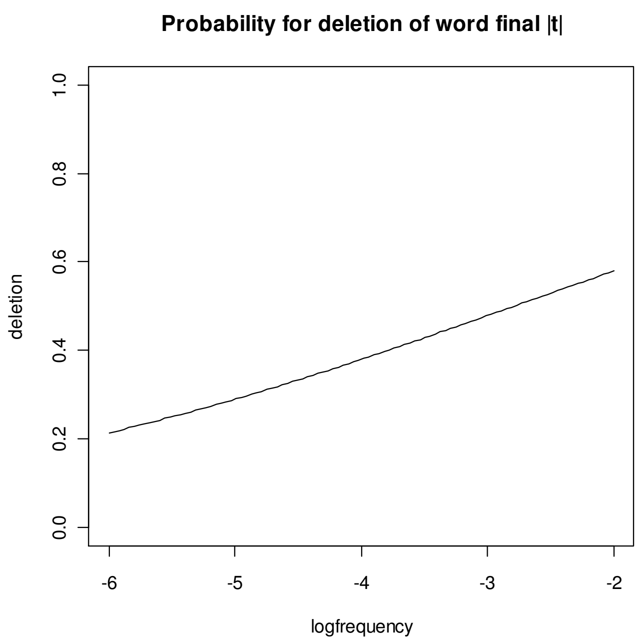
Table 9.2 – Distribution of monophthongized [ɑj] in unstressed syllables by word form frequency					
Word frequency	-6	-5	-4	-3	-2
Number monophthongized	10	9	113	36	173
% monophthongized	8 %	29 %	47 %	88 %	84 %

Table 2 clearly shows the steady increase in rate of monophthongization in unstressed tokens as word frequency increases – with monophthongization being almost obligatory in high frequency unstressed tokens. Note that for the unstressed tokens, only 'jeg' and 'nej' occur in the high frequency category (meaning that 'meget is always stressed in the DanPASS Dialogues). Here, the high frequency words are both more likely to be monophthongized than not, i.e. the variation within the category is smaller than the variation across categories, and therefore frequency emerges as statistically significant, and may be interpreted as a robust effect: when unstressed, (ɑj) is more likely to be monophthongized in high frequency words than in low frequency words.

The significant effect of lexical frequency suggests that tendency for monophthongization is a property stored in the lexicon, but the interaction with stress indicates that a subsequent biasing factor is introduced in the exemplar selection procedure, or in other words, that stress acts as a weight greatly favoring selection from the part of the exemplar space for (ɑj) where diphthongal tokens are located, and for selecting among these exemplars in the procedure of choosing a target for articulation.

9.2 WORD FREQUENCY AND DELETION OF (d)

There are a total of 3385 tokens of words ending in |t|, i.e. [d̥], in the DanPASS Dialogues, but only 3249 of these are tokens of word forms which also occur in the LANCHART corpus, on which the frequency counts are based. In 610 of the 3249 tokens, i.e. 18 %, (d) is deleted, meaning that the subset for which word frequencies can be given have a rate of deletion very similar to that for the entire set of tokens of (d) in the DanPASS Dialogues, in which (d) is deleted in 19 % of tokens.



The deletion of (d), with the standard variant [d̥] does show an effect of word frequency, and the partial effect is illustrated in the graph to the left. No interactions with frequency were found. This is exactly the effect that is expected in a usage-based framework: the more frequent the word, the higher the likelihood that (d) will be deleted, all things being equal. This finding supports exemplar based models of representation, in which the frequency effect on (d) deletion can be modeled as a biasing factor in exemplar selection. In contradistinction to the result found for (ɑj), there is no indication that the effect is cancelled by phonetic factors.

9.3 Lack of a frequency effect

For the variables (ɒw), (p), (b), (k) and (g), word form (log) frequency does not emerge as significant when the factor is added to the models developed in chapter 6 and fit to the subsets of tokens that occur both in the DanPASS Dialogues and the frequency list based on the LANCHART Corpus. While the variation may still be handled within an exemplar based model of lexical representation, the necessity for doing so is less obvious for these particular variables, and furthermore goes against the expectations of the usage based account, in which all reductive phonetic processes are hypothesized to affect high frequency word forms before low frequency ones, a hypothesis that should be reflected in the distribution of reduced variants in a synchronic sample (cf. the treatment of schwa deletion in Bybee (2002) p. 41). These variables thus all give evidence suggesting that non-lexical factors may be more important in influencing reduction than lexical ones. I return to the implications for models of lexical representation in the chapter "Reduction and representation". In the next section, I present the results for the investigation of possible effects of repetition on consonant reduction.

9.4 Lack of a repetition effect

Repetition was measured by classifying tokens of each variable according to whether it occurred in the first mention of a word in a map task recording session, or whether the word had been used previously in the session. However, no effect was found for any of the variables. This contradicts findings reported in Schachtenhaufen (unpublished) on schwa-assimilation in the DanPASS Dialogues. Schachtenhaufen found a difference between first and subsequent mentions for the word form 'kirkegården' [ˈkʰiɐ̯g̊əg̊ɒːˀɐn] *the church yard* and an overall increased tendency for schwa assimilation and deletion as the number of mentions of the word 'kirkegården' increased within a single dialogue (Schachtenhaufen (unpublished) p. 61). Schachtenhaufen notes that other segments are also reduced in subsequent mentions, but not as consistently as schwa.

This failure to find an effect of repetition shows that consonant reduction is unlikely to function as a (crude) cue for information structure in contemporary Copenhagen Danish, even more so given the assumption that all and only first mentions of words represent new information, whereas all subsequent mentions represent given information (Pluymaekers et al (2005) citing Prince (1981)). It is possible that this binary distinction is too coarse and that it may mask an effect emerging in later mentions, especially given that reduction is quite rare for many of the variables studied here. If an effect emerges with a more detailed classification of word repetitions, i.e. where there is a difference in the tendency for reduction not just between first and subsequent mentions, but between second and third or fifth and tenth (as was found in Aylett & Turk (2004)), this would indicate that repetition is more than a reflection of information structure in terms of new and given information, and could instead be interpreted as a recency effect operating at a very local level, i.e. within a single conversation. In other words, repetitions contribute to an incremental adjustment of the individual word form in production, which may not be related to information structure per se. More thorough investigations of the relationship between segmental reduction, word repetition and conversational structure are needed to fully understand the possible functional aspects of reduction of segments.

10. Phonetic variation as change – using synchronic samples to study on-going sound change

The preceding chapters have focused on reduction processes which are best interpreted as processes of stable variation. Although some variables were seen to be more prone to reduction than others, and while two showed evidence of being lexically diffused, none of the variables showed a significant difference in the patterns of realizations between the two generations studied. In the present chapter, I will present the results of analyses of a reduction process which does show signs of being a change in progress. I begin with an introduction to the general methodology of the study of on-going change before introducing the variable.

The study of on-going sound change has traditionally been the purview of sociolinguistics. In this field, identification of an on-going change was until recently based exclusively upon the comparison of the pronunciation of different generations of speakers. Such studies are known as apparent time studies and rest upon the Apparent Time Hypothesis (Labov (1994)). The last decade or so has seen an increase in longitudinal studies of phonetic variation in adults. In studies like these, the pronunciation of groups of adults is studied at (at least) two different points in time typically spaced 2 or more decades apart. In sociolinguistics, these studies are referred to as real time studies. The subsequent analysis of the phenomena studied at the time of the first recording can either be done using the exact same people, a so called panel study, or people from similar social backgrounds as the one studied in the first investigation, a so called trend study. I will return to the methodology of real time studies in more detail in the chapter "Real Time investigations of consonants in Copenhagen Danish". This chapter introduces the Apparent Time Paradigm, the assumptions behind it, and critical evaluations of its usefulness and limitations. Subsequently, an apparent time analysis of deletion of [w] before [ð̞̆] is given, based on the original recordings in the LANCHART Cph Corpus.

10.1 Apparent time

The study of language change lies at the heart of classical variationist sociolinguistics. From the very beginning of the study of social influences on phonetic variation in U.S. English, the objective was also to provide an empirical study of linguistic change and its correlation with social characteristics of speakers (cf. Labov (1963(1972)) and (1966(2006)). At the time of publication of Labov's first studies, the study of sound change relied heavily upon descriptions of pronunciation, or directly on variations in spelling in written texts. No attempt was made to observe changes in progress, and as noted by Labov many thought such an endeavor to be impossible from the outset, cf. Bloomfield's (1933) assertion that the “process of linguistic change has never been directly observed [and] such observation, with our present facilities, is inconceivable.” (Bloomfield 1933 p. 347 quoted in Labov (1963(1972)) p. 22).

Due to the lack of recorded speech material from the areas studied (Martha’s Vineyard and the lower east side of Manhattan in New York City), Labov’s initial investigations of on-going sound change relied on evidence from apparent time investigations[21]. Labov (1966(2006)) states that “[b]y studying the differences between the linguistic behavior of successive age levels in our sample, we can make inferences about linguistic change.” (p. 200). In a strictly synchronic perspective, what the apparent time paradigm does, then, is compare the distribution of variants between different generations, i.e. not studying the change in distributions at different points in time but the difference between distributions of variants at the same point in time, but within different groups in the speech community, groups that are defined by their age.

The limitations of this methodology were of course also recognized immediately, and the validity of the inferences about sound change was acknowledged to be restricted to circumstances where “1)...there are no differences between older and younger speakers *which are repeated in each generation*; and 2)...the older speakers remain isolated from the effects of the language used by younger speakers.” (pp. 200-201, emphasis mine). In other words, if we are to say that the difference in pronunciation of the vowel in words like ‘ret, frem, præst’ [ʁæd̥, fʁæmˀ, pʰʁæsd̥] or [ʁɑd̥, fʁɑmˀ, pʰʁɑsd̥] *straight, forward, priest* that we observe between 20 year old speakers and 60 year old speakers in 2008 reflects a sound change that has progressed within the past 40 years, we must assume that the speech of the 60 year olds today is, at least with respect to their pronunciation of vowels, identical to their speech 40 years ago.

Assumption 1) above concerns the phenomenon of generation specific speech behavior, or age-grading of linguistic variables. While age-grading is perhaps more pervasive in the vocabulary than in the phonology, it is not impossible that a particular phonetic variant, or rather very frequent use of a particular phonetic variant, may be associated with the speech of a particular age group, just as it can be associated with gender, social class, and/or style. In other words, when a variable violates assumption 1), it cannot be interpreted as on-going change, but must rather be seen as stable variation related to specific age groups, possibly related to a speaker’s maturation as suggested by Chambers (2003), p. 206). Chambers (2003) also states, that age-grading is quite rare and that the majority of age-graded variables show evidence of an abrupt change in late adolescence (p. 206). This is supported by the introduction given in Meyerhoff (2006), where age-grading is presented as a phenomenon that is

[21] Note that comparison of the speech of speakers from different generations had been used in dialectology by Gauchat (1905), but that Labov’s study of Martha’s Vineyard is the first one to use this method again to study language change.

connected to the transition from the relatively closed social space of educational institutions to the more varied sphere of the job market, where younger speakers are exposed to the speech of older generations to a greater extent, but where they are also, in some professions, expected to conform to speech community norms, and hence to abandon the use of overtly stigmatized forms (Meyerhoff (2006) pp. 145 -149). Age-graded variation may, according to Chambers (2003) and Meyerhoff (2006) be identified even in a synchronic sample, because it typically involves extremely frequent use of the non-standard variant among adolescent speakers, with a rapid decrease through successive generations in favour of the standard variant. Naturally, the only way to be certain that such a peak *does* constitute age-grading, rather than (abrupt) incipient change, is to examine either previous recordings of different speakers from the same age group, or to conduct a real time study. The present study includes a group of adolescent speakers (referred to as generation 2 in the description of the LANCHART Cph corpus), but it is not possible to compare them to previous recordings of adolescent Copenhageners. Instead, any sign of apparent time change must be re-evaluated in the subsequent real time analysis of the same speakers.

Assumption 2) above regarding the stability of adult speakers can only be evaluated by real time evidence. Naturally, individuals who continue to participate in an on-going change throughout adulthood will blur the picture that can be seen from an apparent time sample, and if the majority of adults have continued to participate in the change being studied, it may not be possible to discover it in an apparent time study. The increasing amount of real time evidence suggests that such a situation is unlikely (cf. Sankoff (2006)), however, these are issues to which I return in the chapter "Real Time". For now it will suffice to say that precisely because the stability of adult speakers had to remain an assumption in the first variationist studies of language change in progress, the apparent time paradigm has remained a *hypothesis* about the relationship between the speech of different generations and diachronic change. Many researchers (cf. Bailey (2002), Chambers (2003), Sankoff (2006)) emphasize the degree of similarity between results from apparent time studies and the few real time studies that exist, while Labov (2001) shows greater caution in the light of the accumulating evidence of real time change, stating: "the assumption of stability for young adults...may have to be revised."(p. 447). This is particularly important in a study like the present one, which involves comparison between adolescent and adult speakers.

11. Apparent Time analysis of consonant reduction in contemporary Copenhagen Danish spontaneous speech

This section consists of an analysis within the apparent time framework using the LANCHART Cph Corpus. The analyses of the auditory classification of variants of the phonetic variable (wəð) is carried out using mixed-effects multiple logistic regression. The purpose of this section is to investigate signs of apparent time change in a well-known variable, and to examine the degree of homogeneity of predefined groups in the population sampled.

11.1 [w] deletion before [ð̩] in apparent time

The variable to be investigated in apparent time in the present study is the realization of [w] before the sequence [əð] when the latter is realized as syllabic [ð̩] due to assimilation of schwa. This process has been described in Copenhagen speech previously and has been known to be going on for some time. Therefore, this apparent time study of the variable can be considered as a further exploration of the variable in order to supplement what was known about the distribution of its variants at the time of recording the original LANCHART Cph data.

It should be noted at the outset, that [w] does not occur after high back vowels, i.e. /u (ː)/ and /o (ː)/ (Grønnum (2005), p. 295-296) in contemporary Copenhagen Danish, and this is irrespective of the succeeding phonetic context. Therefore, tokens of [w] after high back vowels and before [ð̩], as in e.g. 'overhovedet' [ɒwɐˈhoːwəðəd̥] *at all* are excluded from the present investigation.

11.2 History of [w] deletion before [ð̩]

The process has so far been studied as part of the more general process of monophthongization before [ð̩] (Brink & Lund (1975) pp. 346-352). Brink & Lund (1975) connect the process of [w] deletion to the process of [j] deletion in the same context, but they note that whereas deletion of the palatal glide becomes obligatory in this context already for speakers born in the beginning of the 20th century (p.349)), deletion of the labio-velar approximant does not become obligatory even in the speech of their youngest informants. Both processes of monophthongization are more frequent after vowels that are similar to the approximant in place of articulation, and become rarer the further apart the places of articulation are, such that [w] deletion is least frequent after [i], although it does occur. A few vowels seem to block the process entirely, since [w] deletion never occurs after: short [ɑ] and long and short [ɒ] (Brink & Lund (1975), p. 349). The process is conditioned by the process of assimilation of schwa to [ð] resulting in syllabic [ð̩], a pervasive process in Copenhagen speech throughout the 20th century and in contemporary spontaneous speech (cf. Brink & Lund (1975) pp. 191-220). Brink & Lund (1975) state that assimilation of schwa to /ð/ must take place in the individual token in order for [w] deletion to occur. This is also the explicit assumption behind Grønnum (2005)'s auditory motivation for the process as change, where the deletion of [w] is interpreted as a process of auditory reanalysis (along the lines proposed in Ohala (1981): since Danish [ð] is slightly velarized in any style of speech, this velarization can become perceptually more salient and the velar characteristics of the approximant [w] can be attributed to the following [ð], that is the rounding in [w] is in effect drowned by the velarization, leading to deletion of [w] in running speech (Grønnum (2005), p. 334). This factor was not included in the classification of variables, so it will not be possible to quantify to what extent deletion of [w] before [ð̩] is correlated with increased degree of velarization of the [ð̩].

The tendency for deletion of [w] is also linked to stress in Brink & Lund (1975, p. 349). It is observed that when the process of monophthongization applies in contexts that do not favor it, e.g. [w] deletion after the front vowel /i/, it is more prevalent in syllables with weak stress than in syllables with full stress. Thus, the realization [liːˀð̩] for 'livet' [liːˀwəð] *life* (sg. def.) is rare in fully stressed tokens, and more common in compounds such as 'kirkelivet' [ˈkʰiɐ̯g̊əˌliːˀwəð] *church life* where the syllable containing [w] only has secondary stress in the compound.

Recent observations indicate that the process of [w] deletion may be spreading. Grønnum (2005) adds long [ɔː] to the set of vowels where [w] may be deleted irrespective of phonetic and phonological environment, i.e. also when not followed by [ð̩]. However, it is noted that this is only a tendency and not an obligatory process like the deletion of [w] after long [uː] and [oː]. Furthermore, it is noted that [w] deletion also occurs after high front vowels even in stressed syllables, i.e. 'livet' [ˈliːˀwəð] *life (sg. def.)* may be affected by monophthongization regardless of the degree of stress (Grønnum (2005), p. 334) contra Brink & Lund (1975).

11.3 METHODOLOGY: CODING AND BACKGROUND VARIABLES

All tokens of stem final [w] before the suffix [əð] were found by searching for sequences of 'w0D' and 'e:0D' and 'i:0D' , i.e. [wəð], [eːəð] and [iːəð] respectively, in the phonetic annotations of the LANCHART Cph 80s corpus, with two general exceptions: tokens of [w] after [u(ː)] and [o(ː)] were excluded, since original [w] after high back vowels are obligatorily deleted in contemporary Copenhagen Danish irrespective of further phonetic factors (cf. above).

Including the sequence 'e:0D', in the search yielded a case of false tokens of the variable: the word 'steget' [sd̥e:jəð] *risen*, which in the annotations of the LANCHART Cph corpus is represented as 'sde:0D', due to the obligatory deletion of /j/ in this context (cf. above). A second type of false token, which was not included in previous studies of the process of [w] deletion, was the occurrence of [w] before the suffix [əðə], since the intended context, [əð]#, forms a subset of the string [əðə] when, as here, the word boundary is not specified in the search. Since it was found that the initial schwa of this suffix overwhelmingly tended to assimilate to the following [ð], thereby allowing for the possibility that the circumstances needed for the process to apply were present, these tokens were also included in the analysis.

Each token was coded for three target variants: deletion, realization as [w], and realization as [v]. Deletion was defined as 'no audible trace of a rounded, non-syllabic segment intervening between the nucleus of the syllable containing underlying [w] and the [ð] of the suffix'. In cases where the classification was uncertain, the coder put an asterisk, *. For the statistical modeling, only tokens in which both coders were certain of the classification are included. Each token was also coded for the stress on the syllable containing the underlying [w]. Coders only supplied "perceived stress", whereas some of the unstressed tokens were supplied with a mark for secondary stress in order to study the hypothesis that syllables with secondary stress resemble fully stressed syllables segmentally (Grønnum (2005), p. 250 and see the analyses of consonants in the previous chapters). Each classification was aligned to the phonetically annotated orthographic representation of the token, allowing for analysis of the influence from word frequency and underlying preceding vowel. The extra-linguistic background variables age and gender are obtained automatically from the file name of the transcripts of the recordings in which the codings of the variable are saved.

11.4 Deletion of [w] in the mid-1980s

The data were coded by two listeners, and the codings of the full set of 418 tokens were compared. First of all, tokens where at least one coder was in doubt about the classification of the token, i.e. tokens where one or both coders could not determine whether the [w] was realized or deleted and tokens where at least one coder thought there was too much noise to make a reliable classification were discarded. This reduces the dataset to 362 tokens of the variable (38 tokens discarded due to background noise, 18 tokens discarded due to coder uncertainty). The remaining tokens were analyzed for intercoder agreement using Cohen's Kappa. This test yielded a value of 0.61 – a quite modest level of agreement. In order to be certain that the tendencies captured by the statistical modeling is based on valid data, only tokens where the two coders agreed on the classification are included in the final set. This leads to a dataset of 336 tokens of the (wəð) variable.

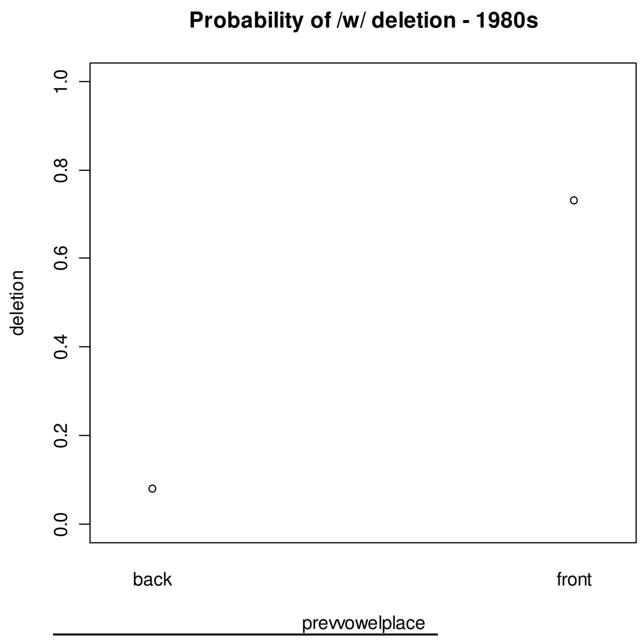
11.5 MODELING THE VARIATION

In order to test whether there is evidence of an apparent time change in the deletion of [w] before [ð̞] while controlling for stress and the quality of the preceding vowel, a logistic regression model with the factors 'stress', 'previous segment' and 'decade' was fit to the data. The latter factor contains information about the decade in which the informants were born, thus providing a more fine grained factor than the two generations which the subjects sampled are intended to represent. To examine whether there is a difference between men and women's tendencies for [w] deletion, the factor 'gender' was included in the model. Finally, in order to control for the influence from individual speakers, a mixed-effects logistic regression model was fit, with 'subject' as random factor. To model deletion of [w], the reference level of the dependent variable was set to 'realized'.

The dataset is quite unbalanced with respect to the representation of different preceding vowels. Not only are there some vowel phonemes which do not occur before (wəð) in the dataset, but for some of the vowels, the data is also very sparse. Furthermore, a preliminary pure fixed effects model that included the quality of the preceding vowel as a factor, had great risk of collinearity (vif for several factors = 10). This risk persisted even when attempts were made to make the dataset more balanced with respect to the quality of the preceding vowel, by conflating sparsely represented vowels with more richly represented and articulatorily similar vowels. In order to investigate the influence of the quality of the preceding vowel, two factors were devised, which encoded articulatory aspects that appear to be relevant both in the light of previous investigations as described above and on the basis of the distributions of variants with respect to preceding vowel quality. These factors are: place of articulation, given here in terms of the place of the highest point of the tongue dorsum during articulation, i.e. either front or back, and vowel height, either high, highmid, lowmid, or low. Only place of articulation of the preceding vowel was included in the initial model.

Recall that two types of stress have been coded in the dataset: perceived stress and lexical stress. It is to be expected that there is a high degree of correlation between the two measures, in particular, that syllables lexically specified for secondary stress, will be realized without stress, whereas syllables lexically specified for primary stress will be realized both with and without stress. The factors show a modest but highly significant correlation ($p < 0.001$, Kendall's tau = 0.247), and a cross tabulation confirms the expectation: 95 % of syllables specified for secondary stress are realized as unstressed (43 tokens out of 45), whereas 61 % of syllables specified for primary stress are realized as unstressed (178 tokens out 281). However, a test for collinearity does not reveal any cause for concern (all vifs < 2), and therefore both factors have been included in the model[22]. The high amount of syllables specified for primary stress that are realized as unstressed are almost all entirely due to unstressed tokens of the word 'blevet' [ˈb̥leːwəð] *got (aux.)*, which make up 110 of the 178 tokens of syllables specified for primary stress that are realized as unstressed. Of the remaining 68 tokens, 30 consist of the verb 'lavet' [ˈlæːwəð] *made.* Both verbs may function as auxilliaries, and when they do they are unstressed (Rischel (1983)).

