

Variability and change in Received Pronunciation

A study of six phonological variables in the speech of television newsreaders

Bente Rebecca Hannisdal

Dissertation presented in partial fulfilment
of the requirements for the degree
Doctor Artium

Department of English
University of Bergen

2006

Variability and change in Received Pronunciation

A study of six phonological variables in the speech of television newsreaders

Bente Rebecca Hannisdal

Dissertation presented in partial fulfilment
of the requirements for the degree
Doctor Artium

Department of English
University of Bergen

2006

Acknowledgements

I wish to thank the late professor Kurt Rydland for his invaluable support and encouragement in the early stages of this project. I also wish to express my sincerest thanks to professor Helge Sandøy at the University of Bergen for taking over the supervision of the study, and for providing excellent guidance throughout the development of the dissertation. My thanks are also due to professor Inger Moen at the University of Oslo for reading parts of the manuscript, and for her encouraging comments.

I am grateful to professor Clive Upton and Thor Sigurd Nilsen for listening to speech samples and providing detailed comments, and to Dr Catherine Sangster at the BBC Pronunciation Research Unit for her assistance in arranging interviews with BBC World newsreaders.

I would also like to thank the Faculty of Arts at the University of Bergen for financially supporting this research project.

Contents

List of tables	vi
List of figures	viii
Abbreviations	ix
1 Introduction	1
1.1 Aim and scope	1
1.2 The variables studied	4
1.3 Why study RP?	5
1.4 Why study broadcast speech?	7
1.5 Research questions	9
1.6 The structure of the dissertation	9
2 Aspects of RP	11
2.0 Introduction	11
2.1 The history of RP	11
2.2 Varieties of RP	16
2.3 The social meaning of RP	18
2.4 Defining RP	20
2.4.1 Introduction	20
2.4.2 The BBC accent	22
2.4.3 An abstract norm	22
2.4.4 Prestige	23
2.4.5 Social definitions	25
2.4.6 Phonological definitions	28
2.4.7 Non-localisability	29
2.4.8 Summary	33
2.5 The phonological core of RP	34
2.6 RP and broadcasting	35
3 Theoretical aspects	44
3.0 Introduction	44
3.1 Language variation and change	44
3.1.1 General	44
3.1.2 Lexical diffusion	47
3.1.3 Accent levelling	50
3.1.4 Gender	54
3.1.5 Style	60
3.1.5.1 Stylistic variation	60
3.1.5.2 Explaining stylistic variation	61
3.1.5.3 “Natural” phonological processes	63
3.1.5.4 Stylistic aspects of broadcast speech	64
3.2 The concept of ‘standard’	67
3.2.1 A spoken standard?	67
3.2.2 C-RP vs. n-RP	72
3.2.3 Summary	74
3.3 The concept of ‘norm’	75
3.3.1 What are norms?	75
3.3.2 Studying norms	76

3.3.3	A broadcast norm?	78
3.3.4	Summary	84
4	The variables	86
4.0	Introduction	86
4.1	The linguistic variable	86
4.2	Variability in RP	88
4.2.1	Types of variability	89
4.2.2	Changes in RP	92
4.2.3	RP versus Estuary English	96
4.3	Choice of variables	98
4.4	The variables studied	99
4.4.1	Cure lowering	99
4.4.2	GOAT allophony	104
4.4.3	R-sandhi	106
4.4.4	T-voicing	112
4.4.5	Smoothing	116
4.4.6	Yod coalescence	120
5	Methodology	125
5.0	Introduction	125
5.1	The speakers	125
5.1.1	Selecting a sample	125
5.1.2	The channels	130
5.1.3	Style	132
5.2	Data collection and analysis	134
5.2.1	Data collection	134
5.2.1.1	Collecting the data	134
5.2.1.2	Sample size	135
5.2.2	Analysing the data	136
5.2.2.1	Identifying variants	136
5.2.2.2	Conditioning factors	137
5.2.2.3	Transcription and identification	138
5.2.2.4	Auditory analysis	139
5.2.3	Quantification	141
5.2.3.1	Quantifying the results	141
5.2.3.2	Statistical analysis	142
6	Analyses and results	146
6.0	Introduction	146
6.1	CURE lowering	147
6.1.1	Introduction	147
6.1.2	The auditory analysis of CURE lowering	147
6.1.3	Identifying CURE words	148
6.1.4	Results	149
6.1.5	Comparison with pronunciation dictionaries	152
6.1.6	Summary	153
6.2	GOAT allophony	154
6.2.1	Introduction	154
6.2.2	The auditory analysis of GOAT allophony	154
6.2.3	Results	155
6.2.3.1	General	155
6.2.3.2	Phoneme split?	156