Table 11.1 - [w] deletion, all speakers			
Factor	**Estimate**	**P**	**Sign**
Perceived stress -unstressed	1.30651	0.001	**
Lexical Stress - secondary	-2.12939	<0.0001	***
Preceding vowel front	3.41366	<0.0001	***
Decade	0.0698	0.0251	*

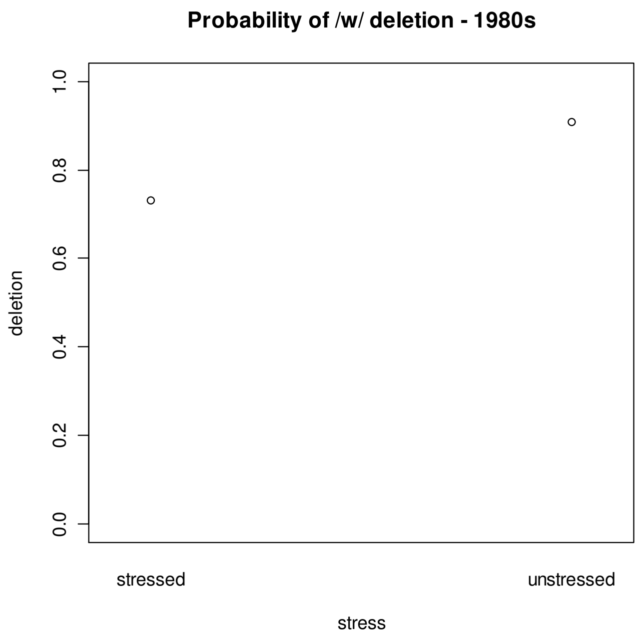


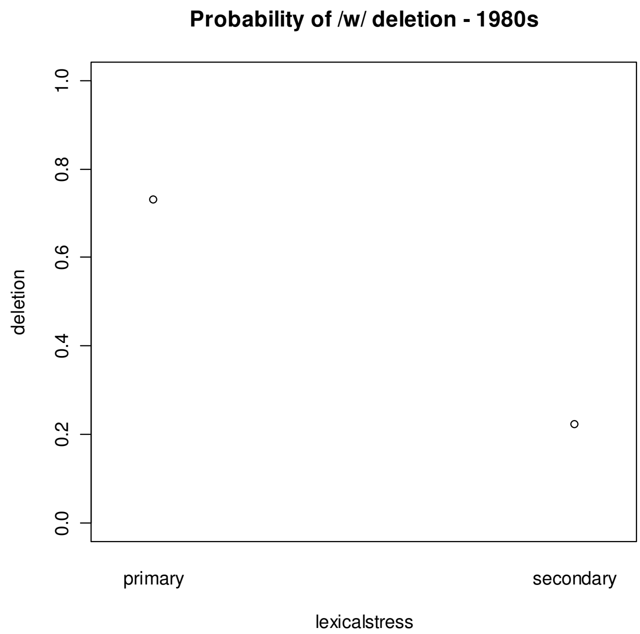
Only four of the factors have a significant effect on the probability for [w] deletion. Gender is non-significant, and has therefore been removed before fitting a final model. The goodness of fit measures for the model are: C = 0.89 and Dxy = 0.78, indicating a very good fit. The estimates and significances of the factors that were retained in this model are given in table 11.1.The table shows that [w] is more likely to be deleted in unstressed syllables than in stressed syllables, but is less likely to be deleted in syllables with secondary stress as compared to syllables with primary stress, that [w] deletion is more likely to occur after front vowels than after back vowels, and that the probability for [w] deletion increases with decade of birth, i.e. younger people are more likely to delete [w] than older people. The graphs below show a transformation of the log odds given as the estimates in the table into probabilities. For each factor, the y axis shows the probability that [w] will be deleted. The relation between the place of the preceding vowel and the tendency for [w] deletion is shown first. Clearly, [w] deletion is most likely to

[22] A test for an interaction of lexical stress and perceived stress showed that the interaction was not significant. Therefore a purely additive model has been chosen for the analysis.

occur after a front vowel. This is in accordance with the findings of Brink & Lund (1975) as described above, although it should be recalled that they found [w] deletion to be least likely after /i/ and stated that this only occurs in syllables with weak stress. Grønnum (2005) states that [w] may be deleted after high vowels in general (p. 260), suggesting that vowel height should be introduced in order to fully understand the influence from the preceding vowel. A model in which this factor was added shows that [w] deletion is least likely to occur after high vowels, and most common after high-mid vowels, i.e. /e/, and not very likely after low-mid and low vowels. The characteristics of the present model of [w] deletion in the 1980s, then, are in accordance with the results presented in Brink & Lund (1975).

Next, I turn to the effects of perceived stress and lexical stress. The reference level for place of articulation for the vowel has been changed to front, in order to show the partial effect of stress in the environment that favors [w] deletion the most. The relative relationship between the levels of the factors is the same, since there is no interaction between neither perceived stress and vowel place nor lexical stress and vowel place.

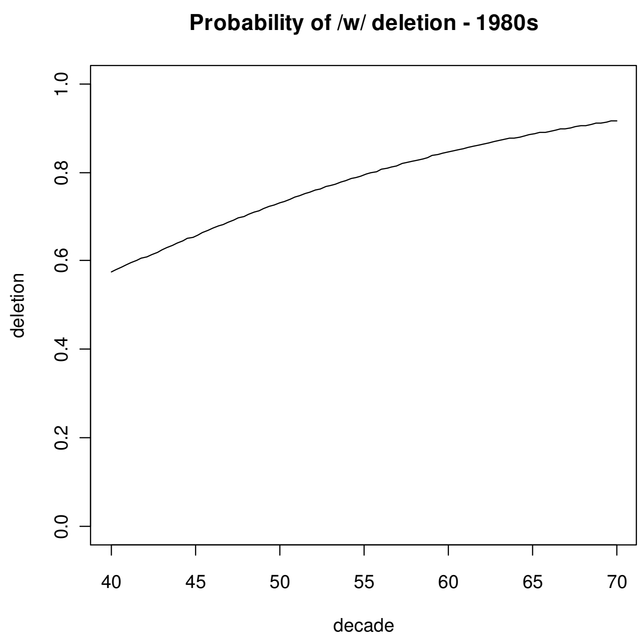




The difference between stressed and unstressed syllables is not very large, so the effect of perceived stress, while statistically significant, can be said to be slight. The effect of lexical stress is larger – there is a clear difference between the likelihood for [w] deletion in syllables specified for primary stress as opposed to syllables specified for secondary stress. As could be seen in the table of coefficients, secondary stress inhibits the process of [w] deletion, at least to some extent. I interpret this as an reflection of how well segments are preserved when they occur in positions in compounds where, had they occurred in non-compounded forms, they would have been fully stressed. In other words, this indicates that the information about lexical stress influences speech production. It contradicts the observation in Brink & Lund (1975) that syllables with weak stress are generally more likely to be affected by deletion of [w], since the majority of tokens (95 % or 43 out of 45) with secondary lexical stress, are realized as unstressed. The higher tendency for syllables with primary stress to be affected by [w] deletion can be ascribed to the loss of primary stress for syntactic (and other) reasons in connected speech.

The factor age has the expected effect, which can be seen quite clearly in the graph below. Again, the reference level for the place of articulation of the preceding vowel is front. The model clearly shows that the younger the speaker, the more prone to [w] deletion before [ð̩], all things being equal.

It is safe to say that [w] deletion is a common process for all speakers in this sample, and it appears to be all but obligatory for the younger generation, already in the mid 1980s. Thus, the process identified in Brink & Lund (1975) has continued, but has not yet been brought to completion. In chapter 9 "Usage based effects in contemporary Copenhagen Danish" I will examine the question of whether word frequency plays a role in exerting the small amount of inhibition of the process seen here.



However, while the analysis so far has shown an effect of speaker age, thereby supporting the hypothesis that [w] deletion before [ð̩] continues to be a change in progress, it remains to be seen to what extent the phonetic factors that influence this process are equally relevant for different subgroups of the speakers sampled. While the modeling shows that gender is not an important factor in [w] deletion, it is not possible to determine whether the two generations sampled differ with respect to the constraints imposed by stress and vowel quality on the basis of the initial model given in table 11.1 above – although, clearly, the effect of decade shows that the tendency for [w] deletion is correlated with age. To examine this possibility, the initial model has been fit to two subsets of the data: one including tokens produced by the older generation and another including tokens produced by the younger generation. The characteristics of these models are given in the two tables of coefficients below.

Table 11.2 - Older generation – 1980s			
Factor	**Estimate**	**Std. Err**	**Pr(>|z|)**
(Intercept)	-10.79	3.20	0.002 **
unstressed	1.29	0.46	0.005 **
secondary	-2.22	0.56	<0.0001 ***
front	3.35	0.86	<0.0001 ***
decade	0.18	0.06	0.01 *

Table 11.3 - Younger generation – 1980s			
Factor	**Estimate**	**Std. Error**	**Pr(>|z|)**
(Intercept)	2.42	12.80	0.85
unstressed	1.44	0.82	0.07 .
secondary	-2.17	1.07	0.04 *
front	5.79	2.11	0.007 **
decade	-	-	n.s.

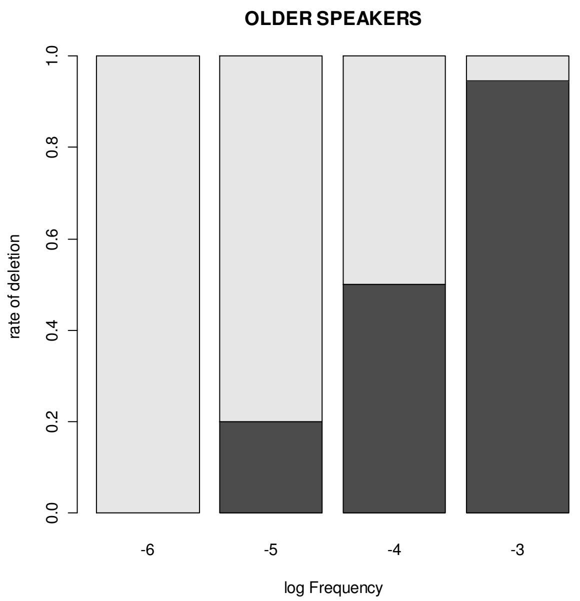
The tables indicate that the two generations are largely alike with respect to the influence from the phonetic environment on [w] deletion. Note, however, that perceived stress only approaches significance for the younger generation. This could be taken as a sign that stress is losing its function as an inhibitor of [w] deletion, indicating that the process is being generalized to new contexts in the younger generation compared to the older generation. Furthermore, the speakers in the older generation are not as homogeneous as those of the younger generation with respect to [w] deletion: decade of birth emerges as statistically significant for the older generation but not for the younger generation.

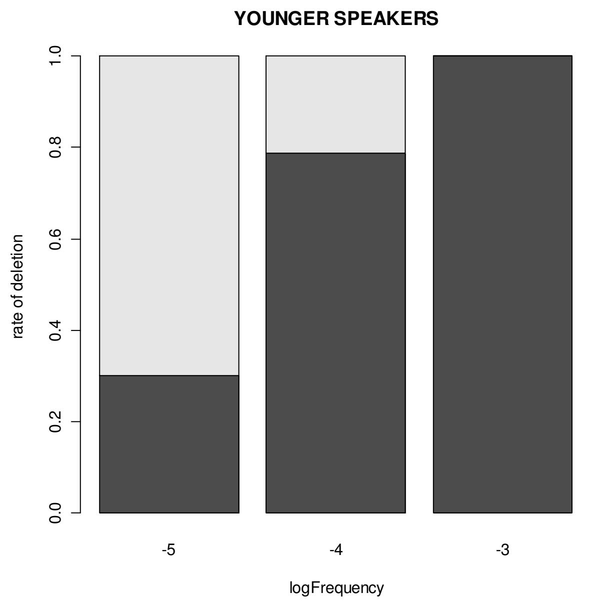
11.6 WORD FORM FREQUENCY IN APPARENT TIME

So far, I have dealt with the influence of phonetic factors on the probability for deletion of [w] before [ð̩]. While phonetic factors involved in the conditioning of a sound change are indicative of the way the properties of the vocal apparatus and hence the process of production itself conspire to induce reductions, quantitative analyses of the influence of phonetic factors like the one shown above imply that reduction processes cannot be simply automatic consequences of specific articulatory mechanisms. The quantitative analysis suggests that phonetic effects are tendencies rather than obligatory rules. Some other factor must be involved in "co-determining" reduction. The factor that I will explore is language use in terms of the frequency of words in running speech. Studies of several languages have shown an effect of word form frequency on the spread of sound change, with reductive sound changes first occurring in high frequency words and gradually spreading to words of lower frequency as the change spreads over time (cf. Bybee (2002). This effect of word frequency has been interpreted as an influence of language use on phonology: the number of times a word is used in running speech affects its propensity for participation in a process of change – or rather, affects the probability that speakers will use reduced variants in the pronunciation of the word. I will return to the question of how the individual speaker's tendency to reduce segments is affected by the frequency of the words in which the segment occurs in the conditioning phonological context after giving an analysis of how word form frequency, and hence word use, can be seen to correlate with change as indicated in apparent time at the group level.

The proportion of tokens where the underlying [w] is deleted is shown for each of the two generations in the bar plots below. Each bar represents a frequency category, rate of deletion is given on the y-axis and word frequency increases from left to right on the x-axis. Because the relative frequencies are all below 1 and have been log transformed, the values for each frequency category have a negative number. The frequencies are based on the entire corpus from which the frequency counts were computed. Therefore the highest frequency category seen here is -3, because there are no word forms with the relevant sequence of distinct phones, i.e. [wəð], in the frequency categories -2 and -1.

Dark grey portions of the bars give the proportion of deleted [w] for that frequency category.





Both groups show the expected pattern: deletion of [w] is more likely in high frequency words and even obligatory for all speakers in the younger generation in the highest frequency category, than in words of lower frequency, and it never occurs in the rarest words in running speech, as can be seen from the bar plot for the older generation. Here, some speakers produced words with a log frequency of -6, but [w] was retained in all of them. No speakers in the younger generation produced any tokens of words from this frequency category, so it is impossible to tell whether the younger generation has in fact extended the process to words of lower frequency than the older generation. That is, the data cannot be used to test whether an on-going sound change spreads to words of lower frequency from one generation to the next, but it is possible to test whether the process of [w] deletion before [ð̞] is intensified in high frequency words. The lack of tokens of words of the lowest frequency is not as surprising as it might first seem, given the circumstances of the interviews: the interviews conducted with the speakers in the younger generation were typically shorter than the interviews conducted with speakers from the older generation. Furthermore, it should be noted that the frequency categories do not contain an equal number of tokens. This can be seen in the tables below, where the number and proportion of tokens for each frequency category is given for each of the two generations.

Table 11.4 – Distribution of tokens by frequency category, older speakers				
log frequency	-6	-5	-4	-3
n realized	5	12	33	7
n deleted	0	3	33	120
% deleted	0%	20%	50%	94%

As the table shows, even the older speakers produce only 5 tokens of the very rarest words, meaning that not all speakers produce words of this frequency category. This is an indication that a more frequently occurring variable would be an even more interesting case for an exploration of reduction in words of very low frequency.

However, what can be explored is whether the apparent difference between propensity for [w] deletion before [ð̞] in high frequency words in the two generations as shown in the bar plots above is statistically significant for each of the three comparable frequency categories. The distribution by word frequency category in the younger generation is given in the table below.

Table 11.5 – Distribution of tokens by frequency category, younger speakers			
log frequency	-5	-4	-3
n realized	7	7	0
n deleted	3	26	80
% deleted	30%	79%	100%

This table shows that the group of younger speakers produced fewer tokens of the variable but also had higher rates of deletion in the frequency categories -5, -4, and -3. The question is whether these higher rates are significantly different from the rates of the older generation. The difference between the distributions for the frequency category -5 is tested using Fisher's exact test, since the expected frequencies are too low to allow the use of Chi square. As expected the significance test reveals no statistically significant difference between the distribution of realized and deleted [w] in words with a log frequency of -5 for the two generations (p = 0.65). For the next highest frequency however, a chi square test is possible and it returns *p* =0.005 meaning that the rate of deletion at 79 % for the younger generation is significantly higher than the 50 % found for the older generation. A Fisher's exact test of the differences between the generations for the high frequency category -3 reveals that this difference is also significant (*p* = 0.04).

Thus one of the expectations at the group level for the frequency effect in apparent time is borne out: the deletion of [w] is continued by the younger speakers by increasing the rate of deletion in relatively high frequency items.

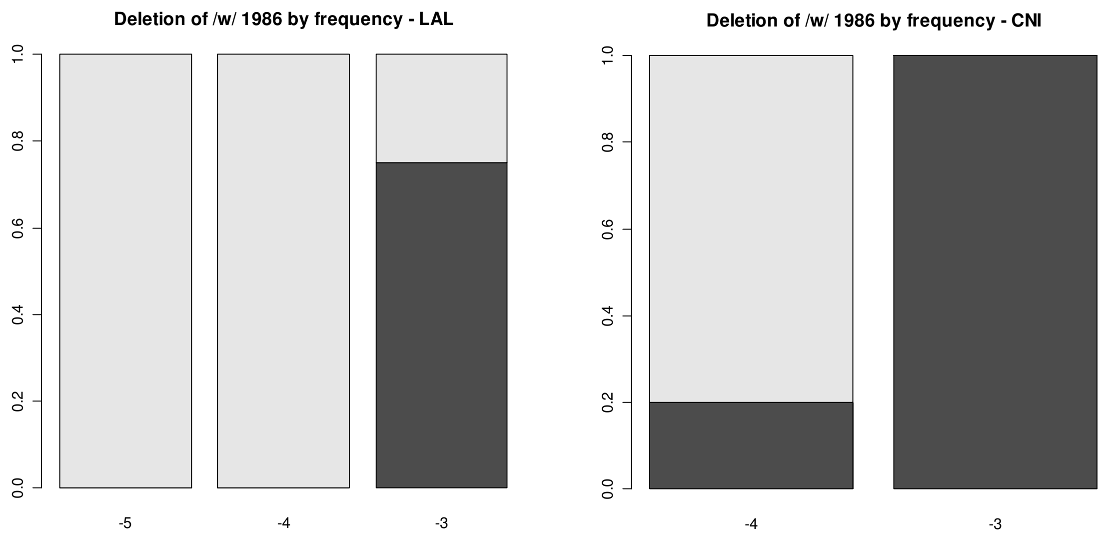
11.7 INDIVIDUAL SPEAKERS AND WORD FREQUENCY

The significant difference between the two generations in their propensities for [w] deletion and the correlation with word frequency shows that the process is still in progress, possibly on its way to completion, but it does not tell us whether the individual speakers all follow the observed trend. The significant effect of age group only shows that we have a greater probability of finding a token of a word containing the sequence [wəð] in which the [w] will not be realized if we listen to the speech of a person from the younger age group.

While the advantage of mixed-effects models allows us to infer that the factors that emerge as significant do in fact hold for the entire sample studied and do not just stem from contributions by individual members of the groups, they do not tell us anything directly about the phonetic behavior of the individual speakers. To know more about the behavior of individuals and thereby about the differences and similaritites between individual phonologies, it is of course necessary to analyze each informant in more detail, in order to see how the effects of phonetic and extra-linguistic factors found for the group pattern in the speech of individuals. As mentioned previously, the dataset is unbalanced, and in particular there is a great deal of difference between the number of tokens produced by individual speakers: the highest amount of tokens produced by one speaker is 38 tokens while the lowest is 4. This means that analysis of individuals will in many cases be qualitative, since speakers who produce very low amounts of tokens of the variable cannot be meaningfully analyzed statistically. But it is still possible to investigate the trend for each speaker with respect to the different factors that have been found for the group. In what follows, I will limit myself to analyzing the behavior of speakers who have produced at least 10 tokens of (wəð). The analyses here concern the relationship between word frequency and probability for [w] deletion in individual speakers. As was shown above, the major difference between the two groups lay in the propensity for deletion of [w] in high frequency words, with speakers of the younger generation having obligatory [w] deletion before [ð̞] in words of frequency category -3, the highest level of frequency for words containing (wəð). The next level of frequency, -4, was seen to be the group of words where the difference between the two generations was most pronounced: in the older group 50% of these words were realized with [w] deletion, whereas [w] deletion applied in 79 % of such words in the younger group. The focus will therefore be on words of what can be termed mid-frequency, i.e. words with a log frequency of -4 in the LANCHART corpus. The words are: 'givet, livet, løbet, opdraget, oplevet, revet, skibet, skrevet, sproget, toget' [ˈg̊iːwəð ˈliːˀwəð ˈløːwəð ʌb̥ˈd̥ʁɑːˀwəð ˈʁæːwəð ˈsg̊iːˀwəð ˈsgʁæːwəð ˈsb̥ʁɔːˀwəð ˈtˢɔːˀwəð] *given, the life, have run, raised, experienced, torn, the ship, written, the language, the train.* Note that two of these words may be realized with [b̥] rather than [w] in distinct pronunciation: 'skibet' and 'løbet' may both be realized with [b̥] instead of [w] which would mean that their phonological representation would not contain the variable investigated here. However, these words were never relaized with [b̥] in the LANCHART Cph subcorpus, and therefore tokens with [w] and tokens where there is no segment between the vowel and the syllabic [ð̩] were included as instances of the variable.

11.8 Word frequency and individuals – the older generation

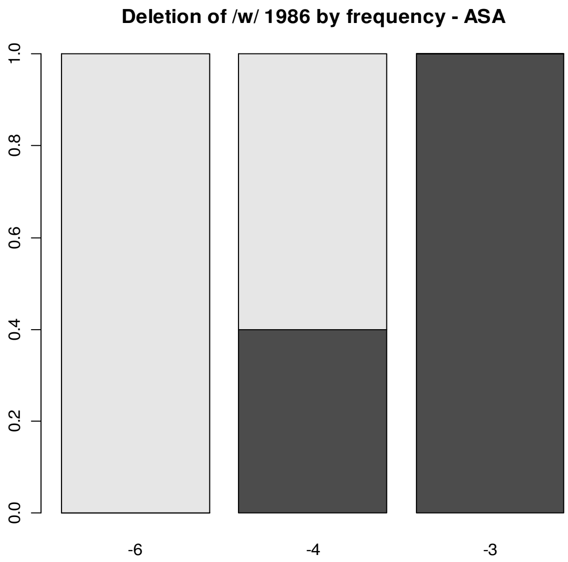
The speakers in the older generation can be split into three sub-groups on the basis of their tendencies for deletion of [w] in the mid-frequency words. The first group consists of two speakers who show a diminished tendency for [w] deletion in comparison with the group average. Their tendencies for deletion by word frequency are shown for each speaker in the bar plots below.

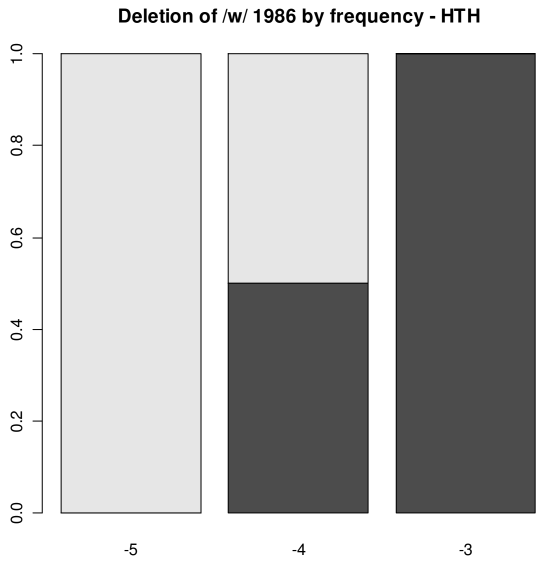


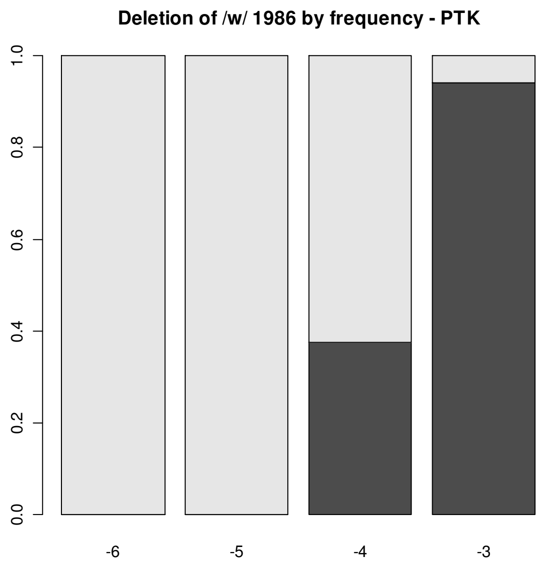
The most conservative speaker, i.e. the informant least likely to delete [w] before [ð̩], is LAL, a female speaker who only deletes [w] in the high frequency word 'blevet'. CNI, a male speaker of the older generation, deletes [w] in all high frequency words, but shows a reduced tendency for [w] deletion in mid-frequency words. In this frequency category, the group average is a deletion rate of 50 %, but CNI only deletes [w] in 2 of the 10 tokens he produces of words with a logfrequency of -4, namely in his productions of 'givet' [ˈg̊iːwəð] and 'oplevet' [ˈʌb̥ˈleːˀwəð]. The instances of deletion produced by CNI cannot be attributed to the stress of the syllable. 9 out of 10 of the tokens of *high* frequency words, in which CNI always deletes [w], are unstressed, which might indicate that stress rather than word frequency is the factor that constrains [w] deletion before [ð̩] for this speaker. If so, the expectation would be that the other tokens in which he deletes underlying [w] would mostly be unstressed, and that the majority of unstressed syllables containing a token of the variable would exhibit deletion. This is not the case: the majority of tokens of mid-frequency words produced by CNI, 7 out of 10, are unstressed, but only in one of these 7 unstressed syllables is [w] deleted, and that is in the word 'oplevet' – the other token with [w] deletion is a stressed realization of the word 'givet'. While it is true overall that a token in which [w] is deleted is more likely to be unstressed than stressed, it is not true that unstressed tokens are likely to have their underlying [w] deleted. However, a clear pattern for CNI is that [w] deletion is more likely than [w] retention in high frequency tokens. Thus the two most conservative speakers can be seen to pattern according to the word frequency hypothesis: when they do delete [w] before [ð̩] they do so in high frequency words.