6.2.3	Summary	157
6.3	R-sandhi	158
6.3.1	Introduction	158
6.3.2	Linking (r)	159
6.3.2.1	General	159
6.3.2.2	Social factors	161
6.3.2.3	Linguistic constraints.....	164
6.3.3	Intrusive (r).....	169
6.3.3.1	General	169
6.3.3.2	Social factors	171
6.3.3.3	Linguistic constraints.....	174
6.3.4	Style and speech rate	178
6.3.5	Summary	181
6.4	T-voicing	182
6.4.1	Introduction	182
6.4.2	Linguistic conditions	182
6.4.3	Results	186
6.4.3.1	General	186
6.4.3.2	Social factors	187
6.4.3.3	Linguistic factors	190
6.4.3.4	Style and speech rate	196
6.4.4	Summary	198
6.5	Smoothing	199
6.5.1	Introduction	199
6.5.2	The auditory analysis of smoothing	200
6.5.3	Results	200
6.5.3.1	General	200
6.5.3.2	Social factors	202
6.5.3.3	Linguistic factors	205
6.5.3.4	Speech rate.....	207
6.5.4	Summary	208
6.6	Yod coalescence	209
6.6.1	Introduction	209
6.6.2	The auditory analysis of yod coalescence	210
6.6.3	Results	210
6.6.3.1	General	210
6.6.3.2	The lexical distribution of yod coalescence.....	212
6.6.3.3	Social factors	214
6.6.4	Summary	217
6.7	Covariation	217
7	Conclusions	221
7.0	Introduction	221
7.1	Summary of results.....	221
7.2	Channels	223
7.3	Gender	224
7.4	Style.....	225
7.5	Variation and change in RP.....	225
7.6	Updating RP.....	228
7.7	Pedagogical aspects.....	229
7.8	Further research.....	229
	References	232
	Appendix	246

List of tables

Table 5.1	List of speakers	127
Table 6.1.1	CURE lowering: total scores	149
Table 6.1.2	The lexical distribution of /ʊə/ and /ɔ:/ in CURE	150
Table 6.2	The use of GOAT allophony	155
Table 6.3.1	Linking (r): total scores	159
Table 6.3.2	Linking /r/: mean scores for channels	162
Table 6.3.3	Linking /r/: ANOVA analysis for channel	162
Table 6.3.4	Linking /r/: mean scores for gender	163
Table 6.3.5	Linking /r/: ANOVA analysis for gender	163
Table 6.3.6	Linking (r) according to the grammatical category of the surrounding words	164
Table 6.3.7	Linking (r) before proper nouns	165
Table 6.3.8	Linking (r) according to the weight of the following syllable	166
Table 6.3.9	Linking (r) according to the presence and absence of another /r/	167
Table 6.3.10	Linking (r) before clause boundary	168
Table 6.3.11	Intrusive (r): total scores	169
Table 6.3.12	Intrusive/r/: mean scores for channels	172
Table 6.3.13	Intrusive /r/: ANOVA analysis for channel	172
Table 6.3.14	Intrusive /r/: mean scores for gender	173
Table 6.3.15	Intrusive /r/: ANOVA analysis for gender	173
Table 6.3.16	Intrusive (r) before proper nouns	174
Table 6.3.17	Intrusive (r) according to the weight of the following syllable	175
Table 6.3.18	Intrusive (r) according to the quality of the preceding vowel	176
Table 6.3.19	Intrusive (r) according to the presence and absence of another /r/	177
Table 6.3.20	R-sandhi according to style	179
Table 6.4.1	T-voicing: total scores	186
Table 6.4.2	T-voicing: mean scores for channels	188
Table 6.4.3	T-voicing: ANOVA analysis for channel	189
Table 6.4.4	T-voicing: mean scores for gender	189
Table 6.4.5	T-voicing: ANOVA analysis for gender	190
Table 6.4.6	T-voicing according to word position	191