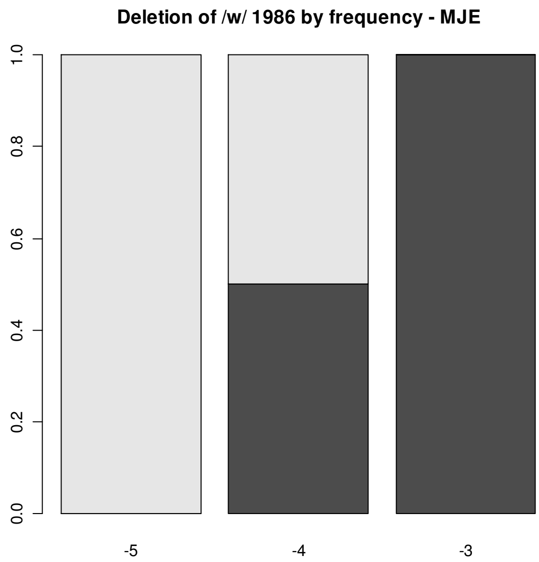
11.8.1 The average speakers

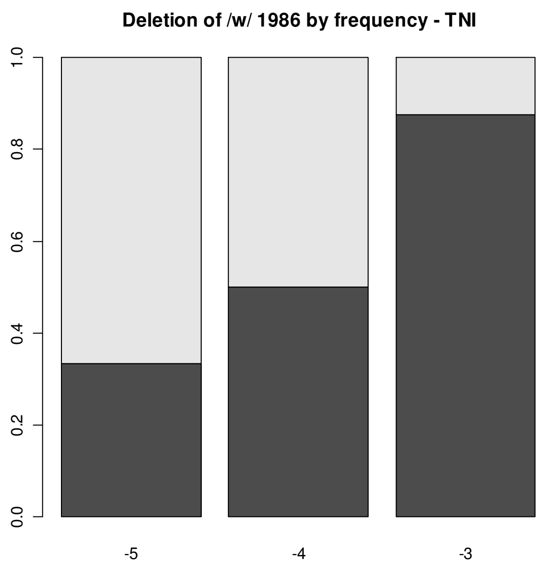
The second subgroup of speakers in the older generation show patterns close to the group average with respect to propensity for deletion in words of different frequencies. There are 5 speakers in this group, 3 of which categorically delete [w] in high frequency words. They all show a tendency for deletion in mid-frequency words that is close to the group average, i.e. they delete [w] in 40 % to 50 % of tokens with a log frequency of -4. Interestingly, only the male speaker TNI produces tokens with deleted [w] of the lower mid-frequency category -5. Apparently, he is responsible for all of the tokens of deleted [w] in this frequency category observed for the group as a whole. The distribution of tokens with [w] deletion is given for each of the 5 speakers in the bar plots below in order of increasing tendency for [w] deletion in mid-frequency word forms. For each speaker, the possibility that stress is the factor that induces [w] deletion is evaluated using Fisher's exact test.











For ASA there is no sign that stress is the underlying factor, since she in fact has more [w] deletion in stressed tokens than in unstressed tokens. She does not always delete [w] in stressed tokens, however, and the likelihood for deletion is not greater for stressed tokens according to the significance test (p = 1!).

For HTH we can observe an overwhelming tendency for [w] deletion in unstressed syllables:

of the 13 unstressed tokens she produces, [w] is deleted in 12 of them. But [w] is also deleted in all of the 6 stressed tokens of high frequency words, indicating again that frequency and not stress is the important factor. As for CNI, a token with [w] deletion is more likely to be unstressed, but HTH is equally likely to delete [w] in stressed and unstressed syllables ($p = 1$!).

PTK produces a larger number of tokens than the rest of the "average" group, a total of 38, in 22 of which [w] is deleted. 18 of these 22 tokens are unstressed, indicating again that a realization without [w] is more likely to also be unstressed than stressed, but unstressed tokens are not more likely to be realized with out [w] than stressed tokens, at least not to a statistically significant degree ($p = 0.69$).

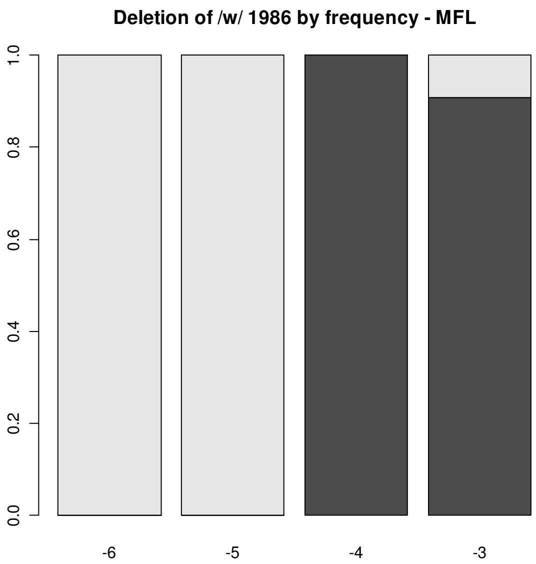
For MJE the pattern is similar to that of ASA. In fact, in all 5 of his stressed tokens the [w] is deleted, as it is in 8 of the 10 unstressed tokens he produces. So again, a token with [w] deletion is more likely to be unstressed, but although the raw numbers suggest that deletion is more likely in stressed than in unstressed tokens, the significance test reveals that this may be due to chance since $p = 0.59$.

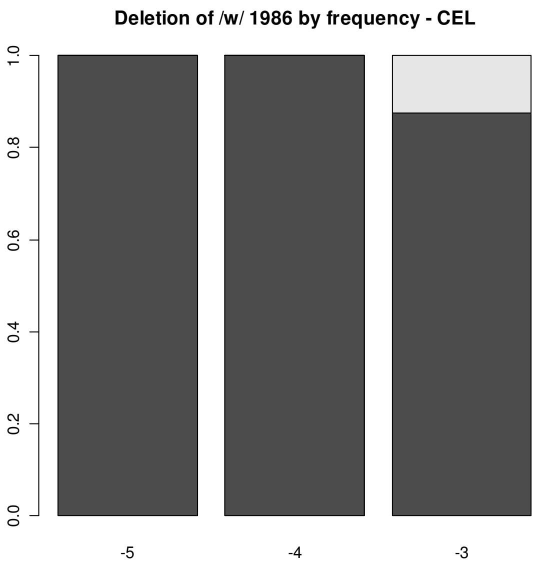
The last speaker in the average group is also the one who shows the greatest range of word frequencies in which he deletes [w]: TNI actually deletes [w] in words with a logfrequency of -5. And in fact he is also the only speaker who shows evidence of the phonetic factor stress underlying his pattern of deletion. He produces a total of 33 tokens, in 25 of which [w] is deleted. 22 of these 25 tokens are unstressed, meaning that a token with [w] deletion is much more likely to be unstressed than stressed. And deletion is also more likely than retention in unstressed syllables according to the test for significance, since $p = 0.04$. This indicates that TNI has a general tendency for deleting [w] in syllables without stress. However, a Fisher's exact test on the subset of *un*stressed tokens produced by TNI shows a significant difference between the tendencies for deletion of [w] for the three levels of frequency, ($p = 0.01$). This shows that even when the lack of stress is held constant, log frequency of word forms still matters, and hence that TNI is more likely to delete [w] before [ð̩] in high frequency words than in words of relatively lower frequency.

As with the more conservative speakers, the general pattern that emerges from the group of "average" speakers is that the tendency for [w] deletion patterns with word frequency even for individual speakers: [w] deletion before [ð̩] is more likely to happen in words of high frequency. The fact that this pattern can be observed for individuals supports the hypothesis that frequency of use is incorporated into individual mental representations of phonological forms of words

11.8.2 The atypical speakers

The last subgroup of speakers in the older generation consists of only two male informants. Their behavior with respect to [w] deletion before [ð̞] is atypical in the sense that they actually have the greatest tendency for [w] deletion in mid-frequency words of all of the individuals in the older generation, but they also retain [w] in high frequency items. This can be seen in the barplots below.





MFL only differs from the "average" group in that he has obligatory deletion of [w] in the frequency category -4, whereas CEL is the most atypical member of the group of older speakers because he has almost obligatory deletion of [w] irrespective of word frequency. But, what is of course most surprising about these two speakers is that they do not always delete [w] in high frequency words despite their proclivity for deleting [w] in words of lower frequency. This is not what is to be expected if the frequency effect were a property of individual phonologies. It should be noted, however, that for both speakers exactly one high frequency token retains [w]. For MFL it is 'lavet' and for CEL it is 'blevet'. MFL does not always retain [w] in 'lavet' (which would indicate a specific lexical effect): he produces 3 more tokens of the word in all of which the [w] is deleted. The realization of [w] in this one token cannot be due to stress either, although 'lavet' is a word that is likely to be realized both stressed and unstressed, due to the possibility for the word to enter into unit accentuation (Rischel 1983?). Although the token of 'lavet' where MFL retains [w] is stressed, so are 2 of the tokens in which he deletes the underlying [w] in 'lavet'. Furthermore, the difference between the tendency for deletion in stressed and unstressed syllables is not significant for MFL. Clearly, the retention of [w] in this token must be explained by other factors than those explored in the present investigation.

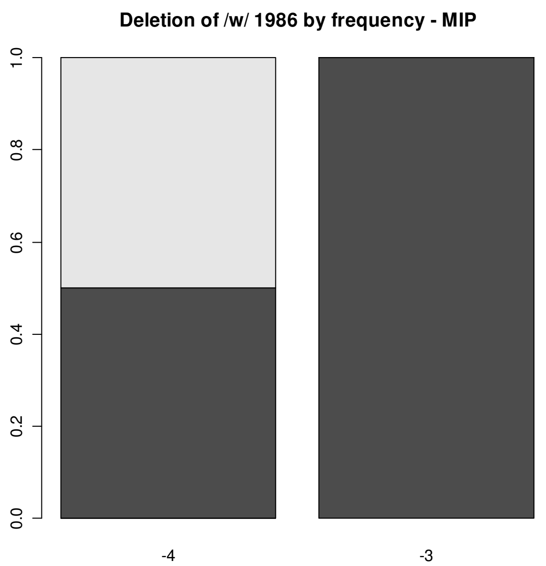
A similar pattern holds for CEL, who retains [w] in one token of the word 'blevet' but deletes it in 3 other tokens of the same word. All of these tokens are unstressed, so again stress or prominenece cannot explain the realization of [w] in this particular token. As for MFL, this one token of retained [w] must be explained by factors not explored here.

Taken together, the individuals in the older generation show that the frequency effect on [w] deletion before [ð̞] found for the group as a whole is by and large reproduced by the individuals, supporting the hypothesis that the frequency effect is a function of the structure of the mental lexicon. It should be recalled that this is not the only possible result. It could have been the case that some individuals only ever deleted [w] in words of low frequency or at least that they had a significantly higher tendency for [w] deletion in low frequency words than in high frequency words. But this was not seen for any one of the 9 individual members of the group of older speakers who produced more than 10 tokens of the variable (wəð). What *was* found, however, was that for two of the speakers it was possible to realize [w] in a high frequency item even when most or all words of lower frequency were affected by [w] deletion in the speech sample for those speakers. While the evidence is sparse it does suggest that the frequency effect is not an automatic process that will always apply in a particular fashion. While the mechanism proposed for exemplar based models of the lexicon does allow for the occasional selection of a relatively low frequency form in the cloud of exemplars representing a particular word, it cannot handle the occurrence of such tokens without appealing to something other than frequency of use of the word. It is beyond the scope of the present investigation to delve further into the reasons behind non-reduction in high frequency words that show an overall strong tendency for deletion both in the speech community and in the speech of the individual producing the non-reduced token. But it can be stated that the mechanism does not appear to be one of lexical exception or generalization of aspects of the phonetic environment, as shown in the analysis presented above.

11.9 Word frequency and individuals in apparent time – younger generation

The analysis of the relationship between [w] deletion before [ð̞] and word frequency in individual speakers must of course also be carried out for the younger generation. This is done both with the purpose of investigating to what extent the frequency effect can be traced for individuals, and thereby extending the findings from the analysis of individuals in the group of older speakers, and in order to further investigate the observed difference between the two generations, namely the increased probability for [w] deletion in word forms of log frequency -4 in the younger generation as compared to the older generation. As with the older group of speakers, the individuals in the younger group are presented here in order of increasing likelihood for [w] deletion in mid-frequency words. Again the analyses will be mainly qualitative due to the small number of tokens produced by each speaker, and is restricted to speakers who produced 10 or more tokens of the variable (wəð).

11.9.1 The conservative speaker

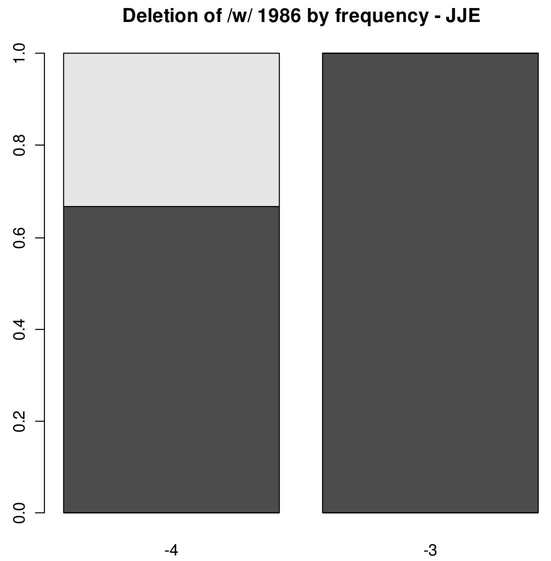


Like the individuals in the older generation, the group of speakers in the younger generation may be divided into 3 separate sub-groups on the basis of their rate of deletion of [w] before [ð̞] in words with a frequency of -4. The first "group" consists of just one person, the male speaker MIP, whose tendency for [w] deletion by frequency is shown in the barplot on the left.

Like all of the speakers in the younger group, he deletes [w] in *all* high frequency word forms. Unlike the rest of the speakers, he only deletes [w] in half of the tokens of words forms of mid-frequency that he produces, meaning that he most closely resembles the speakers of the older group in particular speakers from the "average" group HTH and MJE. MIP produces a total of 18 tokens of the variable, with only 2 of them containing a [w] in the realization. No phonetic factors can be seen to account for the retention of [w] in these two tokens: one is stressed, the other is unstressed, and MIP is equally likely to delete [w] in stressed and unstressed syllables (p = 1!). In all tokens the underlying [w] is preceded by a front vowel, making it impossible to investigate the role of place of articulation, and since [w] is retained in both the word 'livet' and the word 'revet' there is no indication of an effect of vowel height either. The only discernible factor, then, is word form frequency, and MIP is clearly less likely to delete [w] in mid-frequency words than the remaining members of the younger group, as can be seen from the analyses of their behvaior presented below.

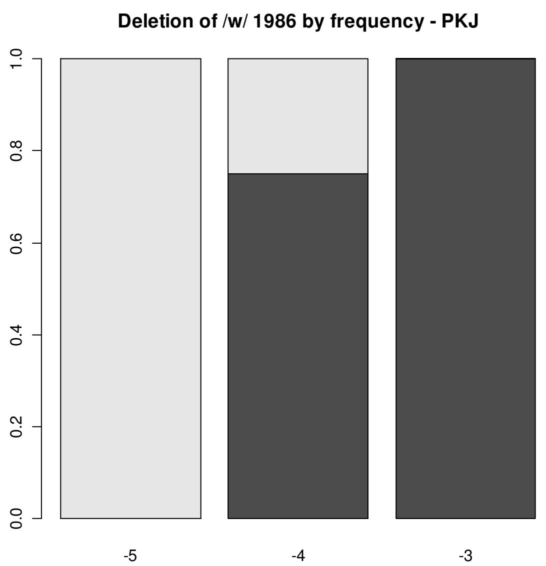
11.9.2 The average speakers

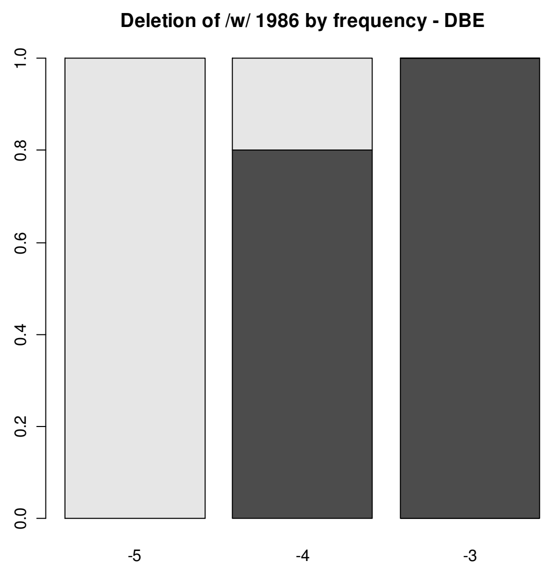
The subgroup that comes closest to the group average in their rate of deletion of [w] in words of log frequency -4 consists of 3 speakers. The association between [w] deletion and word frequency is shown for each of them in the barplots below. Interestingly, while they most closely resemble the group average for words with a log frequency of -4, none of them can be seen to delete [w] in words of frequency -5, although the group as a whole did show a slight tendency for deleting [w] in these words of relatively low frequency. For one of them, this is because she does not produce any words with a frequency lower than -4, but the two others simply do not delete [w] in words of lower frequency.



For JJE, the speaker who only produces tokens of words with frequencies of -3 and -4, there appears to be a possible effect of stress, in the sense that she deletes [w] in all of the 9 unstressed tokens she produces, but only in 1 of her 3 stressed tokens, and this distribution is significantly different from chance according to a Fisher's exact test ($p = 0.04$). Furthermore, the one stressed token in which she does delete [w] is of frequency -4, i.e. the lowest category produced by this speaker, as are the two tokens in which she retains [w]. Since 3 of the unstressed tokens are also of this relatively lower frequency category, and [w] is deleted in all of them, it seems probable that stress is a more important factor than word frequency for JJE.

For PKJ and DBE the situation is different: both are equally likely to delete [w] in unstressed and stressed tokens according to the statistical test ($p = 1$! for PKJ, and $p = 0.59$ for DBE). Actually, PKJ deletes [w] in every stressed token he produces as

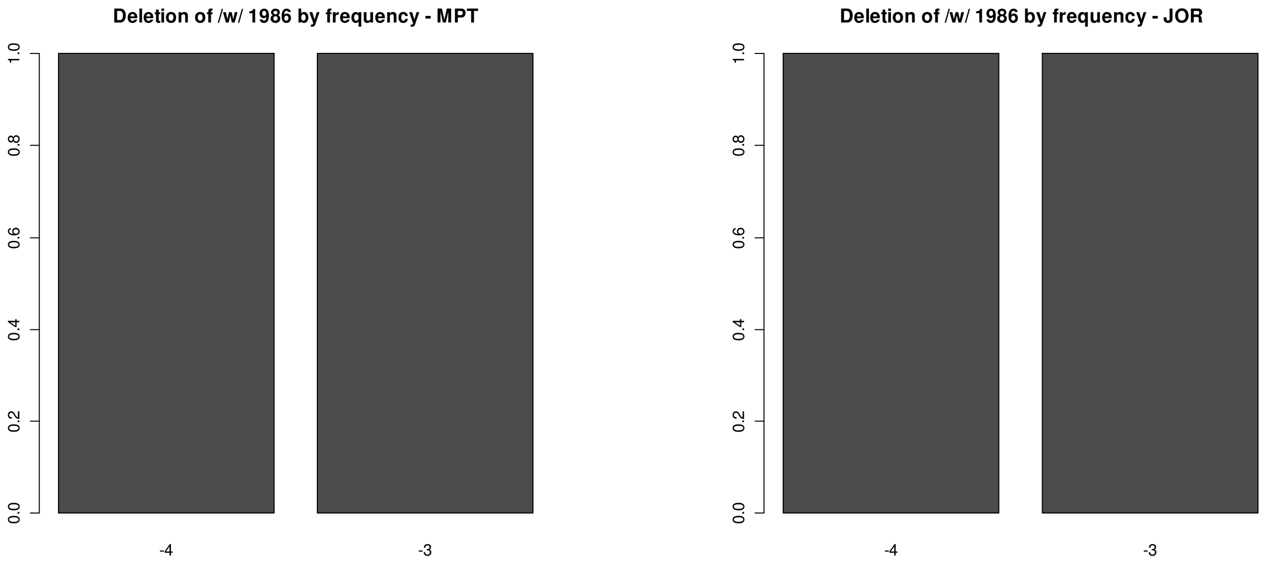




well as in the majority of his unstressed tokens, indicating that frequency is the significant factor for him (which is supported by the statistical test of word log frequency where $p = 0.02$). DBE does not delete [w] obligatorily in neither stressed nor unstressed tokens, and also has a significant effect of word log frequency ($p = 0.009$). For these two speakers, then, the apparent effect of word frequency on their i ndividual tendencies for [w] deletion can be said to be real, thereby further supporting the evidence for the word frequency effect on phonetic reduction at the level of the individual speaker.

11.9.3 Categorical deleters

The two remaining speakers in the younger group who produce more than 10 tokens of the /wəð/ variable constitute the last sub-group in the younger generation. Nothing can in fact be said about the role of word frequency in their probability for deletion of [w] since they obligatorily delete [w] in the sample, as can be seen from the barplots below.



11.10 DISCUSSION OF WORD FREQUENCY EFFECTS IN INDIVIDUAL SPEAKERS

The initial analysis at the group level showed that [w] deletion continues to increase in apparent time and that the increase is correlated with word frequency: younger speakers have a higher rate of deletion in high frequency word forms, indicating an influence of word usage on the spread or transmission of this particular reductive sound change. However, no evidence was found for the change to spread to word forms of lower frequency in the younger generation, but this may be due mainly to limitations of the data: none of the speakers in the younger group produced any words with a frequency lower than -5.

The frequency effect is generally found also to apply to the distribution of [w] deletion for individual speakers: importantly, there are no individuals who only delete [w] in low frequency word forms and not in high frequency ones, but some speakers may realize [w] as [w] in a few tokens of high frequency word forms, showing that frequency of use cannot be an automatic trigger of [w] deletion before [ð̞].

For two other speakers, an effect of stress was found to either override or interact with the effect of word frequency. Both of these speakers showed an overwhelming tendency for [w] deletion in unstressed tokens. For one of them, [w] deletion was obligatory in unstressed syllables, whereas for the other speaker [w] was realized as [w] in a few unstressed tokens, mainly in low frequency words. These speakers indicate that while phonetic reduction may be regulated by word use, speakers may eventually come to generalize some phonetic property of the high frequency words and begin to extend the process according to phonetic criteria – in some cases still regulated by word use, and in other cases simply generalizing the process to all environments that meet the phonetic criteria.

However, the data are quite sparse, making statistical analysis of individual speakers meaningless in many cases. The results for individual speakers should therefore in most cases be interpreted as tendencies. But the tendencies observed do suggest that further investigation of a more frequently occurring variable is warranted.

If the association between word frequency and deletion rate seen here is reproduced, this would support models of the mental lexicon which incorporate frequency of use in lexical representation for individual speakers.

12. Real time study of [w] deletion

The comparison of probabilities for deletion of [w] before [ð̩] in the LANCHART Cph interviews recorded in the 1980s revealed an effect of age of the speaker measured as decade of birth. The probability for deletion of [w] increased with decade of birth, meaning that over all, the younger the speaker, the more likely it is that [w] will be deleted, even when the phonetic factors stress and preceding vowel quality are controlled for. In other words, the 1980s data show evidence of a change in progress, as expected from the analyses of Brink & Lund (1975).

We may now ask whether this tendency has increased in the intervening 20 years: has deletion of [w] become more common? If so, is this true for all speakers, or mainly for the younger generation? Finally, are the restrictions that were found for [w] deletion in the apparent time study still observed by speakers? Or has the process become generalized to other contexts?

As a first step, a model is fit to the entire dataset, to see if there is any evidence of an increase of [w] deletion. The fixed factors perceived stress, lexical stress, quality of the preceding vowel, decade of birth, gender and time of recording are included in a mixed effects model with subject as random factor. Recording is the only new factor, and it is of course added in order to test whether there is a statistically significant tendency for increased [w] deletion in the new recordings, when all other factors that were seen to influence this process in the original recordings, are controlled for. The model characteristics, given in table 12.1 below, support the hypothesis of a real time change.

Table 12.1 – Overall model for time of recording

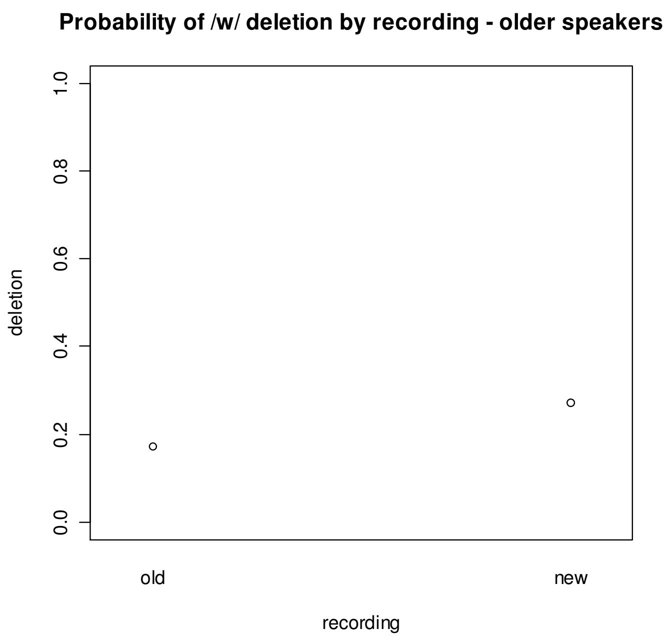
Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	-4.97481	1.10654	6.93e-06 ***
unstressed	1.24061	0.26557	2.99e-06 ***
secondary	-1.65148	0.34359	1.54e-06 ***
front	2.80853	0.35764	4.06e-15 ***
decade	0.06054	0.01842	0.001010 **
rec new	0.84356	0.24035	0.000449 ***

All factors except gender emerge as significant, even perceived stress, which was only found to be significant for the older generation in the data from the 1980s. What is of particular interest here, of course, is that the time of recording also emerges as significant, even when all other factors are controlled for. This indicates a real time change, and the table of coefficients shows that it is in the expected direction: new recordings have a positive coefficient, meaning that the probability for [w] being deleted is higher in the new recordings compared to the old ones.

This is to be expected for the younger group, since some of them were adolescents at the time of original recording (Chambers (2002), pp. 359-360)). To study whether there is an effect for both generations, this same model is fit to two subsets of the data: the set of tokens produced by speakers from the older generation, and the set of tokens produced by speakers of the younger generation. Table 12.2 below gives the coefficients for the group of older speakers for all factors which emerge as significant.

Table 12.2 - Older speakers in real time			
Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	-4.20122	2.10342	0.0458 *
unstressed	1.236	0.2874	1.69e-05 ***
secondary	-1.847	0.3740	7.89e-07 ***
front	2.447	0.3898	3.47e-10 ***
rec. new	0.58316	0.2544	0.0219 ***

The time of recording remains significant for the subset of older speakers and has a positive coefficient, meaning that in this group, across speakers, there is an increase in [w] deletion. The partial effect of recording is illustrated in the graph below.



As is apparent, the effect is quite small, which can also be seen from the size of the estimate for recording in table 12.2. The effects of lexical stress and the place of articulation of the preceding vowel have the same effect as observed in the original sample: [w] deletion is less likely in syllables with secondary stress than in syllables with primary stress, and it is more likely after front vowels than after back vowels. Perceived stress also remains significant here – unstressed syllables are more likely to be affected by [w] deletion than stressed syllables. Before looking at the younger group of speakers, I present a comparison of the factors influencing [w] deletion in the group of older speakers at the time of the original recording with the factors that influence the process in the new recordings. This will show to what degree factors have taken on new roles over time. The effects of each of the factors are shown for each of the two sets of recordings of the older generation in tables 12.3 and 12.4 below. This comparison shows whether any of the constraints on [w] deletion that were found in the original recordings are violated in the new recordings.

Table 12.3 - Older generation – 1980s			
Factor	Estimate	Std. Err	Pr(>|z|)
(Intercept)	-10.79	3.20	0.002 **
Unstressed	1.29	0.46	0.005 **
secondary	-2.22	0.56	<0.0001 ***
front	3.35	0.86	<0.0001 ***
decade	0.18	0.06	0.01 *

Table 12.4 - Older generation – 2000s			
Factor	Estimate	Std. Err	Pr(>|z|)
(Intercept)	3.06	3.53	0.39
unstressed	1.41	0.41	0.0006 ***
secondary	-1.52	0.56	0.0064 **
front	2.19	0.47	2e-06 ***
decade	-	-	n.s.

Examining the two tables of coefficients for the older generation, it first of all becomes clear that whereas there was a difference in probability for [w] deletion in the 1980s depending on decade of birth, this factor disappears in the model of the new recordings. This suggests that the group has become more homogeneous with respect to [w] deletion. Whether this increased homogeneity is due to an increase in [w] deletion for those speakers who showed less of a tendency to participate in the process in the 1980s sample of the older generation, or whether it is due to a convergence towards a more stable level of [w] deletion will be explored further in the analysis of individual speaker behavior below. For now it is possible to conclude that the older group as a whole appears to have participated in the real-time change of increased [w] deletion, and while clearly some speakers must have changed their behavior more relative to others, thereby creating increased homogeneity within the group, the linguistic factors of perceived stress, lexical stress, and preceding vowel quality still influence the process of [w] deletion. That is, the process may have increased in the population as a whole, as the initial model fit to the entire dataset indicates, but the linguistic factors involved in regulating the application of the process are still observed by speakers in the older generation. The real time change seen for this group, then, does not indicate a change in the phonology of the speech community, but a change in usage.

We may now ask whether the younger speakers, as a group, also increase their tendency to delete [w] before [ð̞] and if so, whether they have extended the process to all kinds of environments, meaning that no linguistic factors should be seen to be statistically significant in modeling the process. Or have they come to resemble the older generation more in the 20 years intervening between the two recordings: does the application of [w] deletion become more restricted or do the speakers continue to delete [w] more in contexts which can be seen to disfavor [w] deletion in the older generation? First of all, do the younger speakers also show evidence of real time change?

Table 12.5 - Younger speakers in real time			
Factor	Estimate	Std. Error	Pr(>lzl)
(Intercept)	2.45118	4.06705	0.5467
unstressed	0.93515	0.51179	0.0677 .
secondary	-1.40981	0.67011	0.0354 *
front	3.05572	0.60127	3.73e-07 ***
rec. new	1.26773	0.45470	0.0053 **

Table 12.5 shows the factors that are statistically significant when the model for real time change is fit to the subset of the data pertaining to the younger generation only. As was the case with the older generation, the factor 'recording' is significant, and the coefficient is positive, meaning that the tendency for [w] deletion increases over time for the younger generation as well. The partial effect of recording is illustrated in the graph. The difference between the old and the new recording is relatively larger for the younger generation than for the older generation (cf. above).

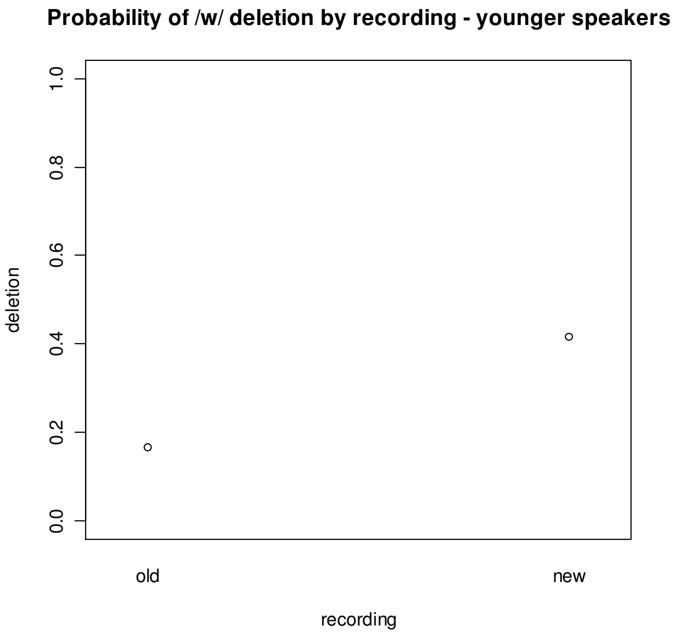


Table 12.6 below gives the characteristics of the model fit to the subset of tokens produced in the new recordings by speakers from the younger generation. All factors that were found to be significant for the older generation in the new recording are presented in the table regardless of their level of statistical significance for the modeling of the behavior of the younger speakers in the new recordings.

Table 12.6 - Younger generation – 2000s			
Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	4.65	4.16	0.26
unstressed	0.72	0.71	0.31
secondary	-1.17	0.93	0.20
front	3.01	0.70	1e-05 ***

As table 12.6 shows, the only significant factor is the quality of the preceding vowel, indicating that the effects of stress that were observed for the younger generation in the recordings from the mid 1980s have now disappeared, i.e. the group of younger speakers have generalized [w] deletion to syllables with secondary stress, and the tendency for inhibition of [w] deletion in syllables with stress is no longer statistically significant at all. The younger group continues to show a higher tendency for [w] deletion than the older group, and they have not adopted any of the constraints imposed by stress in the previous generation, but have instead abolished them altogether. This indicates that the higher tendency for [w] deletion in this younger generation observed already in the recordings from the 1980s should not be interpreted as a sign of age-grading, but rather could be seen as the continuation of a change in progress, a process that some of the speakers have continued to participate in well into adulthood. Taking the modeling of the two generations together, tells us that the continued increase in [w] deletion before [ð̞] may be interpreted as a case of community change in the Sankoff (2006) typology of processes of change.

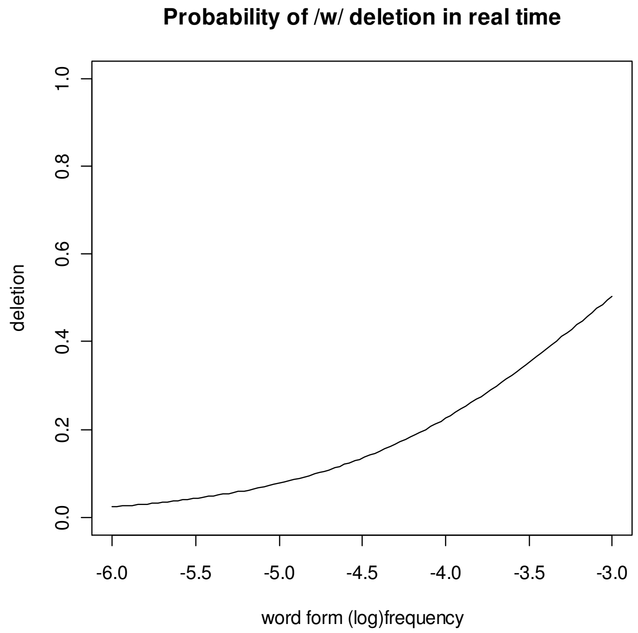
Before looking more closely at the behavior of individuals and the changes to individual phonologies, I want to examine the relationship between [w] deletion and word frequency at the group level.

12.1 Real-time change and the effect of word frequency

So far, I have looked at phonetic and extra-linguistic factors in the studies of [w] as a change in progress. However, as mentioned in the introduction, the purpose of studying change here is to see whether the word frequency effects found in the work of Joan Bybee and summarized in Bybee (2002), an effect that forms the basis of recent models of lexical representation (e.g. Pierrehumbert (2002) and Johnson (1997)), can be replicated in real time change over a relatively short time span (compared to the sound changes studied in Bybee (2002)). I begin by investigating the overall effect of word frequency on [w] deletion viewed as a real-time change in progress. This is done by adding the factor word form log frequency to each of the models presented above.

Adding word frequency to the model and fitting it to the entire dataset removes the effect of lexical stress – it remains as a tendency with $p = 0.07$. This is not a big surprise. The frequency measure is based on word forms, meaning that simple words and compounds receive separate frequency measures. Since secondary lexical stress by definition is a property of compounds, lexical stress and frequency should be correlated to some extent[23]. What is of interest here, is that the factor word frequency is significant – and has an appreciable estimate. Given that the model attempts to "predict" [w] deletion while controlling for time, this can be interpreted as an overall effect of word form frequency on the real time increase in [w] deletion that was found earlier. The coefficient is positive, indicating that [w] deletion increases with word frequency, as is also shown in the graph illustrating the partial effect of word frequency for the entire dataset.

[23] There are no risks of collinearity in the model: all vifs < 2

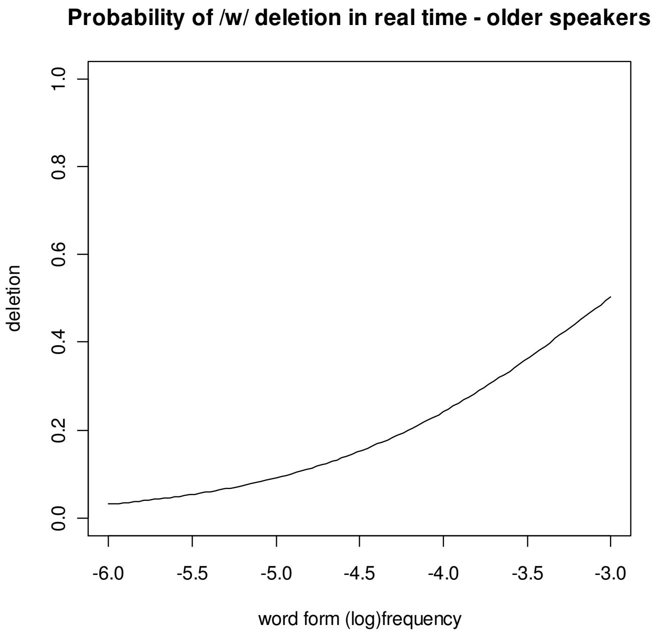


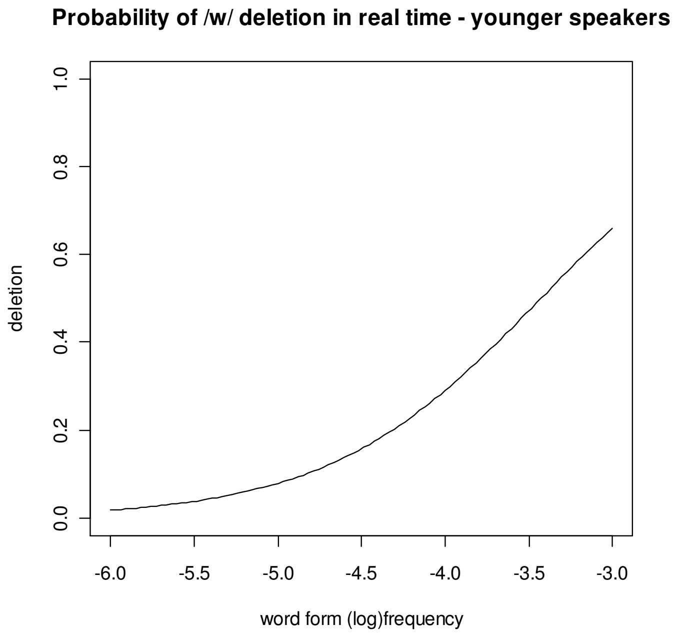
The next question to be answered is whether word frequency matters for both groups of speakers? Since both older and younger speakers show an increase of [w] deletion in real time, we would expect the frequency to obtain for both groups. However, we also saw that the younger group of speakers showed signs of generalizing the process beyond the contextual constraints observed by the older generation. Might this mean that frequency is no longer relevant in the structuring of their tendency to delete [w] in running speech? Tables 12.7 and 12.8 below give the coefficients for the two generations separately. All factors that were significant in the model of older speakers are given in both, but, as can be seen, not all of them remain significant when word form frequency is added.

Table 12.7 - Older speakers – 2000s frequency effect			
Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	1.44	2.38	0.54
unstressed	0.86	0.30	0.0046 **
secondary	-0.77	0.41	0.0589 .
front	1.70	0.41	4e-05 ***
decade	-	-	n.s.
male	-	-	n.s.
new	0.60	0.27	0.0270 *
logfreq	1.15	0.19	1e-09 ***

Table 12.8 - Younger speakers – 2000s frequency effect			
Factor	Estimate	Std. Error	Pr(>|z|)
(Intercept)	9.17	3.52	0.0092 **
unstressed	-	-	n.s.
secondary	-	-	n.s.
front	1.94	0.63	0.0021 **
decade	-	-	n.s
male	-	-	n.s.
new	1.37	0.47	0.0037 **
logfreq	1.56	0.29	1e-07 ***

The coefficients show that frequency is significant for the modeling of real time change for both generations, but stronger for the younger group than for the older group (log odds 1.56 and 1.15 respectively, see further below). Further, we see that the effect of lexical stress is only just significant for the older generation. For the younger generation, the quality of the preceding vowel is the only linguistic factor that is significant when word form frequency is introduced: [w] is more likely to delete after front vowels, meaning that the constraint that a preceding back vowel will tend to inhibit [w] deletion still holds. The same constraint is true for the older generation, who also still show an effect of stress, exactly as before, with unstressed syllables being slightly more likely to be affected by [w] deletion. Graphs illustrating the partial effect of word form frequency for the two generations are given below.





The plots of the partial effect for word form (frequency show the same general trend for both generations. However, they also show that the younger generation has a greater tendency for [w] deletion in high frequency words than the older generation does. So far then, the frequency effect in this data is exactly as predicted by Bybee (2002) and others: high frequency words are more likely to be affected by deletion, and this pattern also holds for real time change. Furthermore, younger speakers show a stronger tendency for [w] deletion in words of relatively high frequency, whereas there appears to be no difference between the two generations in their propensity for [w] deletion when the analysis is restricted to word forms of a log frequency of -4 or less. This difference between the two generations suggests that the effect of word form frequency is also operative in the transmission of an on-going sound change from one generation to the next. The speakers in the younger generation in the LANCHART Cph Corpus are clearly continuing the change of [w] deletion before [ð̞], and they are doing so primarily by regularizing the process in high frequency words.

12.2 *Word frequency and life span changers*

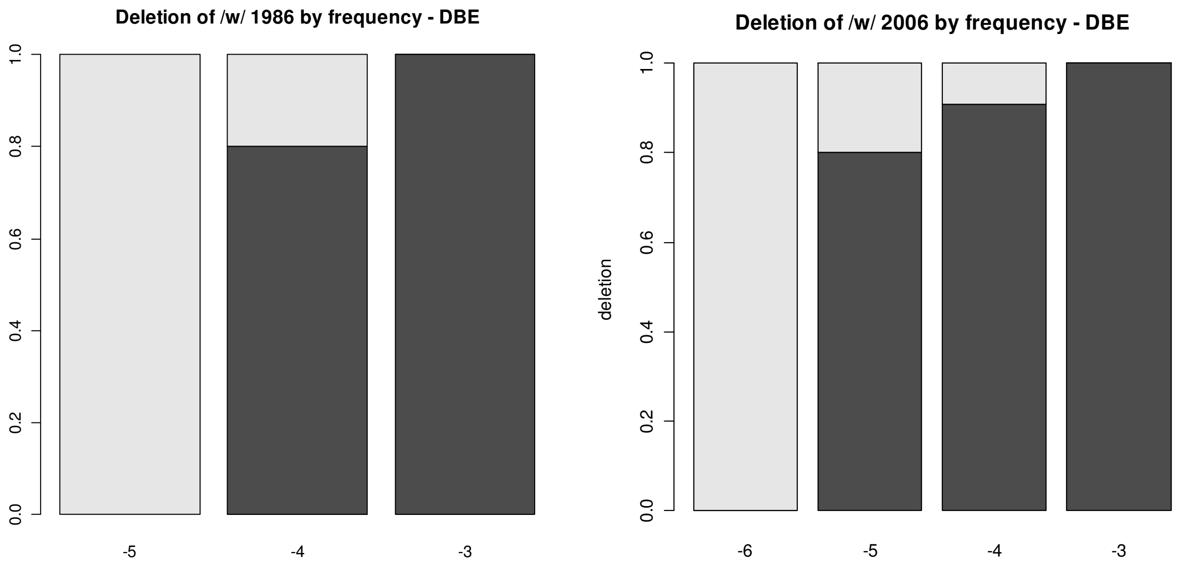
So far I have examined the continued increase in the deletion of [w] before [ð̩] at the group level, by comparing the pattern of phonetic and/or phonological constraints on the process in the two generations sampled, both contrasting the two groups with each other, thereby establishing a diminishing effect of stress on the process, and by comparing each group to itself in the re-recordings, thereby establishing a tendency for a continuation of the process. However, an important aspect of the real time panel study is the possibility for investigating change within the individual across the life span in order to test the assumption of stability of the phonology of the vernacular throughout adulthood, which forms the premise of the apparent time paradigm. In this section, then, I will address the question of how many of the speakers in the original sample that show evidence of changing their proportion of deleted [w] before [ð̩] from the original recording in the mid-1980s to the re-recordings in the early 2000s? In order to test this, the distribution of deleted and realized [w] was calculated for each of the 22 subjects, and the distributions of the variants in each of the two recordings were compared and tested for significant differences. As already mentioned, the data are quite sparse at the level of individuals, with some of the individual speakers producing a very low number of tokens in the original recordings. This means that Fisher's exact test had to be used when comparing each speaker to him- or herself. Doing so for all of the speakers reveals that only 3 of the 22 speakers can be shown to change in real time, or in the terminology of Sankoff (2006), 3 of the speakers are life span changers. The remaining speakers show some fluctuations in their distributions across the two recordings – some have a slight increase in rate of deletion of [w], others have a slight decrease – but the 3 life span changers all show a significant increase in [w] deletion from the original to the new recording. The distribution of variants together with proportion of deletions and the result of the significance tests are given for each of the 3 speakers in the table below.

Table 12.9 – distribution of variants of [w] before [ð̩] for the life span changers

speaker	n realized 1980s	n deleted 1980s	% deleted 1980s	n realized 2000s	n deleted 2000s	% deleted 2000s	p	Age group	Gender
DBE	5	12	70 %	3	38	92 %	0.04	Younger	Female
LAL	8	3	27 %	0	10	100 %	0.001	Older	Female
PTK	16	22	58 %	4	29	88 %	0.007	Older	Male

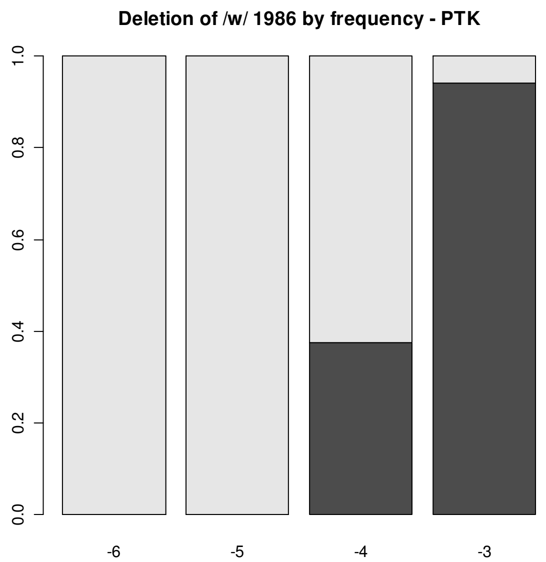
The table shows that the life span changers were not necessarily adolescents at the time of the original recordings, supporting the findings from other studies that even adult phonologies do not necessarily remain stable in adulthood. The most striking change is found for the female speaker LAL from the older generation, who goes from being a reluctant deleter of [w] in the original recording to becoming a speaker who always deletes [w] before [ð̞]. LAL was in fact the speaker who was seen to be most conservative in the analysis of individuals in apparent time: all of her deletions occurred in high frequency words only in the original recording.

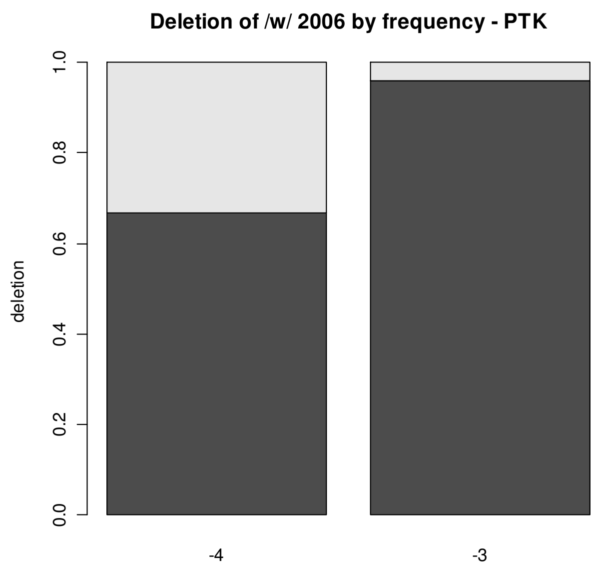
It obviously does not make sense to examine the relationship between word frequency and spread of the sound change in the individual lexicon for a speaker who becomes a categorical user of one of the two variants. We simply cannot know how the sound change may have progressed for LAL in the 20 years intervening between the two recordings. For DBE and PTK however, we can ask whether life span change progresses according to frequency, just as was seen to be the case in apparent time change. The tendency for deletion by word frequency is given below for both speakers.



DBE shows a clear association between the spread of [w] deletion and word frequency from the original to the second recording: she never deleted [w] in words with a log frequency of -5 in the original recordings whereas now she deletes [w] in 80 % of such tokens.

For PTK, the picture is not so clear as can be seen in the bar plots overleaf. PTK produced words at all levels of frequency that pertain to the /wəð/ variable in the recording from 1986, and he showed a clear tendency for deleting [w] in words of relatively high frequency while never deleting [w] in words of low frequency. As described in the section on apparent time change, this pattern could not be explained by any of the phonetic factors included in the analysis. In the re-recording, however, PTK only produces words of relatively high frequency, i.e. the ones in which the data from the first recording had already shown that he was likely to delete [w].