Table 6.4.7	Word-final t-voicing according to the length of the target words	193
Table 6.4.8	Lexical distribution of word-final t-voicing in words with $N \geq 10$	194
Table 6.4.9	Word-final t-voicing by stress	196
Table 6.4.10	T-voicing according to style	197
Table 6.5.1	Smoothing: total scores	201
Table 6.5.2.	Smoothing: mean scores for channels	203
Table 6.5.3	Smoothing: ANOVA analysis for channel	203
Table 6.5.4	Smoothing: mean scores for gender	204
Table 6.5.5	Smoothing: ANOVA analysis for gender	204
Table 6.5.6	Smoothing according to the morphological complexity of the target words	206
Table 6.5.7	Smoothing of <i>-r(e)</i> items according to the presence and absence of a following linking /r/	207
Table 6.5.8	Smoothing according to the presence and absence of a following pause ...	207
Table 6.6.1	Yod coalescence: total scores	211
Table 6.6.2	Yod coalescence: the lexical distribution of /tʃ, dʒ/ and /tj, dj/	212
Table 6.6.3	Yod coalescence: mean scores for channels	215
Table 6.6.4	Yod coalescence: ANOVA analysis for channel	215
Table 6.6.5	Yod coalescence: mean scores for gender	216
Table 6.6.6	Yod coalescence: ANOVA analysis for gender	217
Table 6.7.1	Overview of individual frequencies for variables 2-6	218
Table 6.7.2	Bivariate correlational analysis for variables 3-6	219

List of figures

Figure 6.3.1	Percentage frequency of linking /r/, individual speakers	160
Figure 6.3.2	Percentage frequency of intrusive /r/, individual speakers	170
Figure 6.4	Percentage frequency of [ʔ], individual speakers	187
Figure 6.5	Percentage frequency of smoothed [aə], individual speakers	202
Figure 6.6	Percentage frequency of /tʃ, dʒ/, individual speakers	212

Abbreviations

RP	Received Pronunciation
EE	Estuary English
C	any consonant
V	any vowel
LPD	<i>Longman pronunciation dictionary</i> (Wells 2000)
EPD	<i>English pronouncing dictionary</i> (Jones 2003)
ODP	<i>Oxford dictionary of pronunciation</i> (Upton et al. 2001)

1 Introduction

1.1 Aim and scope

Received Pronunciation, or RP, is probably the most thoroughly described accent of English. There has, however, been very little quantitative empirical research into the accent, both phonological and sociolinguistic. The present study is an empirical investigation of current RP as spoken by television newsreaders.

RP is widely perceived as more static and homogenous than other varieties, and the entrance of new features is often interpreted as signalling the death of the RP accent. The view underlying the present work is that there exists a living variety that can be identified as RP, and that, as all living varieties, it accommodates a substantial amount of variability, at the realisational as well as the systemic level. The present study aims to explore aspects of this variability from a variationist perspective, by a detailed analysis of a set of phonological variables. Some of the variables reflect ongoing change and recent developments, and an investigation into the current status of these features may give some indication of the direction in which RP is heading.

The empirical basis for the investigation is 30 hours of speech produced by 30 news presenters from three different television news channels. The spoken data have been analysed auditorily and quantified with statistical methods. The variables are correlated with intra- and extralinguistic factors, and the findings are interpreted in light of previous descriptions and general principles of variation and change. By providing concrete linguistic facts, the study can contribute to an updated description of RP based on empirical data.

The primary focus is thus RP, not broadcast speech as such. The investigation will, however, also explore the possible effects of the newscast situation on the linguistic behaviour of the speakers. Moreover, the three news channels will be compared in order to ascertain whether usage patterns are homogeneous across channels, or whether there are systematic differences that reflect 'station styles' (cf. Bell 1982).

In addition to the empirical aspects, the study addresses some theoretical issues related to RP as a linguistic variety and an object of study. These issues are often ignored or treated sketchily in many writings on RP. The most important such topic is perhaps the question of

how RP should be defined and delimited. Section 2.4 provides a critical review of the many definitions of RP, and argues in favour of non-localisability as an operational criterion. The theoretical discussion also involves an assessment of the social significance of RP and its status in British society in general and newscasting in particular. In this connection, the terms *standard* and *norm* are of particular relevance. RP is often referred to as a standard variety or a prestigious norm, without further clarification of the terms. Chapter 3 includes a detailed discussion of the two concepts and how they relate to RP. As a supplement to the theoretical treatment of this topic, a small-scale attitudinal study was carried out among a group of BBC newsreaders, with the aim of shedding some light on the presenters' linguistic choices and level of awareness about their own accent. The results of this study are reported in section 3.3.

One of the basic assumptions for the present investigation, then, is that RP is subject to variation and change and can be explored from a sociolinguistic viewpoint.