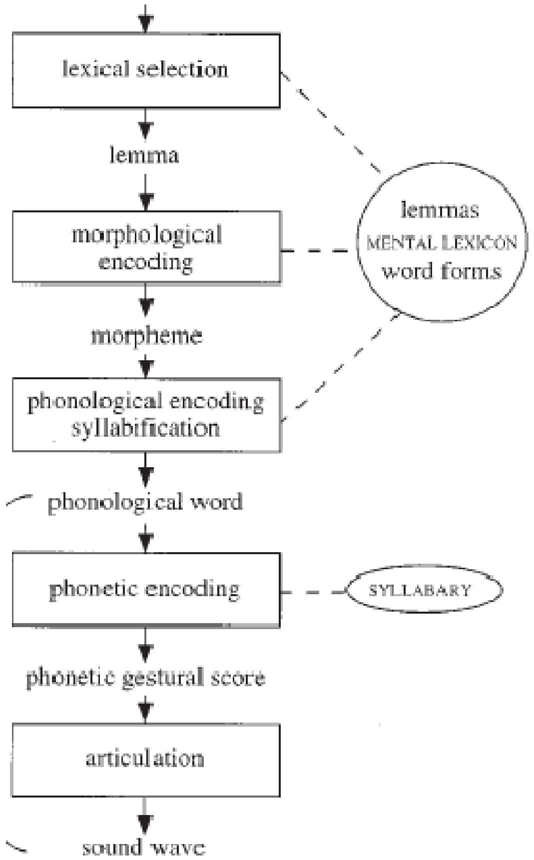
All that can be said about PTK's statistically significant increase in deletion of [w] across the life span with respect to word frequency is that he has increased (significantly) his rate of deletion in words of the second highest frequency. We cannot know whether he has also extended the process to occur in words of lower frequencies.

Taken together, the two informants who can be seen to extend the process of [w] deletion according to frequency indicate that the spread of a sound change through the lifespan may also be related to frequency. Clearly however, analysis of a more frequently occurring variable is needed to further substantiate this finding.

13. Reduction and representation

13.1 Phonetic reduction in a single entry model of the mental lexicon

In a single entry model of the representation of lexical forms, or lexemes, like the "Blue-print for the Speaker" as presented in Levelt (1989) and Levelt et al. (1999) it is assumed that each lexeme has one and only one representation in the long-term memory store of words, the mental lexicon. The model is represented schematically in figure 13.1 on the left, a facsimile of figure 1 from Levelt et al (1999).



The lexemes go through a number of operations before being realized as overt speech. The effects found in the present investigation take place at the level of the construction of phonetic plans for connected speech, the module labeled 'phonetic encoding' in the figure. Prior to the application of reduction processes, lexemes have been retrieved from the lexicon and have been concatenated into complex word forms and utterances. The pertinent phonological processes have changed the phonological shape of the lexemes, such that positional allophony has been derived, and prosodic phenomena like stress and final lengthening have been ascribed according to the rules of the language. The details of stress assignment involve a procedure of activating metrical frames on the basis of on-line syllabification of the phonological word (Levelt et al (1999), pp. 19 – 23), and I will return to the specifics of this process and how it relates to the reduction of consonants below. Taking these contextual conditions into account leads to activation of the segments or phonemes that make up the phonological words (with specific mechanisms for the sequential ordering of the segments, which are not, however, relevant to the discussion here). This representation is passed on to a procedure of phonetic spell out in the module labeled "phonetic encoding" in figure 13.1, where the activation of the phonemes, which were activated when the lexeme was retrieved, spreads to the syllables stored in the mental syllabary. These syllables are conceptualized as comprising the articulatory commands needed to produce the (allo)phones in the syllables (Levelt (1989), pp. 324 – 329). This leads to a phonetic plan for the word or words to be uttered, represented in citation form, or what I have called distinct phones. A second step in the phonetic spell out procedure involves taking into account contextual factors above the level of the syllable, e.g. regressive place assimilation across syllable boundaries. The prosodic structure assigned to the planned utterance further influences parameterization of the motor commands in the syllables retrieved as does the rate of speech (Levelt (1989), pp. 406-410, and Levelt et al. (1999), p. 23). It is at this stage in the speech production process that reduction phenomena like the ones investigated here can arise in the single entry model. They are, in other words, interpreted as the result of contextual conditioning on the stored motor programs for articulation of syllables. Until this point in the process, all information that has been retrieved in order to convey the intended content which leads to the selection of lemmas and subsequent retrieval of lexemes and further activation of the constituent phonemes is abstract information. Even the output from the phonological encoding is represented in terms of abstract units, which are explicitly invariant and discrete. In the phonetic encoding, information about the temporal overlap of the gestures required to turn the phonetic plan into overt speech is activated. In other words, implementing the phonetic plan requires the use of procedural knowledge. In the model of Levelt et al (1999), the phonetic plans consist of specifications of particular tasks to be accomplished rather than exact specification of movements that ensure completion of the task. For example, the phonetic plan for an unvoiced bilabial plosive only specifies that closure at the lips should be achieved, as far as the oral gestures are concerned, not that both the lower lip and the jaw have to be moved upwards. This means that the phonetic plan for [b̥] 'spids' [sb̥es] *point* and 'oppe' [ˈʌb̥ə] *up* can be identical, with the implementation procedure being adjusted according to the segmental context – in this case it is probable that the two implementations will require different specifications for the movement of the jaw, since lowering of the jaw is only required for the preceding vowel in 'oppe', not (necessarily) for the preceding fricative in 'spids'. The model proposed in Levelt et al (1999) has little to say about the mechanisms involved in the procedure from phonetic spell out to articulation, nor does it fully specify the contents of the mental syllabary. Here, I will explore how the suggestions made in Levelt et al (1999) and Levelt (1989) may be reconciled with the results obtained in the present investigation of consonant reduction in contemporary Copenhagen Danish. It should be noted that the model of Levelt et al. (1999) is specifically intended to handle the production of a single phonological word. I will therefore discuss how the model might be extended in order to also deal with larger utterances.

As already mentioned, the approach taken in single entry models of lexical representation is compatible with the phonetic conditions on consonant reduction found in the present investigation, when these conditions are considered as isolated effects applying to a given sequence of activated syllable sized motor programs. Effects of prosodic factors like the increased tendency for reduction of consonants in unstressed syllables, syllable position-effects and segmental factors like an increased tendency for target undershoot of plosives in intervocalic environments and increased deletion rates in inter-consonantal environments can be handled by making the parameterization of the implementation of the phonetic plans an on-line probabilistic procedure in which the contextual factors act in concert to produce the specific token. In fact, the results found here show that the implementation procedure must be probabilistic, since reduction of consonants may occur in any context, although it is more likely to occur in some contexts than in others. Conversely, while the intervocalic position of [g̊] in words like 'ligger' *lie (pres.)* and 'stykke' *piece* may favor reduction to [ɣ], such that the words are likely to be realized as [ˈleɣɐ] and [ˈsd̥øɣə] respectively, the process does not always apply in this context in running speech. Since the single entry model does not allow for variability at higher levels in the speech production process, these probabilistic rules must apply at the level of implementation of the phonetic plan. This is precisely because the model only allows one invariant representation of each lexeme. Regular, across-the-board transformations, such as the transformation of morphophonological |k| to /g/ in syllable final position, can occur at the level of phonological encoding, meaning that the forms retrieved from the lexicon may be represented by morphophonemes which are then transformed to phonemes on the basis of position in the structural frame during phonological encoding. One of the major criticisms of single entry models has been the inability for the models to account for the probabilistic nature of reduction in connected speech. As stated in Pierrehumbert (2002, p. 102) "[i]n modular, feed-forward models [of speech production NP], the (categorical) form of the lexeme wholly determines the phonetic outcome." However, since the single entry model does involve on-line operations on the string of segments that a phonological word is comprised of, it should in principle also be possible to incorporate the effects found in the present investigation. In the model of Levelt et al. (1999), metrical frames are assigned during phonological encoding, unless the retrieved lexeme has a stress pattern which deviates from prototypical pattern. In Danish, stress is largely predictable from the segmental structure of the phonological word (cf. Grønnum (2005), pp. 245-246). Stress is always assigned to the syllable with a full vowel in words where only one full vowel occurs. In words with more than one full vowel, stress is assigned to the first long full vowel in the word. Even if there are no long full vowels in the word, stress is still largely predictable from the segmental structure, since the last full vowel followed by a consonant will be assigned stress, implying that if the last syllable is closed, this syllable will be assigned primary stress. There are several more or less regular exceptions to this last pattern. Some can be made with reference to the type of consonant which closes the final syllable, and others must be made with reference to the origin of the word. For example, words which end in /n/ and only contains short full vowels like ‘turban’ *turban* are always stressed on the penultimate syllable, i.e. [ˈtˢuɐ̯b̥an] except for French loan words like ‘balkon’ *balcony*, since these always have final stress, i.e. [b̥alˈkʰʌŋ]. Since the model of Levelt et al (1999) deals with completely predictable rules for the assignment of stress on the basis of segmental structure, it would probably be necessary to store the metrical specification with all words which contain only short full vowels directly in the lexicon, while allowing stress to be assigned in phonological encoding for all other types of words. Similarly, the syllable accent stød is also largely predictable from the metrical and segmental structure of the phonological word, and even the absence of stød in heavy syllables which have been assigned more than null stress (and are thus eligible for stød) is predictable on the basis of morphological structure (as in the model for the interplay of prosody and word structure developed in Basbøll (2005)).

Thus, information about syllable level prosody will be available once the lexeme has undergone phonological encoding, and is therefore also available in the spell out of the phonetic plan. The final specifications for stress at the level of utterances should allow some degree of look-ahead in order to avoid assignment of stress to verbs which enter into unit accentuation. Recall for example that the verb ‘drejer’ [ˈd̥ʁɑjɐ] *turn (pres.)* is often realized without stress in the DanPASS Dialogues. This can be accounted for by letting the words following the verb be available during stress assignment, since the loss of stress happens by default in syntactic phrases where a verb is followed by a preposition (Rischel (1983)). For example, in the utterance ‘så drejer du til højre efter vandfaldet’ *then you turn right after the waterfall* the lack of stress on the word ‘drejer’ is predictable from the following prepositional phrase ‘til højre’ *to the right* even when the subject of the verb, ‘du’ *you* intervenes. Whether this happens at the level of phonological encoding or phonetic spell out is immaterial for the present purposes. All that is necessary is that the specification for stress is available during implementation of the phonetic plan.

Information about the segmental context is of course also available during the spell out of the phonetic plan, since the entire phonological string of segments has been passed on from phonological encoding. The gradient rules of consonant reduction would thus have all the information necessary to allow the application of the processes of reduction and deletion in phonetic implementation. The rules could be in the manner of the variable rule format first proposed in Labov (1969) and further developed in Labov & Sankoff (1979) and Fasold (1991) among others, with the modification that these are rules of implementation and not rules of phonology proper. This would allow us to retain a model of the mental lexicon in which lexemes are represented by a single invariant form, and letting the variability observed in running speech be handled by the phonetic implementation module. Probabilities for the application of processes could thus be computed on the basis of corpus studies like the present investigation. However, the implementation module would have to contain information about more than the phonetic conditions under which a particular phoneme is to be realized. There are in particular three aspects which have been uncovered in the present study that necessitates this. First of all, some words are more likely to be affected by reduction processes than other words. In some cases, this is because the words consistently provide a phonetic environment which favors reduction, such as the aforementioned reduction of target [g̊] in words like 'ligger, stykke'. But for the processes of (ɑj) monophthongization and word final (d) deletion, the probability for application of the process was seen to be affected by word form frequency. A similar effect was found for the apparent time and real time studies of the on-going sound change of deletion of [w] before [ð̞], where the process was seen to be more advanced in high frequency words than in low frequency words, both when comparing the two generations and when examining the distribution of deleted [w] for individual speakers. Furthermore, the effect of frequency persisted in the real time comparison of both groups and individuals. Without information about the frequency of the word form, information about the process of realization of these variables and the development of the process over time would be lost. In addition, the variable (g) showed an effect of the gender of the speaker, with male speakers being more likely to reduce a token of (g) than female speakers (cf. the section "Variation of (g)" in chapter 6). And, finally, there were rather substantial differences between speakers in the groups studied in the LANCHART Cph Corpus, indicating that probabilities based on the aggregate behavior of groups is also inadequate for capturing the details of the speech production process of the individual speakers.

13.2 External factors in a single entry model

It is possible to incorporate some of the effects into the traditional single entry model of speech production. Since word form frequency is related to ease and hence speed of retrieval of lexemes, with high frequency lexemes being retrieved faster than low frequency lexemes, a fact that is also observed by Levelt et al (1999) (p. 18). Thus, the speed of retrieval could be incorporated as a weighting factor in the implementation of the phonetic plan, if this information is retained from the process of retrieval through phonological encoding and phonetic spell out to the implementation of the plan, much in the same way as information about prosodic factors are passed on from phonological encoding. However, if the probabilities are based only on the effects which can be observed at the level of the speech community, in effect modeling the behavior of an "average" speaker of the particular variety studied, it becomes difficult to see how differences between groups in a speech community may arise, both with respect to apparently stable synchronic variation, as in the case of the effect of gender on the reduction of (g), or between different generations. While the speech production model of Levelt et al (1999) is not explicitly designed to handle changes in progress, it would seem a welcome addition to such models, if they incorporated mechanisms which can motivate the changes in the speech production process observed across generations. Furthermore, even though word form frequency might be incorporated into the single entry model, this would entail that every application of a reductive process is qualitatively the same. As discussed in chapter 6, the large differences in rate of reduction for particular words suggest that while the phonetic outcome is the same, it might be more fruitful to consider for example the near-categorical monophthongization in ‘jeg’ [jɑj] *I* as a process that is different from the occasional monophthongization in ‘indianerlejren’ [endiˈæːˀnɐlɑjˀʁɐn] *the Indian camp*. In the next section I turn to a description of the type of models which are able to handle such usage-based effects and to more directly capture the differences between lexemes.

13.3 Variability in the lexicon – exemplar based models of speech production

On the basis of a growing body of research which shows structured variability in the phonetic realization of individual words both in synchronic samples and in studies of the spread of sound changes, a new type of models of lexical representation and speech production have been proposed in recent years, namely the so-called usage based or exemplar based models of speech processing. In the following I will outline some of the principles of these models and argue that they are better suited for handling the results observed in the present investigation of consonant reduction. I will also argue that exemplar based models should retain the levels of abstraction assumed in single entry models in order to adequately capture the patterns of variation observed.

I will conclude by suggesting further work in the study of phonetic reduction in contemporary Copenhagen Danish which may contribute to a fuller understanding of the processes than can be gained on the basis of the present investigation.

13.4 Exemplar models

Exemplar based models of lexical representation are radically different from the single entry models of speech production in the way in which lexical items are represented in long term memory. Rather than operating with a limited number of invariant lexemes that have had all predictable phonetic detail abstracted away, exemplar models posit a vast store of all encountered tokens of a given unit in long term memory. The memories of these individual tokens are the exemplars. The approach has its roots in psychology, where work on perception has shown that episodic memory can be taken to be the basic form of representation of incoming sensory stimuli in long term memory (cf. Johnson (2007)), and categories may emerge from abstractions over the exemplar space, but the detailed percepts are retained in long term memory. The relevance for such a detailed and hence heavily redundant store of remembered tokens of words and segments in speech processing come from a series of experiments on the role of individual voices in the recognition of spoken words presented in Goldinger (1996). Here it was found that listeners recognize words faster if they are presented by a voice that is familiar to the listener than if the word is presented in a voice that is unfamiliar to the listener. Familiarity was induced by having listeners participate in two separate word identification tasks. In the first experiment, all listeners were required to respond to words presented by voices that were all unfamiliar to them. In the next experiment, the original group of listeners was divided into two sub-groups: one group heard words produced by a new set of unfamiliar voices, and another group heard words presented by a mix of voices, some of which were the same voices as those they had heard in the first word recognition experiment. The results showed that identification of words was faster in the second experiment for the group that responded to stimuli produced by voices that they had heard in the first experiment. This suggests that details of the individual voices had been retained in memory in the period intervening between the two experiments, leading Goldinger (1996) to posit that such details are a part of the long term memories of words.

In another series of experiments, presented in Goldinger & Azuma (2004), the effect of individual voices was found to spill over into the speech production system. A group of subjects first participated in a word naming task, in which they said aloud words presented on a screen. The sessions were recorded and used as stimuli for a subsequent experiment, described below. The same group of speakers participated in a word recognition experiment the following day, in which they were presented with the same words they had said aloud in the naming task the previous day, but in this second experiment the words were presented to them in different, unfamiliar voices. Five days later the same group of subjects participated in another word naming experiment, similar to the first experiment, that is the same words used in the two previous experiments were again said aloud by the subjects. This session was also recorded. The production of words from the first experiment and the production of words from the third experiment were then used as stimuli in a similarity judgment task performed by a new group of subjects. A set of AXB stimuli were constructed from the three sets of recordings. The X was always a token used in the word recognition experiment with the first group of subjects, whereas the A and B items were tokens of the same word taken either from the very first recordings of the first group, i.e. tokens produced before the speakers had been presented with the voice saying the X stimulus, and the other stimulus was the token produced by the speaker in the third experiment, i.e. a token produced five days after the speaker had heard the word produced by the voice speaking the X stimulus. The group of listeners who participated in this AXB experiments were asked to rate whether the A or B stimulus was most similar to the X stimulus. Listeners consistently judged the token produced in the third experiment as the best imitation of the X stimulus, indicating that the tokens the first group of subjects had heard in the word recognition experiment had affected their own productions of these words. Interestingly, the effect of "goodness of imitation" was higher for low frequency words which had been used several times in the word identification experiment. The effect of word frequency suggests that words which are encountered rarely in everyday speech are most heavily affected by increased exposure to the word, whereas frequently encountered words are more robust in their representations, since as far as the second group of listeners was concerned, these realizations did not constitute as good an imitation of the X stimulus as realizations of the low frequency items did.

Taken together, these experiments suggest that detailed memories of lexical items are retained in long term memory, memories which include voice specific details, and that hearing words affects subsequent productions of the words. The fact that high frequency and low frequency words are affected to different degrees is explained by taking into account that the words naturally existed in speaker memory before participation in the series of experiments. High frequency items will be represented by a large number of many different tokens, whereas low frequency items will be represented by a relatively low number of encountered tokens. When listeners are exposed to new items, high frequency words are affected less than low frequency words, because the relative weight of the novel tokens will be larger in a distribution with few tokens than in distributions with a large number of tokens.

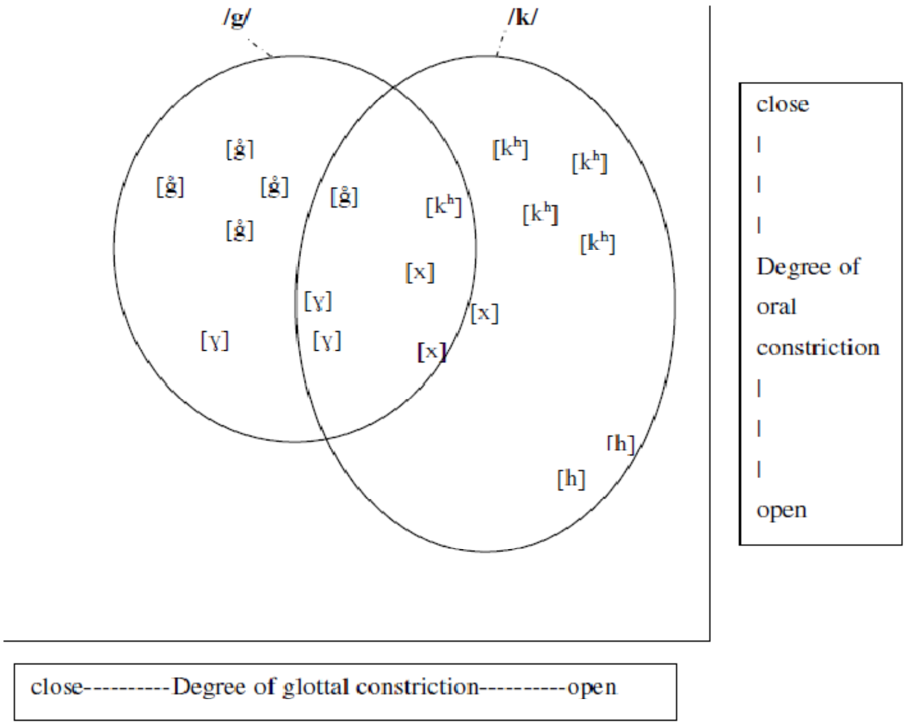
The evidence that voice specific detail facilitates word recognition and may influence subsequent word production suggests that the stored representations consist of richly detailed memories of encountered speech. Thus, an exemplar based model of the lexicon allows for direct representation of the variability observed in running speech, without the need for specific rules which strip away redundant information in perception and add predictable information to stored abstract representations in production. Instead, word specific biases for particular non-contrastive variants of phonemes can follow directly from the association between the word form and specific parts of the exemplar space.

Naturally, this requires that abstraction does occur. In radical versions of exemplar based models, categorization of exemplars is done only through similarity matching of incoming tokens and no privileged role is given to the categories that may emerge from the distribution of tokens in exemplar space. Such radical models fail to account for the fact that speakers are able to produce spoken forms of words they have only encountered in writing, and hence also the ability to provide a spoken form for nonce words and to judge the degree to which spoken nonce forms are good examples of possible words in their native language, as pointed out in Pierrehumbert (2002). This is because novel forms and nonce forms cannot be represented in a model which relies exclusively on the storing of encountered forms in speech, since by definition novel and nonce forms have not been encountered previously. By allowing abstraction over the exemplar space through the use of labels that assign exemplars to specific categories at different levels of abstraction, e.g. phonemes, morphophonemes, and word forms (as suggested in Pierrehumbert (2001) and (2002)), it becomes possible to account for the fact that spoken forms of novel and nonce words can be produced: a phonological shape can be assembled by the use of the category labels attached to the exemplar space, and an overt speech form can be produced by selecting among the clouds of exemplars associated with each label. Clearly, then, abstraction over the linguistic input with reference to structural descriptions is still necessary, and in the following I will restrict the discussion to such hybrid models of representation, i.e. models where abstract categorization is a ubiquitous aspect of the representation of words and segments in memory alongside the storing of concrete, episodic memories.

Although exemplar based models were developed to handle perception and categorization over perceptual stimuli, the results presented in Goldinger & Azuma (2004) show that detailed representations also affect speech production and the exemplar framework has been extended to handle speech production, notably in Pierrehumbert (2002) and (2006), but also Plug (2005). The exemplar space is posited to be the same for perceived and produced tokens, but where the process in perception involves the categorization of incoming stimuli, production consists of selection of tokens to be produced.

In the exemplar based production models, the exemplar space is structured by an association of word form labels and phoneme labels with the stored exemplars. Each unit thus has a particular distribution of variants associated with it, and the most frequent variants form the center or peak of the distribution with less frequent variants being located at the periphery. For example, at the level of phonemes, Danish /g/ will be associated with a range of variants, which may be represented by the variants attested in the DanPASS Dialogues corpus, i.e. [g̊ ɣ x kʰ j ç] and nil. The vast majority of tokens will consist of [g̊] with the other variants being located at the periphery of the distribution. The phoneme /g/ shares parts of the exemplar space with the phoneme /k/, which may also be realized with the same variants as /g/ in addition to the variants [h ɦ], but the centers of the distributions for the two phonemes will be located at different points in the exemplar space. This is illustrated schematically below.

Figure 13.2 – A schematic representation of the association between phonemes and variants in an exemplar based model



With this basic structure, the explanation that /g/ is most often realized as [g̊] and /k/ is most often realized as [k^{h}] can be made by simply positing a random selection over the exemplar space associated with the phoneme, since the two variants will always be the most likely candidate for an articulatory target.

However, as has been shown, the two phonemes are not always realized by the most typical variant. This is handled in the exemplar model by introducing biases in the selection procedure. These biases may arise through a phonological decomposition of the word form labels that are also associated with the exemplar space. The word form labels are categories in the same sense that phonemes are categories. Word forms can emerge as categories because the identification of incoming speech streams is eventually connected to the meanings associated with the words, and hence the variable acoustic input comes to be associated with word forms. Note that this also allows for morphophonemes to arise as abstract categories, since stems and derivatives are linked by their meaning, and hence the alternations observed in stem final position in Danish, e.g. between [g̊] and [k^{h}] in word pairs like ‘lak’;’lakere’ [lɑg̊];[lɑˈkʰeːˀɐ] *varnish; to varnish,* may come to be associated and emerge as connections in the mental lexicon. Since the variables studied here do not allow a full analysis at the morphophonological level, I will not discuss the representation of morphophonological alternations in an exemplar based model further here – for a proposal, see Bybee (2002, p. 117-118).

In the following section, I will outline how biases based on the phonetic conditions found to be relevant in the processes of consonant reduction found in contemporary Copenhagen Danish may guide the selection procedure.

13.5 PHONETIC FACTORS IN EXEMPLAR MODELS

The phonetic factors that influence the tendency for producing variants that are reduced relative to a distinct pronunciation can be incorporated as biases or weights in the production of a specific message. Up to this point, the single entry model and the exemplar model are very similar in the conceptualization of the speech production process. An intended message is assumed to activate labels associated with the form representation of that message. These labels can then be assembled in a phonological buffer, and implementation of the targets for production can proceed from here. In a single entry model, the factors can be seen as weights in the on-line production of a specific implementation, and the weights then adjust the duration of the motor gestures stored in the mental syllabary. In an exemplar model, the factors can be conceptualized as biases on the selection procedure in establishing an articulatory goal for the particular utterance.

In an exemplar model without any representation of the influence from contextual factors, selection of articulatory goals is simply a random selection from the exemplar cloud associated with the category that has been activated. The selection procedure is assumed to be random, since otherwise there would be no variation in speech production, which is clearly at odds with the observed empirical facts. However, as found in the present study, speech variation is not entirely random, but probabilistically regular in the sense that contextual factors may increase or decrease the likelihood that a given variable will be produced differently from its target in distinct pronunciation. In a model of speech production that incorporates phonological representations as labels used in the structuring of the exemplar space, the contextual factors are incorporated as biases guiding the selection procedure. When a word like ‘ligger’ [ˈleg̊ɐ] *lie (pres.)* is selected for production, the phonological category /g/ is activated during the selection procedure. Initially, the system is simply biased towards selecting randomly from the set of exemplars of /g/, and since the majority of exemplars in this category resemble the variant [g̊], it is most likely that the instantiation will include a [g̊] as the intervocalic plosive. By introducing a bias based on the segmental environment it becomes possible to guide the selection procedure, such that an articulatory goal is selected from the part of the exemplar set where [ɣ] is more frequent than [g̊]. This bias can arise through the simultaneous association of this part of the exemplar space with labels at both the phonological level and the word form level. The exemplars in this part of the mental map of speech sounds are thus an aggregate of the tokens of /g/ occurring in word forms where the /g/ is intervocalic. But the word form label can only provide this information if phonological decomposition of word forms is a part of the model. If we were to assume instead that the only relevant level for categorizing tokens, and thus structuring the exemplar space, was the word form level, we should expect to find that all effects were word specific. To the extent that a quantitative study like the one presented here adequately represents the variation found in the speech of middle class speakers of contemporary Copenhagen Danish, we should have found that the likelihood for producing reduced variants would be influenced only by factors belonging to the word form level, i.e. specific tendencies for particular words and effects of word frequency. However, this is not the case, since contextual factors like segmental environment, stress, stød and speech rate often emerged as significant in the statistical analyses, and more so than word form frequency. Incorporating phonological decomposition as part of the process allows the exemplar model to capture these regularities. It furthermore provides the possibility for accounting for how processes which are phonetically motivated, such as target undershoot of plosives in intervocalic contexts, may extend to all words containing the triggering context. In this way, the phonetic motivation for the process can be linked to its application in each individual word.

Thus, the process of speech production in exemplar based models also contain a phonological buffer, like the single entry model of Levelt et al (1999), in which the word form is analyzed and where prosody is assigned in the same way as it was done in the single entry model. In other words, the selection procedure is not guided exclusively by the word form, but by the structural and metrical description arrived at in the phonological buffer. Under the assumption that the statistical procedure used to model the effects of phonetic extra-linguistic factors approximates the generalizations made in the cognitive process of selection of articulatory goals, the models developed for each variable can be said to reveal at least part of the information used to guide speech production.

Why should these effects be tied to lexical representation? If word forms are analyzed into their phonological segments, the model might as well use structural descriptions as labels over the exemplar space. In other words, why should the variability be incorporated into the lexicon instead of formalizing them as gradient or variable rules that apply in phonological encoding and phonetic spell out? I would argue that precisely the three aspects which are difficult to handle in a single entry model are the ones which support the use of exemplar based models in the account for consonant reduction. First, for two of the variables, (ɑj) and (d), an effect of word form frequency was found for the probability of reduction relative to the distinct form., and a frequency effect was found for the on-going process of [w] deletion before [ð̞] both in apparent time and real time. Second, some words were found to exhibit much higher reduction rates than others. Third, differences between groups of speakers were found in the analyses of the likelihood for reduction.

13.6 FREQUENCY EFFECTS IN EXEMPLAR MODELS

The effect of word form frequency on synchronic reduction and in on-going sound change is a classic argument in favor of exemplar models and usage-based phonology (cf. Bybee (2002) and Pierrehumbert (2001)). In traditional single entry models of lexical representation, there is no way of motivating why frequent words are affected more often by a reduction process than low frequency words. All the words contain the phonological characteristics that trigger the process, and therefore should all be adjusted during phonological encoding prior to the construction of the phonetic plan for the words. Similarly, when a synchronic state of variation becomes an on-going sound change, traditional single entry models would predict that it would spread through the lexicon in an across-the-board fashion, i.e. that all sound changes should be Neogrammarian in their path of diffusion: since all word forms contain the triggering environment, all word forms should be equally affected by the process. In so far as the methods employed for investigating the process of [w] deletion before [ð̞] adequately capture the diffusion of the process through the lexicon, this uniform path of diffusion is not supported by the data. Instead, diffusion of the process is correlated with word form frequency, both at the group level, and at the level of individual speakers who continue to participate in an on-going sound change during adulthood (although admittedly the evidence for the effect at the level of individual speakers is limited).

As discussed above, the frequency effect on synchronic variation might be incorporated into a single entry model, to the extent that word form frequency can be reliably derived from the speed of retrieval of lexemes. However, this cannot account for the fact that probability for reduction increases in high frequency items over time, and that the process also spreads to words of low frequency. Presumably, the relative frequency of occurrence of the words do not change over time, nor are there large differences in the frequency of individual words for individual speakers, at least not for very highly frequent words. In other words, the ease of access as a function of frequency should not change over time, nor should it be different for subsequent generations of speakers. Yet this is exactly what was found in the study of [w] deletion: younger speakers showed an increased probability for deletion compared to older speakers, and deletion increased in high frequency words over time and spread to words of lower frequency.

This effect of word form frequency can be explained in an exemplar model of lexical representation: highly frequent words are affected more often than rare words, because they are more often available for the process of [w] deletion to apply. Since each application of the process is retained in memory and associated with the word form in which it was heard, the representation for the word form is incrementally updated, with more and more reduced tokens being collected in the exemplar space associated with the word form, and thereby gradually increasing the likelihood that speakers will produce the words with the reduced variants. Because the effect is connected to properties of word forms, it is desirable to conceive of it is a property affecting lexical representations. If it were instead to be a property of the inventory of segments used to implement word forms, the process should only be sensitive to phonological factors like stød, stress and segmental environment, or articulatory factors like articulation rate. The word form frequency effects on the spread of a sound change like the one found in the present study of contemporary Copenhagen Danish show the benefit of incorporating phonetic variability into the representation of word forms. It may also be used to explain the difference between generations. When a particular phonetic process begins, it is of course very rare in the vocabulary of any speaker. As the process progresses, it becomes more and more common, particularly in high frequency words. This also means that speakers in a younger generation will have a relatively larger set of reduced tokens of high frequency words in memory as they acquire the speech production patterns of the surrounding speech community, than the speakers from the older generation did during speech acquisition. Hence, the process will be more strongly associated with the high frequency word forms for speakers in the younger generation.

13.7 Differential rates of reduction for particular words

The existence of high rates of reduction for particular words may be seen as a special case of word form specific representation. Recall from the discussion of monophthongization of (ɑj), that some words showed near-categorical tendencies for monophthongization. I would argue that for such words the monophthongization is qualitatively different than for words in which monophthongization is relatively rare. Such a distinction is of course impossible to make in a single entry model. In an exemplar model, however, the difference between monophthongization in a word which is very often affected by the process and a word which is only occasionally affected can be captured directly in the lexical representation in exactly the same way that the effect of word form frequency can. As reduced variants of these words accumulate in memory, the association of exemplars with reference to the word form enables the word form to act as a bias in the selection of articulatory goals. In this way, the articulatory process affecting the two types of words is at some level still the same, e.g. monophthongization or deletion, but by associating the word form with a phonetic distribution, the qualitative difference arises: for a word that is frequently affected by the process, the word form label itself may act as a strong bias towards selecting from a particular part of the exemplar space in the production process. For a word form which is only occasionally realized as reduced, reduction must be the result of a phonological decomposition of the word through which characteristics of the prosodic and segmental context are obtained and used to guide the selection of an articulatory target. In this way, reduced forms can become the default articulatory target for a word form, in a sense making it more relevant to account for the non-reduced realizations that do occasionally occur. This does not mean that phonological decomposition or prosodic characteristics cannot affect the selection procedure for words with reduced default targets. This is perhaps best illustrated by the effect of word form frequency that was observed for (ɑj). For highly frequent words, like the pronouns 'jeg, dig, mig, sig' [jɑj dɑj maj sɑj] *I, you, me, oneself*, monophthongized tokens constituted the predominant realization of these words in the DanPASS Dialogues. Recall from chapter 9 that the probability for monophthongization of (ɑj) increased with word form frequency for unstressed tokens only. For stressed tokens, the rate of monophthongization was generally low except for a few, but not all, high frequency items. This can be described as an effect of a prosodic parse of the utterance in which the word form containing (ɑj) is to be produced. If the syllable containing (ɑj) is assigned stress in the phonological buffer, a general bias for selecting among the diphthongal exemplars is brought to bear on the selection procedure. This affects all of the pronouns such that they are more likely to be realized as [jɑj d̥ɑj mɑj sɑj] when they are stressed, but it affects the relatively low frequency words 'dig, mig, sig', which are never monophthongized when they are stressed, more than the very high frequency word 'jeg', which can be monophthongized when stressed, but not as often as when it is unstressed. Thus, the bias introduced by the word form itself may be mediated by prosodic factors. Furthermore, words of the same frequency as 'dig, mig, sig', i.e. with a log frequency of -3 or -4, which are not predominantly realized with monophthongized variants of (ɑj) may none the less be realized with a monophthongal variant even when stressed. However, this only happens for words where the glide in (ɑj) is intervocalic, as in 'drejer' [ˈd̥ʁɑjɐ] *turn (pres.)* and 'vejen' [ˈvɑjˀən] *the road*. Thus, while the word form label itself may contribute to the process of selection of an articulatory goal, the metrical and phonological characteristics still enter into the selection procedure and are particularly relevant when a particular word is realized differently from its most common or default form.

Clearly, not all of the effects found in the statistical analysis used here require lexical representation of detailed phonetic percepts. For most of the variables no effect of word frequency or word specific effects were found. Although reduction rates of (g) in the high frequency words 'ikke' and 'ligger' can be seen to be relatively high (cf. table 6.16 in chapter 6, where the percentage of tokens containing reduced variants of (g) is 50 % or greater in these words), this effect can be ascribed to the high probability for reduction of (g) in intervocalic position. While this effect *could* be modeled as a word specific one, it would have to include all words in which (g) occurs intervocalically, since the reduction rates for 'kirkegård', 'byggelegeplads(en)' and 'sikkert' are also relatively high. It would seem disingenuous to claim that such an effect was word specific or regulated by frequency, since this would entirely miss the structural generalization to be made over these word forms, namely that the (g) occurs in intervocalic position in all of them.

As argued above there is nothing which precludes exemplar models from incorporating such effects, as long as phonological analysis of activated word labels is part of the speech production process. In a hybrid model allowing association between abstract categories and concrete and detailed phonetic representations in the lexicon, both biases based on general phonetic properties as a triggering effect and word specific biases can be captured. Any tendency for reduction may be phonetically motivated, such as the tendency for reduction of plosives in intervocalic position observed for some of the variables in the present investigation. The inclusion of an abstract word form level in the model allows for additional bias in the selection procedure from this level. The bias can be derived from what are arguably the original conditioning factors which initiated a particular process, but may become so strongly associated with the category at the word form level, that the word form acts as a bias in the selection procedure. Since hybrid exemplar based models can handle all of the effects that are handled by traditional single entry models and in addition can also account for word frequency effects and even more word specific effects, I suggest that the data presented here argue in favor of an exemplar based approach to the modeling of representation of variability in speech production. As has already been emphasized, I find it important that the exemplar based approach includes abstraction over the speech stream, since the phenomena can clearly be seen to be structured with respect to contextual properties, be they segmental or supra-segmental. Thus, what in the framework of single entry models can be termed low-level phonetic processes still take place, and may still be described in terms of general tendencies that obtain across lexical items. What is attractive about the mode of representation in exemplar models is the emphasis that it places on the probabilistic nature of speech variation, and the focus on the emergence of structural categories from the use of language. Rather than splitting language into abstract competence and concrete performance, the connection between the two is maintained in a hybrid, exemplar based model.

13.8 Social factors in an exemplar based model

Just as words and phonemes may be used as labels to structure the exemplar space, information about language external factors may also be used in the categorization of exemplars. In fact, there is no limit to the kinds of labels that can be linked to a portion of the exemplar space. Thus it is assumed that exemplars are also associated with representations of the gender and age of the speaker who uttered a particular token and even the specific speaker. Such association obviously arises from the results found in the experiments reported in Goldinger (1996) and Golding & Azuma (2004) as described above. It also means that associations between particular phonetic variants and social categories can be directly stored in lexical representation. The association between gender and reduction of (g) found in the present study can be accounted for by having relatively more exemplars of [ɣ] and [x] associated with the category "MALE" than with the category "FEMALE" in the labeling of the exemplar space. This happens automatically, when any aspect of the context in which tokens are encountered can be used to categorize the individual token. Presumably, since men are more likely than women to produce [ɣ] or [x] for (g), any given listener will on average have more exemplars of [ɣ] and [x] stored in the category "MALE" than in the category "FEMALE". To what extent this regularity influences the production of the individual male and female speaker is of course another matter, one which has not been addressed in work within exemplar theory.

This ability to readily capture the association between particular variants and any given external trait may, however, also be seen as something of a weakness for the model. Since exemplars are detailed representations of individual encounters with a given unit and its contextual embedding, there is no way to say which aspects of the context that can be used in this categorization process, e.g. room temperature or hair color of the speaker may in principle both be used as features to form category labels which then become associated with for example [β]. This means that any spurious correlation between a phonetic variant and a particular aspect of the external world may be associated in memory, and thereby any correlation may also be explained by the theory, although the connection may not provide much insight into the speech production process. However, there are restrictions on the strength of the individual memory traces and on the strengths of the associations between category labels and exemplars. Since any speech sound is more likely to be associated consistently with a category at the phonological or word form level than with a category label based on the ambient temperature or hair color of the speaker, the model can be said to be restricted not in the kinds of external properties that form labels, but in the strength of the association between exemplars and labels: frequently co-occurring traits will have stronger associations than less frequently co-occurring traits. Thus, a limitation is placed on how the associations are maintained and therefore indirectly on the kinds of abstractions that a given speaker-hearer is likely to conceptualize as relevant in speech processing.

13.9 Speaker and listener agency in an exemplar based approach – suggestions for further work

One problem remains to be accounted for, namely the difference in propensity for deletion of [w] before [ð̞] between individual speakers that was observed in the apparent time and real time studies in chapters 11 and 12. The automaticity implied in the constant restructuring of the lexicon fails to account for the observation that individual members of a subgroup of the speech community behave differently in spite of statistically robust associations between a particular speech production process and the particular subgroup. As an example, consider the data on deletion of [w] before [ð̞] for the informant LAL (cf. table 12.9 in chapter 12). LAL is a female middle class speaker, who at the time when she was first recorded, in 1986, was already an adult. In this first recording, her tendency to delete [w] before [ð̞] was quite low, since she only did it in the high frequency word form 'blevet' [b̥leːwəð] *got (aux.)*. In fact she was the most conservative speaker with respect to the process of [w] deletion in the entire set of informants analyzed in the present study. In the re-recording from 2006, however, she has become a categorical deleter of [w]. In both the old and the new recording, then, she is extreme in her behavior with respect to the process studied relative to the age group she belongs to. This raises the question of how individuals form generalizations over the exemplar space and to what extent distributions based on observations across speakers reflect the distributions stored in the lexicons of individual speakers. While I have attempted to control for differences between word forms and individual speakers in the quantitative analyses, speakers like LAL pose a challenge for the automaticity currently implicit in much work within the exemplar based approach. A way of dealing with this would be to examine in greater detail the speech behavior of individual speakers in the groups studied here. This would provide a reflection of the extent of both intra- and inter-individual variation, and one could look for correlations between speech behavior and more fine grained information about the speakers, as is already common in sociolinguistic studies of communities of practice, where ethnographic field work methods are employed to gain an understanding of how speakers use linguistic variation to convey social meaning (cf. Eckert (2000), Podesva (2007), and, for Danish, Quist & Maegaard (2005)). However, such studies still rely on the association between variation in speech production and social characteristics of the speakers. In order to gain a deeper understanding of the differences in patterns of production, combining the studies with perception of the variation seems an important next step. As far as the social aspect of phonetic variation is concerned, studies of language attitudes like the ones summarized for Danish in Kristiansen (fthc.) strongly suggest an association between covert prestige and the diffusion of phonetic features from Copenhagen to other varieties of Danish. However, these studies reflect the average evaluations of groups of adolescents without a direct observation of their behavior in speech production. Maegaard (2008) presents a study of more locally constructed social meanings and the evaluation of the speech behavior associated with particular sub-groups, as established through ethnographic field work at a school in Copenhagen. She finds that the perception of certain features do not always match the values indirectly ascribed to them by the speakers and their peer group, suggesting that the structuring of the phonetic exemplar space with respect to categories that are important in the identity construction of adolescents takes place also with reference to locally constructed categories, and not necessarily only on the basis of patterns that can be observed at the macro-level of social categorization (e.g. class, gender, age). While associations and discrepancies may be inferred on statistical grounds in language attitude studies where listeners evaluate phonetic variation that is not necessarily a part of their own variety or register, it would be interesting to study the relationship at the level of individuals, in order to see whether the degree to which participation in a particular phonetic process is correlated with the positive evaluation of speech in which the outcome of this process is a prominent feature.

It would also be interesting to study the perception of variation in speech from a more directly linguistic point of view, for example through the use of lexical decision tasks. Since the processes of (ɑj) monophthongization and (d) deletion were observed to correlate with word frequency, the hypothesis in an exemplar based approach is that high frequency words contain denser distributions of reduced variants of the word forms. If this is a feature of lexical representation, the differences should affect lexical decision. For example, one might investigate whether tokens of a word in which (ɑj) is often monophthongized, like 'drejer' [ˈd̥rɑjɐ] *turn (present)*, is recognized more quickly and accurately than monophthongized tokens of a word like 'streger' [ˈsd̥ʁɑjɐ] *lines*, where the process is rare. Similarly, the strength of the associations between reduced variants and contextual factors like the segmental context and stress could be tested. Recall that (g) was seen to be more likely to be reduced between two vowels than between a vowel and a consonant. If this difference is represented lexically, one would expect recognition of [ˈtˢɑɣɐ] for 'takker' *spikes* to be faster than recognition of [tˢaɣd̥ɐ] for 'takter' *beats*.

Such associations between statistical tendencies observed in speech production and listener behavior in tasks requiring lexical processing would provide a method for evaluating the role of the patterns that can be observed for groups of speakers in the structuring of the mental lexicon of the individual language user. These tests would help us understand how representations of lexical form are structured with respect to different linguistic and extra-linguistic factors which may form the basis for abstraction. One might also explore the relationship between the different domains of categorization. If [ɣ] is associated as a variant of /g/ with both intervocalic context and male speech, one might wonder whether hearing [ɣ] in intervocalic position in a word like 'ligger' [ˈleɣɐ] *lie (pres.)* spoken in a female voice is harder to recognize than when the same form of the word is presented in a male voice. This would provide strong evidence that the association between social characteristics and phonetic variation is lexically stored.

In short, while the principle of pushing all observed phonetic variation into the mental lexicon provides a good framework for understanding how word specific effects can arise and how sound change may be diffused lexically through successive generations, more work is needed to study further how prevalent such lexical effects are in the variety of processes that occur in running speech, and to what extent they influence the categorization and processing of phonetic variation. The present study has shown that reduction of consonants in contemporary Copenhagen Danish is affected by many of the same processes that have been observed both historically and synchronically in other languages, and that these processes are influenced by phonetic factors in a way that further supports the conceptualization of speech production as a complex process of coordination of articulatory gestures. In addition, the investigation emphasizes that the regularity with which these factors influence speech production is probabilistic rather than categorical. It has been argued that effects at abstract, structural levels of description, like effects of position in the syllable and position in the word, should be viewed as emergent properties of the speech signal. The extent to which the statistical generalizations that have been found for the reduction of consonants in contemporary Copenhagen Danish spontaneous speech influence the linguistic and sociolinguistic processing of the speech signal requires further work. It has been suggested that perceptual work involving tasks which can elicit both lexical and attitudinal responses would be an interesting extension in order to further understand the relationship between language use and abstract language structure.

References

Albris, Jon (1991) "Style Analysis" in F. Gregersens & I. L. Pedersen (eds.) (1991) *The Copenhagen Study in Urban Sociolinguistics 1*, pp. 45-106.

Anderson, A.H., Bader, M., Bard, E.G., Boyle, E.,Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H.S.and Weinert, R. (1991) "The HCRC Map Task Corpus" *Language and Speech* 34, 4, pp. 351-366.

Arrild, Josephine B. (unpublished) "To [b] or not to [b]. En akustisk undersøgelse af eksplosionfaser i danske /p t k/-lyde under forskellige trykbetingelser", unpublished BA-paper, University of Copenhagen.

Aylett, M. & Alice Turk (2004) "The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech" *Language and Speech* 47 pp. 31-56

Baayen, R. Harald (2008) *Analyzing Linguistic Data*, Cambridge University Press, Cambridge, UK.

Basbøll, Hans (2005) *The Phonology of Danish,* Oxford University Press, Oxford, UK.

Bailey, Guy (2002) "Real and Apparent Time" in J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, Blackwell, Oxford, UK.

Blondeau, Hélène & Gillian Sankoff (2008) "Language Change across the Lifespan: /r/ in Montreal French" *Language* 83,3, pp. 560-588.

Bloomfield, Leonard (1933)*Language* Henry Holt, New York, NY, USA.

Brink, Lars & Jørn Lund (1975) *Dansk Rigsmål 1- Lydudviklingen siden 1840 med særligt henblik på sociolektern i København*, Gyldendal, Copenhagen, Denmark.

Brink, Lars & Jørn Lund (1975) *Dansk Rigsmål 2- Lydudviklingen siden 1840 med særligt henblik på sociolektern i København*, Gyldendal, Copenhagen, Denmark.

Brink, Lars, Steffen Heger, Jørn Lund & Jens Normann Jørgensen (1991) *Den Store Danske Udtaleordbog* Munksgaard, Copenhagen, Denmark.

Browman, Catherine & Louis Goldstein (1990) "Articulatory Gestures as Phonological Units" *Phonology* 6, pp. 201-231.

Browman, Catherine & Louis Goldstein (1995) "Gestural syllable position effects in American English" in F. Bell-Berti & L. J. Raphael (eds.) *Producing Speech: Contemporary Issues*, A.I.P. Press, Woodbury, NY, USA, pp. 19-33.

Bybee, Joan (2002) *Phonology and Language Use*, Cambridge University Press, Cambridge, England

Byrd, Dani & Cheng Tan (1996) "Saying consonant clusters quickly" *Journal of Phonetics* 24, pp. 263-282.

Campbell-Kibler, Kathryn (2007) "Accent, (ING) and the social logic of listener perceptions" *American*

Speech 82,1, pp.32-64.

Chambers, J.K. (2002)"Patterns of Variation including Change" in J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.) *The Handbook of Language Variation and Change*, Blackwell, Oxford, UK, pp. 349-372.

Chambers, J.K. (2003) *Sociolinguistic Theory* 2nd ed., Blackwell, Oxford, UK.

Cho, Taehong, James McQueen & Ethan Cox (2007) "Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English", *Journal of Phonetics* 35,2, pp. 210-243

Clements , G.N. & Elizabeth Hume (1995) "The Internal Organization of Speech Sounds" in J. Goldsmith (ed.) *Handbook of Phonological Theory,* Blackwell, Oxford, UK, pp. 245-306.

Clopper, Cynthia G. & David B. Pisoni (2006) "Effects of region of origin and geographic mobility on perceptual dialect categorization" *Language Variation and Change* 18, pp. 193-221.

Cutler, Anne (1992) "Psychology and the Segment" in G.J. Docherty & D.R. Ladd *Papers in Laboratory Phonology II - Gesture, Segment, Prosody* Cambridge University Press, Cambridge, UK, pp. 290-295.

Docherty, Gerard & Paul Foulkes (2000) "Speaker, Speech and Knowledge of Speech Sounds" in N. Burton-Roberts, P. Carr & G. Docherty *Phonological Knowledge – Conceptual and Empirical Issues* Oxford University Press, Oxford, UK, pp. 105-130-

Dubois, S. & B. Horvath (1999) "When the music changes you change too: Gender and language change in Cajun English" *Language Variation and Change* 11, 3, pp. 287-313

Eckert, Penelope (2000) *Linguistic Variation as Social Practice*, Blackwell, Oxford, UK.

Fasold, R. (1991) "The quiet demise of variable rules" *American Speech* 66, pp. 3-21.

Fischer-Jørgensen, Eli (1995) *Trends in Phonological Theory Until 1975*, C.A. Reitzel, Copenhagen, Denmark.

Fosler-Lussier, E. & N. Morgan (1999) "Effects of speaking rate and word frequency on pronunciations in conversational speech" *Speech Communication* 29, pp. 137-158.

Francis, W. Nelson & Henry Kucera (1982) *Frequency analysis of English usage*, Houghton Mifflin, Boston, MA, USA.

Fromkin, Victoria (1971) "The Non-anomalous Nature of Anomalous Utterances" *Language* 47,1, pp. 27-52.

Goldinger, Stephen D. (1996) " Words and voices: Episodic traces in spoken word identification and recognition memory" *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 22, pp. 1166-1183.

Goldinger, Stephen D. & Tamiko Azuma (2004) "Episodic memory reflected in printed word naming" *Psychonomic Bulletin & Review* 11, 4, pp. 716-722.

Greenberg, Steven (1999) "Speaking in shorthand – a syllable-centric perspective for understanding pronunciation variation", *Speech Communication* 29, pp. 159-176.

Greenberg, Steven, H. Carvey & L. Hitchcock (2002) "The relation between stress accent and pronunciation variation in spontaneous American English discourse" *Proceedings of the International Speech Communication Association on Prosody and Speech Processing*, pp. 351-354.

Gregersen, Frans (fthc.) "The data and design of the LANCHART study" to appear in *Acta Linguistica Hafniensia* 41, pp. 4-37

Gregersen, Frans, Søren Beck Nielsen & Jacob Thøgersen (fthc.) "Stepping into the same river twice: On the Discourse Context Analysis in the LANCHART Project" to appear in *Acta Linguistica Hafniensia* 41, pp. 38-73.

Gregersen, Frans, Marie Maegaard & Nicolai Pharao (fthc.) "The long and short of (a) variation in Danish" to appear in *Acta Linguistica Hafniensia* 41, pp. 74-97.

Gregersen, Frans & Inge Lise Pedersen (eds.) (1991) *The Copenhagen Study in Urban Sociolinguistics 1*

Grønnum, Nina (1992) *The Groundworks of Danish Intonation – An Introduction*, Museum Tusculanum Press, Copenhagen, Denmark.

Grønnum, Nina (1996) "Danish Vowels: scratching the recent surface in a phonological experiment" *Acta Linguistica Hafniensia* 28, pp. 5-63.

Grønnum, Nina (1998) "Danish" *Journal of the International Phonetic Association* 28, 1-2, pp. 99-105.

Grønnum, Nina (1999) "Syllables at multiple levels of representation in Danish" in *Syllabes : IIèmes Journées d'Etudes Linguistiques, Nantes, 25-26-27 mars 1999*, Équipe AAI - Université de Nantes, Nantes, France, pp. 24-29.

Grønnum, Nina (2005) *Fonetik og fonologi: almen og dansk* (3rd edt.), Akademisk Forlag, Copenhagen, Denmark.

Grønnum, Nina (2006) "DanPASS – A Danish Phonetically Annotated Spontaneous Speech Corpus" in N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, D. Tapias (eds.) *Proceedings of the 5th International Conference on Language Resources and Evaluation,* European Language Resources Association, Genova.

Grønnum, Nina (2007) *Rødgrød med fløde – en lille bog om dansk fonetik*, Akademisk forlag, Copenhagen, Denmark.

Grønnum, Nina (2009) "A Danish phonetically annotated spontaneous speech corpus (DanPASS)", *Speech Communication* 51, pp. 594-603.

Grønnum, Nina & Hans Basbøll (2007) "Danish Stød: Phonological and Cognitive Issues" in M.-J. Solé, P. Speeter Beddor & M. Ohala (eds.) *Experimental Approaches to Phonology*, Oxford University Press, Oxford, UK, pp. 192-206.

Guy, Gregory (1980) "Variation in the group and the individual" in W. Labov (ed.) *Locating language in time and space*, Academic Press, New York, USA, pp. 1-36.

Handbook (1999) *Handbook of the International Phonetics Association*, Cambridge University Press, Cambridge, UK.

Hawkins, Sarah and Rachel Smith (2001) "Polysp: a polysystemic, phonetically-rich approach to speech understanding" *Italian Journal of Linguistics - Rivista di Linguistica* 13, pp. 99-188.

Henrichsen, Peter Juel (2002) "Fyrre Kilometer Kryds og Bolle - metoder til grammatisk opmærkning i største skala" *NyS* 30, pp.68-88.

Holmberg, Henrik (1991) "The Sociophonetics of some Vowel Variables in Copenhagen Speech" in F. Gregersen & I.L. Pedersen (eds.) (1991) *The Copenhagen Study in Urban Sociolinguistics 1*, pp. 107-240.

Holmes, Janet (1997) "Setting new standards: Sound change and gender in New Zealand English", *English World-Wide* 18, 1, pp. 107-142.

Horvath, Barbara M. (1985) *Variation in Australian English*, Cambridge University Press, Cambridge, UK.

Jaeger, Florian T. & Victor Kuperman (2009) "Issues and Solutions in Fitting, Evaluating and Interpreting Regression Models", handout of presentation given at CUMY Workshop on mixed-effects models.

Johnson, Daniel Ezra (2009) "Getting off the GoldVarb Standard: Introducing Rbrul for Mixed Effects Variable Rule Analysis" *Language and Linguistics Compass* 3,1, 359-383.

Johnson, Keith (1997) "Speech perception without speaker normalization" in K. Johnson & J.W. Mullenix (eds.) *Talker Variability in Speech Processing* Academic Press, San Diego, CA, USA, pp. 145-166.

Johnson, Keith (2007) "Decisions and Mechanisms in Exemplar-based Phonology" in M.-J. Solé, P. Speeter Beddor & M. Ohala (eds.) *Experimental Approaches to Phonology*, Oxford University Press, Oxford, UK, pp. 25-40.

Johnson, Keith (2008) *Quantitative Methods in Linguistics*, Blackwell Publishing, Malden, MA, USA.

Jurafsky, Daniel, Alan Bell & Cynthia Girand (2002) "The role of the lemma in form variation" in C. Gussenhoven & N. Warner (eds.) *Papers in Laboratory Phonology 7*, Mouton de Gruyter, Berlin, Germany, pp. 3-34.

Krakow, Rena (1989) *The articulatory organization of syllables*, unpublished ph.d. dissertation, Yale University, USA.

Keating, Patricia, Taehong Cho, Cécile Fougeron & Chai-Shune Hsu (2003) ”Domain-initial articulatory strengthening in four languages” in J. Local, R. Ogden & R. Temple (eds.) *Phonetic Interpretation – Papers in Laboratory Phonology VI*, Cambridge University Press, Cambridge, UK, pp. 145-163

Kiparsky, Paul (1982) “From Cyclic Phonology to Lexical Phonology**”** in H. V. D. Hulst & N. Smith (eds.) *The Structure of Phonological Representations*, reprinted in J. Goldsmith (ed.) *Phonological Theory – The Essential Readings*, Blackwell Publishing, Malden, MA, USA.

Kochetov, Alexei (2006) “Syllable position effects and gestural organization: Articulatory evidence from Russian” in L. Goldstein, D.H. Whalen & C. T. Best (eds.) *Laboratory Phonology 8*, Mouton de Gruyter, Berlin, Germany, pp. 565-588.

Kristensen, Kjeld & J. Normann Jørgensen (1994) *Moderne sjællandsk,* C.A. Reitzel, Copenhagen, Denmark.

Kristiansen, Tore (fthc.) “The macro-level social meanings of late-modern Danish accents” to appear in *Acta Linguistica Hafniensia* 41, pp. 167-192.

Labov, William (1966(2006)) *The Social Stratification of English in New York City* 2nd ed., Cambridge University Press, Cambridge, UK.

Labov, William (1969) “Contraction, deletion and inherent variability of the English copula” *Language* 45, pp. 715-762.

Labov, William (1972) *Sociolinguistic Patterns* University of Philadelphia Press, Philadelphia, PA, USA.

Labov, William (1990) “The intersection of sex and social class in the course of linguistic change” *Language Variation and Change* 2, pp. 205-254.

Labov, William (1994) *Principles of Linguistic Change 1 – Internal Factors*, Blackwell Publishing, Malden, MA, USA.

Labov, William (2001) *Principles of Linguistic Change 2 – Social Factors*, Blackwell Publishing, Malden, MA, USA.

Ladd, D. Robert (2006) “ “Distinctive phones” in surface representation” in L. Goldstein, D.H. Whalen & C. T. Best (eds.) *Laboratory Phonology 8*, Mouton de Gruyter, Berlin, Germany, pp. 1-26.

Lavoie, Lisa M. (2001) *Consonant strength: phonological patterns and phonetic manifestation,* Garland Publishing, New York, USA.

Lenneberg, Eric H. (1967) *Biological Foundations of Language*, Wiley, New York, NY, USA.

Levelt, Willem J.M. (1989) *Speaking – From Intention to Articulation*, The MIT Press, Cambridge, MA, USA.

Levelt, Willem J.M., Ardi Roelofs & Antje S. Meyer (1999) “A theory of lexical acces in speech production”, *Behavioral and Brain Sciences 22*, pp. 1-75.

Maegaard, Marie (2007) *Udtalevariation og –forandring i københavnsk (danske talesprog 8)* C.A. Reitzel, Copenhagen, Denmark.

Marslen-Wilson, William D. (1984) “Function and process in spoken word recognition - a tutorial review” in H.Bouma. & D.G. Bouwhuis (eds.) *Attention & performance X* , Lawrence Erlbaum, London, UK.

Mendoza-Denton, Norma, Jennifer hay & Stefanie Jannedy (2003) “Probabilistic Sociolinguistics: Beyond Variable Rules” in R. Bod, J. Hay & S. Jannedy (eds.) *Probabilistic Linguistics*, The MIT Press, Cambridge, MA, USA, pp. 97-138.

Meyerhoff, Miriam (2006) *Introducing Sociolinguistics*, Routledge, New York, NY, USA.

Milroy, Lesley & Matthew Gordon (2003) *Sociolinguistics – Method and Interpretation*, Blackwell Publishing, Malden, MA, USA.

Nielsen, Bent Jul & Magda Nyberg (1993) ”Talesprogsvariationen i Odder kommune I” *danske folkemål* 34, pp. 45-202.

Nolan, Francis (1992) “The descriptive role of segments: evidence from assimilation” in G.J. Docherty & D.R. Ladd *Papers in Laboratory Phonology II - Gesture, Segment, Prosody* Cambridge University Press, Cambridge, UK, pp. 261-280.

Norris, Dennis, James M. McQueen & Anne Cutler (2003) “Perceptual learning in speech” *Cognitive Psychology* 47, pp. 204-238.

Ohala, John (1981) ”The Listener as a Source of Sound Change” in C.S. Masek, R.A. Hendrick & M.F. Miller (eds.) *Papers from the Parasession on Language and Behavior* Chicago Linguistics Society, Chicago, USA, pp. 178-203.

Pharao, Nicolai (unpublished) *Inde bag ved ordene – nogle fonologiske repræsentationer og processers psykologiske realitet for danske talere*, unpublished MA thesis, University of Copenhagen.

Pharao, Nicolai (2003) ”The psychological reality of ambisyllabic consonants in Danish” in D. Rescasens & M.-J. Solé (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*

Phillips, Betty S. (1984) “Word frequency and the actuation of sound change” *Language* 60, pp. 320-342.

Pierrehumbert, Janet (2001) “Exemplar dynamics: Word frequency, lenition and contrast” in J. L. Bybee P. Hopper (ed.) *Frequency and the Emergence of Linguistic Structure*, John Benjamins, Philadelphia, USA, pp. 137-157.

Pierrehumbert, Janet (2002) “Word-specific phonetics” in C. Gussenhoven & N. Warner (eds.) *Papers in Laboratory Phonology 7*, Mouton de Gruyter, Berlin, Germany, pp. 101-140.

Pierrehumbert, Janet (2006) "The next toolkit" *Journal of Phonetics* 34, pp. 516-530.

Plug, Leendert (2005) "Phonetic reduction and categorization in exemplar-based representation" in S. Blaho, L. Vicente & E. Scoorlemmer (eds.) *Proceedings of ConSOLE XIII*, pp. 287-311.

Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen (2005) "Articulatory planing is gradient and sensitive to informational redundancy" *Phonetica* 62,2-4, pp.146-159.

Podesva, Robert (2007) "Phonation type as a stylistic variable: The use of falsetto in constructing a persona" *Journal of Sociolinguistics* 11, 4, pp. 478-504.

Prince, E. (1981) "Toward a taxonomy of given-new information" in Cole (ed.) *Radical pragmatics,* Academic Press, New York, USA, pp. 225-256.

Quist, Pia & Marie Maegaard (2005) "Etnografi, praksis og sproglig variation – om etnografisk metode i udforskningen af sproglig variation som social praksis" *Nydanske Sprogstudier* 33, pp. 42-73

Rathje, Marianne (2008) *Generationssprog I mundtlig interaction*, unpublished ph.d. dissertation, University of Copenhagen, Copenhagen, Denmark.

Raymond, William D., Robin Dautricourt & Elizabeth Hume (2006) "Word-internal /t,d/ deletion in spontaneous speech: modeling the effects of extra-linguistic, lexical, and phonological factors", *Language Variation and Change* 18, 1, pp. 55-98.

Rischel, Jørgen (1983) "On unit accentuation in Danish – and the distinction between deep and surface phonology" *Folia Linguistica* XVII,1-2, pp. 51-97.

Rose, Mary (2006) *Language, Place and Identity in Later Life* unpublished ph.d. dissertation, Stanford University, CA, USA.

Sankoff, David & William Labov (1979) "On the uses of variable rules" *Language in Society* 8, pp. 189-222.

Sankoff, Gillian (2006) "Age: Apparent time and real time" in *Elsevier Encyclopedia of Language and Linguistics* 2nd ed. Article number: LALI:01479.

Schachtenhaufen, Ruben (unpublished) *Scwa-assimilation i dansk spontan tale*, unpublished MA thesis, University of Copenhagen, Copenhagen, Denmark.

Shriberg, L. & G. Lof (1991) "Reliability studies in broad and narrow phonetic transcription" *Clinical Linguistics and Phonetics* 5, pp. 225-279.

Taft, Marcus & Gail Hambly (1986) "Exploring the Cohort Model of spoken word recognition" *Cognition* 22, pp. 259-282.

Trubetzkoy, N.S. (1962) *Grundzüge der Phonologie*, Vandenhoeck & Ruprect, Göttingen, Germany.

Turk, Alice & Laurence White (1999) "Structural influences on accentual lengthening in English" *Journal of Phonetics* 27, pp. 171-206.

Tøndering, John (2008) *Skitser af prosodi i spontant dansk,* unpublished ph.d.-dissertation, University of Copenhagen, Copenhagen, Denmark.

Venneman, Theo (1988) *Preference Laws for Syllable Structure and the Explanation of Sound Change*, Mouton de Gruyter, Berlin, Germany.

Yaeger-Dror, Malcah (1994) "Phonetic evidence for sound change in Quebec French" in P.A. Keating (ed.) *Phonological Structure and Phonetic Form – Papers in Laboratory Phonology III*, Cambridge University Press, Cambridge, UK, pp. 267-292.

Zue, Victor W. & Martha Laferriere (1979) "Acoustic study of internal /t,d/ in American English" *Journal of the Acoustical Society of America* 66, pp. 1039-1050.

Consonant Reduction in Copenhagen Danish – English summary

This dissertation explores reduction of consonants in spontaneous connected speech. The quantitative study investigates the influence from both linguistic and extra-linguistic factors on monophthongization and reduction and deletion of plosives using the DanPASS corpus. The influence from the different factors is studied using mixed-effects multiple logistic regression. In addition, the dissertation focuses on the role of language use in consonant reduction. The effect of word frequency is included in the statistical analyses in order to explore whether frequently used words are more likely to contain reduced consonants than words which are rarely used, as has been found in a number of other languages. Consonant deletion is also studied as a process of on-going change using a part of the LANCHART corpus. Deletion of [w] before [ð̩] is studied in the speech of 22 middle class speakers that represent two different generations. The tendency for deletion is compared both across the two generations, in a traditional apparent time study, but the speakers are also compared to themselves over time in a real time panel study, in order to examine the extent to which speakers continue to participate in a process of change through adulthood.

Monophthongization and plosive reduction and deletion are found to be processes of stable variation which are probabilistically regular and mainly affected by stress, segmental context and rate of speech. Reduction of consonants is more likely in unstressed syllables and in intervocalic contexts and the tendency increases in faster speech. Deletion of plosives is also more likely in unstressed syllables, but tends to occur more often in consonant clusters. The expected effect of word frequency is found for the monophthongization of [ɑj] and the deletion of word final [d̥], showing that consonant reduction may be influenced by frequency of use, but that the effect is limited in the stable variation of consonants in middle class Copenhagen Danish.

The effect of frequency is also found in the apparent and real time studies of [w] deletion, together with an increased tendency for deletion of [w] in unstressed syllables and after front vowels. Both in apparent time and real time, the spread of the process is conditioned by word frequency, being most likely in high frequency words, and spreading to words of lower frequency over time. This shows that a reductive sound change spreads incrementally through the vocabulary rather than being generalized to all words at the outset. The implications for models of speech production are discussed and it is suggested that the results on consonant reduction both as a synchronic process of variation and a diachronic process of on-going change is best handled by models of lexical representation which can incorporate usage based factors.

Konsonantreduktioner i københavnsk – dansk resumé

Denne afhandling udforsker konsonantreduktion i spontan, sammenhængende tale. Det kvantitative studie undersøger inflydelsen af såvel sproglige som ikke-sproglige faktorer på monoftongering og reduktion og bortfald af lukkelyde i DanPASS-korpusset. Indflydelsen fra de forskellige faktorer studeres ved hjælp af multipel logistisk regression med blandede effekter. Afhandlingen fokuserer også på sprogbrugens indflydelse på konsonantreduktion. Ordfrekvenser inddrages i de statistiske analyser for at udforske, om højfrekvente ord har større sandsynlighed for at indeholde svækkede konsonanter end ord, der bruges sjældent – en sammenhæng man kender fra en række andre sprog. Konsonantbortfald studeres tillige som en igangværende lydforandring ved brug af en del af LANCHART-korpusset. Bortfald af [w] før [ð̞] studeres hos 22 sprogbrugere fra middelklassen, der repræsenterer to generationer. Tendensen til bortfald sammenlignes på tværs af de to generationer i et traditionelt "apparent time-studie", men talerne sammenlignes også med sig selv over tid i et "real time" panelstudie, for at undersøge i hvor høj grad sprogbrugere fortsætter med at deltage i en igangværende lydforandring op gennem voksenlivet.

Monoftongering og lukkelydssvækkelse viser sig at være processer, der er i stabil variation, og som udviser probabilistisk regelmæssige sammenhænge med de sproglige faktorer. Det er hovedsageligt tryk, nabolyde og talehastighed, der er afgørende for, om processerne appliceres. Konsonantreduktion forekommer oftere i trykløse stavelser og i intervokalisk position, og tendensen til reduktion stiger med talehastighed. Bortfald af lukkelyde er også mest sandsynligt i trykløse stavelser, men forekommer oftere mellem konsonanter. Den forventede effekt af ordfrekvens findes for monoftongering af [ɑj] og bortfald af ordfinalt [d̥], hvilket viser, at konsonantreduktion kan påvirkes af forekomstfrekvens, men at effekten er begrænset i den stabile variation, der kan observeres i konsonanternes udtale i middelklassens københavnsk.

Frekvenseffekten finds også i "apparent" og "real time"-studierne af bortfald af [w], sammen med en høj tendens til [w]-bortfald i trykløse stavelse og efter fortungevokaler. Både på tværs af generationer og over tid påvirkes spredningen af denne proces af ordfrekvens: den forekommer oftest i højfrekvente ord og spreder sig til ord af lavere frekvens over tid. Det viser, at en svækkende lydforandring spredes løbende gennem ordforrådet, snarere end at den generaliseres til alle relevante ord fra begyndelsen. Implikationerne for modeller for taleproduktion diskuteres, og det foreslås, at resultaterne for konsonantreduktion både som en synkron variationsproces og som en igangværende lydforandring bedst kan håndteres i modeller for leksikalsk repræsentation, der inkorporerer sprogbrugsbaserede faktorer.

APPENDICES

Contents

Appendix A – Datasets for the phonetic variables

This appendix contains the full word lists for each of the phonetic variables studied in the chapter "The descriptive study of the role of linguistic and extra-linguistic factors in consonant reduction".

Full dataset for (ɑj)			
Word form	**English**	**sum**	**% monophthongized**
nej	no	313	15
jeg	i	244	94
drejer	turn (pres.)	147	49
dreje	turn	71	6
meget	very	68	96
mig	me	66	47
indianerlejren	the indian camp	53	25
dig	you	51	76
vej	road	41	2
byggelegepladsen	the playground	29	3
flodleje	riverbed	25	0
vejen	the road	25	8
indhegning	fencing	24	4
lejr	camp	22	14
byggelegeplads	playground	13	8
indhegningen	the fencing	12	0
tegner	draw (pres.)	11	0
tegnet	draw (ptcpl.)	11	0
sig	oneself	9	56
halvvejs	halfway	7	0
indianerlejr	indian camp	7	14
sejlklubben	the sailing club	7	0
tegne	draw (inf.)	7	0
sejlklub	sailing club	5	20
tegning	drawing	5	0
tegningen	the drawing	5	0
legepladsen	the playground	4	0
undervejs	under way	4	0
flodlejet	the riverbed	3	100
Stationsvej	proper noun	3	0
allright	all right	2	0
drejede	turned	2	100
drejet	turned (ptcpl.)	2	100
egen	own	2	50

Full dataset for (ɑj) - continued			
Word form	**English**	**sum**	**% monophthongized**
hej	hi	2	0
legeplads	playground	2	0
dejlig	lovely	1	0
dejligt	lovely	1	0
eget	own	1	100
ej	nah	1	0
fejl	error	1	0
fejltagelse	mistake	1	0
fortegnelse	index	1	0
gelejdet	ushered	1	0
guide	guide	1	0
Ibsvej	proper noun	1	0
indhegnede	fenced	1	0
indianerlejre	indian camps	1	0
indtegner	draw (pres.)	1	0
kendetegn	characteristic	1	0
leger	play (pres.)	1	100
midtvejs	midway	1	0
pegefinger	index finger	1	0
regne	calculate	1	0
regner	calculate (pres.)	1	0
samarbejde	cooperate	1	0
stavefejl	spelling error	1	100
strege	strike (vb.)	1	0
streger	strike (pres.)	1	0
tegn	draw (imp.)	1	0
veje	roads	1	0
vejs	road’s	1	0

Full dataset for (ɒw)			
Word form	**English**	**sum**	**% monophthongized**
over	over	187	15
og	and	164	41
jernbaneoverskæring	railway crossing	45	44
okay	okay	41	54
skov	forest	37	3
oven	above	34	35
jernbaneoverskæringen	the railway crossing	33	33
hov	oops	10	10
sovende	sleeping	8	25
ovre	over at	7	29
dog	though	6	0
derovre	over there	3	67
herovre	over here	3	33
overhovedet	at all	2	0
baneoverskæringen	railway crossing	1	0
grov	rough	1	0
jernbaneoverskæringens	the railway crossing's	1	0
jernbaneoverskæringer	railway crossings	1	100
ovenfor	up above	1	0
overfor	across from	1	0
sjovt	fun	1	0
skovs	forest's	1	0

Full dataset for (b)			
word form	**English**	**sum**	**% reduced**
op	up	131	4
klippe	cliff	74	31
oppe	up at	72	31
klippehave	rock garden	28	18
opad	upwards	25	0
kalkstensklipper	lime stone cliffs	18	28
simpelthen	simply	18	78
klippehaven	the rock garden	17	12
klippeskred	avalanche	16	44
skarpt	sharply	16	0
jep	yup	12	33
kalkstensklipperne	the limestone cliffs	8	25
klipper	cliffs	8	13
sejlklubben	the sailing club	8	0
toppen	the top	6	0
skarp	sharp	5	0
jeps	yup	3	0
Jesper	proper name	3	0
sejlklub	sailing club	3	0
klippen	the cliff	2	0
klippeskredet	the avalanche	2	0
knap	scarce	2	0
opdager	discover (pres)	2	0
dropper	drop (pres.)	1	100
eksempel	example	1	0
hjælpe	help	1	0
hoppet	jump (ptcpl.)	1	0
Ibsvej	proper name	1	0
kæmpebue	giant arch	1	0

Full dataset for (b) - continued			
word form	**English**	**sum**	**% reduced**
klippehavens	the rock garden's	1	100
klippehaver	rock gardens	1	0
opdage	discover (inf.)	1	0
opdaget	discover (ptcpl.)	1	0
opfatte	perceive	1	0
oppefra	from above	1	0
opringning	call (n.)	1	0
opstår	arises	1	0
Philip	proper name	1	100
stoppede	stopped	1	0
stopper	stop (pres.)	1	100
tabt	lost	1	0
tabte	lost (past)	1	0
toppunktet	the top point	1	0
tramper	trample (pres.)	1	0
ups	Oops	1	0

Full dataset for (p)			
word form	**English**	**sum**	**% reduced**
på	on	893	8
passerer	pass (pres.)	87	2
passeret	passed	73	5
bananpalmen	the banana palm	51	0
passere	pass (inf.)	49	0
udsigtspunkt	vantage point	43	0
byggelegepladsen	the playground	35	3
parkeret	parked	33	0
startpunktet	the starting point	33	3
vagtpost	guard post	32	0
parkerede	parked	28	7
private	private (pl.)	27	4
pakhus	warehouse	26	0
totempæl	totempole	23	0
par	pair	21	0
totempælen	the totempole	20	10
præcis	exactly	18	11
passer	fit (pres.)	17	0
løvepark	lion park	16	0
løveparken	the lion park	16	6
punkt	point	15	0
tidspunkt	point in time	15	33
passe	fit (inf.)	14	0
udsigtspunktet	the vantage point	14	0
bananpalme	banana palm	13	0
byggelegeplads	playground	13	8
pakhuset	the warehouse	10	0
papiret	the paper	9	33
parken	the park	6	0
startpunkt	starting point	6	0
gårdsplads	courtyard	5	0
gårdspladsen	the courtyard	5	0
papirets	the paper's	5	20
penge	money	5	0
placeret	placed	5	0
byggepladsen	the construction site	4	0
kompasrosen	the compass rose	4	0
legepladsen	the playground	4	0
plads	space	4	0
pæn	nice	3	0
parallelt	parallel	3	0
parasol	parasol	3	0
Teaterpassagen	the theater passage	3	0

Full dataset for (p) - continued			
word form	**English**	**sum**	**% reduced**
bananpalmer	banana palms	2	0
legeplads	playground	2	0
parallel	parallel	2	0
prøv	try	2	0
prøver	try (pres.)	2	0
prik	dot	2	0
problem	problem	2	0
bananpalmens	the banana palm's	1	0
bananpalmerne	the banana palms	1	0
byggeplads	construction site	1	0
kaktusplante	cactus plant	1	0
kompasretningerne	the compass corners	1	0
kompasset	the compass	1	0
kompromis	compromise	1	0
midtpunktet	the midpoint	1	0
p.t.	presently (abbrev.)	1	0
pænt	nicely	1	0
palmen	the palm	1	0
palmer	palms	1	0
panden	the forehead	1	0
papir	paper	1	0
papirhjørnerne	the paper corners	1	0
papirkanten	the paper's edge	1	0
parallelle	parallel (pl.)	1	0
park	park	1	0
parkens	the park's	1	0
parkerer	park (pres.)	1	100
parkeringsplads	parking space	1	0
pas	pass	1	0
passage	passage	1	0
pegefinger	index finger	1	0
periferien	the periphery	1	0
person	person	1	0
perspektivet	the perspective	1	0
Peter	proper name	1	0
placere	place (inf.)	1	0
placerer	place (pres.)	1	0
planen	the plan	1	0
posthus	post office	1	0
præcisere	clarify	1	0
præcist	exact	1	0
prøve	try (inf.)	1	0
pragtfuldt	marvelous	1	0
problemer	problems	1	0

Full dataset for (p) - continued			
word form	**English**	**sum**	**% reduced**
proviant	goods	1	0
puh	phew	1	100
pukkel	hump	1	0
punktet	the point	1	0
punktum	period	1	0
pyt	nevermind	1	0
ticentimeterspunkt	ten centimeter point	1	0
tilpasse	adjust	1	0
totempælens	the totem pole's	1	0
udgangspunkt	point of departure	1	0

Full dataset for (g)			
Word form	**English gloss**	**sum**	**% reduced**
ikke	not	587	54
ligger	lies (vb)	473	74
stik	due	217	6
nok	probably	74	18
stykke	piece	64	44
faktisk	actually	59	36
udsigtspunkt	vantage point	43	35
kirkegård	churchyard	39	62
byggelegepladsen	the playground	35	66
måske	perhaps	33	3
vagtpost	guard post	32	6
rigtigt	right	31	61
telefonboksen	the telephone booth	31	10
tak	thanks	28	14
kornmarker	fields	27	44
marker	fields	27	30
direkte	directly	26	8
pakhus	storage facility	26	27
kirkegården	the churchyard	25	44
kigger	looks (vb)	24	75
telefonboks	telephone booth	17	12
løveparken	the lionpark	16	13
ligge	lie (vb)	16	38
stråtækt	thatched	16	0
løvepark	lionpark	14	21
udsigtspunktet	the vantage point	14	14
byggelegeplads	playground	13	77
seks	six	13	15
gik	went	12	0
sikkert	certainly	11	91
stadigvæk	still	11	9
takker	thanx	11	9
ganske	quite	10	20
hvilken	which	10	30
ryggen	the back	10	10
teksten	the text	10	30

Full dataset for (g) - continued			
Word form	**English gloss**	**sum**	**% reduced**
pakhuset	the warehouse	9	22
væk	gone	9	33
begge	both	7	43
kaktussen	the cactus	7	14
kan	can	7	71
stråtække	thatched	7	14
tænkte	thought	7	86
mexicaner	mexican (n.)	6	50
parken	the park	6	0
rigtig	right	6	33
straks	immediately	6	50
trækker	pull (present)	5	0
øjeblik	moment	4	50
byggepladsen	the construction site	4	100
halvcirkel	semi-circle	4	0
kaktus	cactus	4	25
sikker	certain	4	75
virkeligheden	the reality	4	0
øjeblikket	the moment	3	67
banket	knocked	3	0
kigge	look (vb.)	3	67
rigtige	right (pl.)	3	0
sagt	said (past ptcpl.)	3	33
sigter	aim (present)	3	33
slags	kind (n.)	3	33
strækning	stretch	3	33
tænke	think	3	67
virkelig	real	3	33
boksen	the box	2	0
enkelt	simple	2	50
flok	herd/crowd	2	0
flugte	follow	2	0
kornmarkerne	the fields	2	100
nøjagtig	exactly	2	0
prik	dot	2	50
Rock	rock	2	0
sigte	aim (inf.)	2	50

Full dataset for (g) - continued			
Word form	**English gloss**	**sum**	**% reduced**
snakkede	talked	2	100
stikker	stab (present)	2	50
vække	awaken (inf.)	2	50
vinkel	angle	2	0
øjesteblik	momentito	1	100
østnordøstagtigt	east-north-east-ish	1	0
afmærket	demarked	1	100
agtigt	ish	1	0
akse	axis	1	0
bænke	benches	1	0
bank	bank	1	0
banke	knock (inf.)	1	0
banken	the bank	1	0
boks	box	1	100
brugt	used	1	0
bugter	wobbles	1	0
byggeplads	construction site	1	100
bygget	built	1	100
cirka	approximately	1	100
cirkel	circle	1	0
cykel	bicycle	1	100
eksempel	example	1	0
falske	fake	1	0
fik	got	1	0
flugter	follow (present)	1	0
græsmarken	the grass field	1	0
hak	notch	1	0
huske	remember (inf.)	1	100
hverken	neither	1	0
kaktusplante	cactus plant	1	0
kalksten	lime stone	1	0
kigget	looked	1	0
kikkert	binoculars	1	0
kirke	church	1	100
kornmarken	the field	1	0
lægger	lay (present)	1	100
liggende	lying	1	0

Full dataset for (g) - continued			
Word form	**English gloss**	**sum**	**% reduced**
logikken	the logic	1	0
lykkedes	succeeded	1	0
mærkelig	strange	1	100
mærkeligformet	strangely shaped (compound)	1	0
mærker	marks	1	0
målestok	yard stick	1	0
musikhus	music hall	1	0
musikhuset	the music hall	1	0
nøjagtigt	exactly	1	0
niks	nope	1	0
nixen	no way	1	100
nord-syd-aksen	the north-south axis	1	0
park	park	1	0
parkens	the park's	1	0
perspektivet	the perspective	1	0
pragtfuldt	glorious	1	100
pukkel	hump	1	0
sektion	section	1	0
slurk	sip	1	0
smuk	beautiful	1	0
snakke	talk (inf.)	1	100
snakker	talk (present)	1	100
snakket	talk (past ptcpl.)	1	100
strækker	stretch (present)	1	100
stykker	pieces	1	100
stykket	the piece	1	0
tænkt	thought (past ptcpl.)	1	100
tekst	text	1	0
tekstens	the text's	1	100
telefonboksens	the telephonebooth's	1	0
tjek	check	1	0
trække	pull (present)	1	0
trykker	push (present)	1	0
tysk	german	1	0
udsigtskløft	vantage point (lit: view cliff)	1	0
vagtstuen	the guard room	1	0
vigtige	important (pl.)	1	0

Full dataset for (k)			
Word form	**English**	**sum**	**% reduced**
okay	okay	582	12
kan	can	252	25
kommer	come (pres)	234	6
kloster	monastery	107	0
klippe	cliff	74	1
kommet	come (ptcpl.)	62	5
kort	map	42	0
kirkegård	churchyard	39	0
kulminen	the coal mine	35	0
parkeret	parked	33	0
kolonihaver	allotments (gardens)	32	9
komme	come	30	0
kvæg	cattle	30	0
klippehave	rock garden	28	0
kløft	ravine	28	0
parkerede	parked (past)	28	7
kornmarker	fields	27	0
omkring	around	27	7
kirkegården	the churchyard	25	0
kortet	the map	25	8
kigger	looks (vb)	24	0
kavaleri	cavalery	23	0
kunne	could	23	22
bjergkløft	ravine	22	0
kavaleriet	the cavalery	19	0
kalkstensklipper	limestone cliffs	18	6
klippehaven	the rock garden	17	0
klippeskred	avalanche	16	0
kolonihaverne	the allotments (gardens)	16	0
kom	come	16	0
bjergkløften	the ravine	15	7
klosteret	the monastery	14	0
kryds	cross	13	0
kun	only	11	0
kulmine	coalmine	10	0
klar	clear	9	0
kalkstensklipperne	the limestone cliffs	8	0
klipper	cliffs	8	0
krydset	the cross	8	0

Full dataset for (k) - continued			
Word form	**English**	**sum**	**% reduced**
sejlklubben	the sailing club	8	13
kaktussen	the cactus	7	0
stakit	fence	7	0
kvæget	the cattle	5	0
sejlklub	sailing club	5	20
forkert	wrong	4	25
kaktus	cactus	4	0
kaldte	called	4	0
kant	edge	4	0
kanten	the edge	4	0
kompasrosen	the compass rose	4	0
korrekt	correct	4	0
kører	drive (pres.)	4	0
startkryds	starting intersection	4	0
kalder	call (pres.)	3	33
kender	know (pres.)	3	0
kigge	look (vb.)	3	0
knap	scarce	3	33
knækker	break (pres.)	3	0
kortets	the map's	3	0
krydser	cross (pres.)	3	0
kasse	box	2	0
kavalerikaserne	cavalery barracks	2	0
kendemærke	characteristic	2	0
kilometer	kilometer	2	50
klippen	the cliff	2	50
klippeskredet	the avalanche	2	0
knæk	crack	2	0
kornmarkerne	The fields	2	0
kulminerne	the coalmines	2	0
kurve	curve	2	0
kørt	driven	2	0
lokaliseret	located	2	0
startkrydset	the starting intersection	2	0
sydkanten	the southern edge	2	0
trekant	triangle	2	0
udkant	outskirts	2	0
velbekomme	please	2	0
østkanten	the eastern edge	2	0

Full dataset for (k) - continued			
Word form	English	sum	% reduced
ankom	arrived	1	0
café	café	1	0
caféen	the café	1	0
chikane	harassment	1	0
fokus	focus	1	100
forkerte	wrong (pl.)	1	0
forklaring	explanation	1	0
ikon	icon	1	0
k	k	1	0
kager	cakes	1	0
kaktusplante	cactus plant	1	0
kalde	call (inf.)	1	0
kaldt	called	1	0
kassen	the box	1	0
katten	the cat	1	0
kendetegn	characteristic	1	0
kendingsmærker	hallmarks	1	0
kendt	known	1	0
kigget	looked	1	0
kikkert	binoculars	1	0
kinden	the cheek	1	0
kirke	church	1	0
klat	speck	1	0
klippehavens	the rock garden's	1	0
klippehaver	rock gardens	1	0
klokket	messed up	1	0
klosters	monastery's	1	0
klostre	monasteries	1	0
kløften	the ravine	1	0
knæ	knee	1	0
knækket	cracked	1	0
kolonihaveforening	allotment union	1	0
kolonihaveområdet	allotment area	1	0
kompasretningerne	the compass corners	1	0
kompasset	the compass	1	0
kompromis	compromise	1	0
koordinatsystem	coordinate system	1	0
kornmarken	the field	1	0
koster	cost	1	0

Full dataset for (k) - continued			
Word form	**English**	**sum**	**% reduced**
krystalklar	crystal clear	1	0
kunnet	have could	1	0
kurs	course	1	0
kurven	the curve	1	0
kvarteret	the neighborhood	1	0
kvægindhegningen	the cattle enclosure	1	0
kysten	the coast	1	0
kæmpebue	giant arch	1	0
købmanden	the grocer	1	0
køre	drive	1	0
papirkanten	the paper's edge	1	0
parkerer	park (pres.)	1	0
parkeringsplads	parking space	1	0
risiko	risk	1	0
sekund	second	1	100
selvklart	self-evident	1	0
socialklasser	social classes	1	0
trekvart	three quarters	1	0
underkanten	the underside	1	0

Full dataset for (d)			
newword	**English**	**sum**	**% deleted**
et	a/an	429	14
øst	east	327	5
vest	west	270	5
lidt	little	251	3
godt	good	233	6
at	to/that	175	59
rundt	around	135	11
sydvest	south west	105	8
sydøst	south east	87	6
nordvest	north west	74	5
helt	all	71	51
langt	far	64	28
nordøst	north east	55	4
start	start	51	4
udsigtspunkt	vantage point	43	9
kort	short	42	2
forladt	abandoned	39	3
skråt	askew	36	0
vagtpost	guard post	32	6
midt	middle	29	7
kløft	quite	28	0
ret	my	28	4
mit	right	27	4
rigtigt	your	26	73
dit	south west	23	13
bjergkløft	south east	22	5
først	first	22	27
set	seen	18	17
fint	fine	17	47
højt	high	17	6
hvert	eash	17	94
nærmest	nearly	16	25
skarpt	sharply	16	56
stråtækt	thatched	16	25
punkt	point	15	40
tidspunkt	point in time	15	60
nyt	new	14	7
umiddelbart	immediately	13	0
halvt	half	11	27
mest	Mostly	11	27

Full dataset (d) - continued			
Word form	**English**	**Sum**	**% deleted**
sikkert	surely	11	18
snart	soon	11	18
afbrændt	scorched	10	10
fort	fort	9	11
let	easy	9	0
nordligt	northern	9	11
vent	wait	9	89
stort	big	8	13
gjort	done	7	14
glemt	forgotten	7	29
stakit	picket fence	7	0
alt	everything	6	17
galt	wrong	6	67
nordnordvest	north-northwest	6	0
startpunkt	startingpoint	6	0
tæt	close	6	0
farligt	dangerous	5	20
svært	difficult	5	40
vestligt	western	5	40
vist	(particle)	5	40
åbenbart	apparently	4	25
forkert	wrong	4	50
fortsat	continued	4	0
kant	edge	4	0
korrekt	correct	4	0
nordnordøst	north-northeast	4	0
slet	delete	4	25
sydligt	southern	4	25
østligt	eastern	3	67
øverst	top-most	3	33
bortset	except	3	100
fast	firm	3	33
granitbruddet	the granite quarry	3	33
højst	at most	3	67
parallelt	parallel	3	33
rart	nice	3	0
sagt	said	3	67
sandt	true	3	0
svagt	faint	3	33
vidt	widely	3	0
østvest	east-west	2	0

Full dataset for (d)			
Word form	**English**	**sum**	**% deleted**
delvist	partly	2	50
enkelt	simple	2	50
fat	hold	2	0
forskudt	slanted	2	0
fremmest	foremost	2	50
grøft	ditch	2	50
højdemæssigt	hieght-wise	2	0
helst	rather	2	100
kørt	driven	2	0
længst	furthest	2	50
muligt	possible	2	50
rent	clean	2	50
sæt	set	2	0
sat	sat	2	0
smalt	narrow	2	0
spist	eaten	2	0
trekant	triangle	2	0
udkant	outskirts	2	50
vandret	horizontally	2	0
vestnordvest	west-northwest	2	0
yderst	outermost	2	50
ærgerligt	pity	1	100
østnordøstagtigt	est-northeastish	1	0
agtigt	ish	1	0
allernederst	bottom-most	1	0
art	kind	1	0
behageligt	pleasant	1	0
blødt	soft	1	0
bort	gone	1	0
brugt	used	1	0
centimet	centimete	1	0
dårligt	bad	1	100
dejligt	lovely	1	0
delt	shared	1	0
diamant	diamond	1	0
dobbelt	double	1	100
egentligt	actually	1	100
eventuelt	perhaps	1	100
forfærdeligt	terrible	1	100
fuldstændigt	completely	1	100
granit	granite	1	0

Full dataset for (d) - continued			
newword	**English**	**sum**	**% deleted**
hårdt	hard	1	0
haft	had	1	100
herligt	glorious	1	100
holdt	held	1	100
idiot	idiot	1	0
interessant	interesting	1	0
kaldt	called	1	0
kendt	known	1	0
kikkert	binoculars	1	0
klat	speck	1	0
klippeskredet	the avalanche	1	0
kolonihaveområdet	the allotment area	1	0
lodret	vertical	1	0
mægtigt	magnificent	1	0
modsat	opposite	1	0
nævnt	mentioned	1	100
nøjagtigt	exactly	1	100
nemmest	easiest	1	0
omvendt	reverse	1	0
ordentligt	proper	1	100
pænt	nice	1	100
præcist	precisely	1	0
pragtfuldt	marvelous	1	0
proviant	goods	1	100
relativt	relatively	1	0
rendt	ran (into)	1	0
særligt	special	1	100
sekund	second	1	0
selvklart	self-evident	1	0
sendt	sent	1	0
sidst	last	1	0
sjovt	funny	1	100
skridt	step	1	0
smart	clever	1	0
sort	black	1	0
spændt	excited	1	0
sydsydøst	south-southeast	1	0
sydsydvest	south-southwest	1	0
sydvestligt	southwesterly	1	0
tænkt	thought	1	100
tørt	dry	1	0

Full dataset for (d) - continued			
newword	**English**	**sum**	**% deleted**
tabt	lost	1	100
teateret	the theater	1	100
tekst	text	1	0
ticentimeterspunkt	ten centimeter point	1	0
tidligt	early	1	100
trekvart	three-quarters	1	0
trist	sad	1	0
tydeligt	clearly	1	100
udgangspunkt	point of departure	1	0
udsigtskløft	vantage spot	1	0
vådt	wet	1	100
vildt	wild	1	0

Appendix B – Word forms particular to the DanPASS Dialogues

This appendix gives, for each variable, a list of the word forms which were excluded from the analyses of word frequency effects, because the word forms do not occur in the corpus on which the frequency counts are based, the LANCHART corpus.

Word forms excluded from the dataset for (ɑj)			
Word form	**English**	**sum**	**% monophthongized**
allright	all right	1	0
flodlejet	the river bed	2	100
fortegnelse	index	1	0
gelejdet	ushered	1	0
Ibsvej	proper noun	1	0
indhegning	fencing	13	0
indianerlejr	indian camp	4	0
indianerlejre	indian camps	1	0
indianerlejren	the indian camp	29	31
indtegner	draw	1	0
sejlklub	sailing club	3	33

Word forms excluded from the dataset for (ɒw)			
Word form	**english**	**sum**	**% monophthongized**
baneoverskæringen	railroad crossing	1	0
jernbaneoverskæring	railroad crossing	23	35
jernbaneoverskæringen	the railroad crossing	16	38
jernbaneoverskæringens	the railroad crossing’s	1	0
skovs	forest’s	1	0

Word forms excluded from the dataset for (p)			
Word form	**english**	**sum**	**% reduced**
bananpalme	banana palm tree	7	0
bananpalmen	the banana palm tree	25	0
bananpalmens	the banana palm tree’s	1	0
bananpalmer	banana palm trees	1	0
kompasrosen	the compass rose	3	0
løvepark	lion park	9	0
løveparken	the lion park	7	0
pakhuset	the warehouse	5	0
palmen	the palm tree	1	0
palmer	palm trees	1	0
papirets	the paper’s	2	50
papirhjørnerne	the paper corners	1	0
parkens	the park’s	1	0
passere	pass (inf.)	26	0
perspektivet	the perspective	1	0
præcisere	clarify	1	0
pukkel	hump	1	0
startpunkt	startingpoint	3	0
startpunktet	the startingpoint	17	6
Teaterpassagen	the theater passage	1	0
ticentimeterspunkt	ten centimetre point	1	0
totempælens	the totem pole’s	1	0

Word forms excluded from the dataset for (b)			
Word form	**English**	**sum**	**% reduced**
kalkstensklipper	limestone cliffs	9	22
kalkstensklipperne	the lime stone cliffs	4	25
klippehave	rock garden	15	20
klippehaven	the rock garden	8	25
klippehavens	the rock garden's	1	100
klippen	the cliff	1	0
klippeskred	avalanche	8	38
klippeskredet	the avalanche	1	0
sejlklub	sailing club	2	0
toppunktet	the top point	1	0

Word forms excluded from the dataset for (k)			
Word form	**English**	**sum**	**% reduced**
østkanten	the east edge	1	0
bjergkløft	ravine	11	0
bjergkløften	the ravine	7	14
caféen	the café	1	0
ikon	icon	1	0
kæmpebue	giant arch	1	0
kaktus	cactus	2	0
kaktusplante	cactus plant	1	0
kaktussen	the cactus	3	0
kalkstensklipper	limestone cliffs	9	0
kalkstensklipperne	the lime stone cliffs	4	0
kavaleri	cavalery	12	0
kavaleriet	the cavalery	9	0
kavalerikaserne	cavalery barracks	1	0
kendemærke	characteristic	1	0
kendingsmærker	characteristics	1	0
kikkert	binoculars	1	0
kløften	the ravine	1	0
klippehave	rock garden	15	0
klippehaven	the rock garden	8	0
klippehavens	the rock garden's	1	0
klippen	the cliff	1	100
klippeskred	avalanche	8	0
klippeskredet	the avalanche	1	0
klokket	messed up	1	0
klosteret	the monastery	7	0
klosters	monastery's	1	0
kolonihaveforening	alottment union (gardens)	1	0
kolonihaverne	alottments (gardens)	8	0

Word forms excluded from the dataset for (k) – cont'd			
Word form	**English**	**sum**	**% reduced**
kompasretningerne	the compass directions	1	0
kompasrosen	the compass rose	2	0
kompasset	the compass	1	0
koordinatsystem	coordinate system	1	0
kornmarken	the field	1	0
kornmarker	fields	13	0
kornmarkerne	the fields	1	0
kortets	the map's	2	0
krystalklar	crystal clear	1	0
kulmine	coalmine	5	0
kulminen	the coalmine	17	0
kulminerne	the coalmines	1	0
kvægindhegningen	cattle enclosure	1	0
papirkanten	the paper's edge	1	0
sejlklub	sailing club	3	33
selvklart	self-evident	1	0
socialklasser	social classes	1	0
startkryds	starting intersection	2	0
startkrydset	the starting intersection	1	0
sydkanten	the south edge	1	0

Word forms excluded from the dataset for (g)			
Word form	**English**	**sum**	**% reduced**
øjesteblik	momentito	1	100
østnordøstagtigt	east-northeast-ish	1	0
byggelegepladsen	the playground	2	100
byggepladsen	the construction site	1	100
flugte	follow	2	0
halvcirkel	semi-circle	2	0
kaktus	cactus	2	0
kaktusplante	cactus plant	1	0
kaktussen	the cactus	3	0
kalksten	limestone	1	0
kikkert	binoculars	1	0
kornmarken	the field	1	0
kornmarker	fields	13	38
kornmarkerne	the fields	1	100
løvepark	lion park	7	0
løveparken	the lion park	8	13
logikken	the logic	1	0
mærkeligformet	strangely-shaped	1	0
mexicaner	mexican	3	67
nixen	nope	1	100
pakhuset	the ware house	5	40
parkens	the park's	1	0
perspektivet	the perspective	1	0
slurk	sip	1	0
tekstens	the text's	1	100
telefonboks	telephone booth	1	0
telefonboksens	the telephone booth's	1	0
udsigtskløft	vantage point (lit: view cliff)	1	0
vagtstuen	the guard room	1	0

Word forms excluded from the dataset for (d)			
Word form	**English**	**sum**	**% deleted**
østligt	eastern	3	67
østnordøstagtigt	east-northeast-ish	1	0
bjergkløft	ravine	17	0
diamant	diamond	1	0
fort	fort	9	11
højdemæssigt	altitudinally	2	0
kikkert	binoculars	1	0
nordligt	northern	9	11
nordnordvest	nort-northwest	1	0
selvklart	self-evident	1	0
startpunkt	starting point	5	0
sydligt	southern	3	33
sydvestligt	southwestern	1	0
ticentimeterspunkt	ten centimetre point	1	0