Variation is included already in the early descriptions of the accent. Daniel Jones, for example, incorporated a large number of variable forms in his *English pronouncing dictionary* from 1917. In recent decades, both the social status and the linguistic properties of RP have been undergoing change. Many of these developments are attributed to the radical changes in British society and the subsequent increasing influence of non-standard varieties, along with the rise of competing "standards" such as Estuary English (see further 4.2). Aspects of variation and ongoing trends in RP are described in detail in e.g. Wells (1982: 279-301, 1994a, 1997a), Cruttenden (2001) and Upton (2004). Empirical quantified research into this variability is, however, scarce.

A quantitative corpus-based investigation of modern RP speech raises certain difficulties with regard to the concept of 'speech community'. Unlike most other varieties, RP is not limited to a geographical region or a particular social group (cf. 2.4). This complicates the process of selecting a relevant speaker sample. The notion of speech community is generally used in empirical linguistics to define a unit of analysis within which to study variation and change. The concept of speech community is however in itself problematic, and there is not agreement as to how it should be defined (cf. Patrick 2002). The term typically implies a synthesis of residence and daily activity. It is, however, not clear whether the speech community is primarily a social or a linguistic object. Some definitions emphasise linguistic uniformity, while others focus on social membership and shared attitudes and evaluations. Many recent approaches propose that people can belong to several simultaneous speech communities, as a function of their class, gender, ethnicity, etc. (cf. Eckert 2000: 34-41).

The purpose of the notion of speech community in linguistic research is to provide a target for generalisation. In the case of the present study, a relevant community can be all television newsreaders, or all RP speakers. With regard to the former, they can be said to constitute a socially coherent group, in that they have the same occupation and income, similar education, etc., but they do not all speak alike. If, on the other hand, all RP speakers are defined as a speech community, the members obviously share the same accent, but they do not form a socially homogeneous group. The present study is limited in both directions to include only newsreaders who speak RP. The main focus, however, is RP, not broadcast speech. The newsreaders will therefore be compared not only with each other, but their characteristics will also be related to previous descriptions of RP which (presumably) are based on observations of a more socially diverse mass of speakers. The newsreaders are thus regarded as representatives of the RP-speaking community, but with a recognition of the limitations and conditions of the speech situation. Newscasting is a relatively formal format, and generalisations can therefore not automatically be made to RP speech in general or to casual conversational style.

In view of the homogeneous nature of the sample, which observations can be made with regard to variation and change in RP? A large part of the analysis deals with language-internal factors. A corpus of the present type and size is well suited for investigating the effects of the linguistic context and establishing patterns of usage. In addition, comparisons will be made between male and female presenters, and between the three channels. Some of the variables will also be analysed stylistically, by comparing usage levels in reading style and interview style. In order to establish whether a pattern of variation represents a change in progress, one should ideally have a sample of speakers from all age groups. Differences across generations can be interpreted as evidence of change in line with the ‘apparent time hypothesis’ (cf. Chambers 2002: 212).¹ Researchers can, however, also look to earlier accounts of the variety under investigation, as a way of contextualising the contemporary data and providing a reference against which the results can be interpreted. In the present study of broadcast RP, the findings will be compared with previous descriptions of the RP accent. Various accounts by Wells, Gimson, Upton and others, and major pronunciation dictionaries such as *Longman pronunciation dictionary* (LPD), *Cambridge English pronouncing*

¹ Age-stratified differences can, however, also reflect *age-grading*, which is a typical component of stable sociolinguistic variation. Speakers are commonly found to use more innovative forms as adolescents, and then become increasingly conservative in their speech as they age (cf. Eckert 1997b).

dictionary (EPD) and *Oxford dictionary of pronunciation* (ODP) will be referred to throughout.

1.2 The variables studied

The study investigates six phonological variables which represent different aspects of variation and different stages in the development of RP. The variables are outlined briefly here. They are described in more detail in section 4.4.

1. CURE lowering, which involves the increasing use of /ɔ:/ instead of traditional /ʊə/ in words such as *sure*, *pure*, *tourist*, etc.
2. GOAT allophony, which refers to the realisation of /əʊ/ as [pʊ] before non-prevocalic /l/ in *goal*, *hold*, *shoulder*, etc.
3. R-sandhi, which comprises the use of linking and intrusive /r/ to avoid vowel hiatus, in phrases such as *here/r/ and there*, *law/r/ and order*, etc.
4. T-voicing, which involves the realisation of intervocalic /t/ as a voiced tap rather than a fortis plosive, in e.g. *getting*, *British*, *a lot of*, etc.
5. Smoothing, which refers to the reduction of the sequences /aɪə, aʊə/, as in *fire* and *power*, to [aə] or [a:].
6. Yod coalescence, which involves the coalescent assimilation of /tj, dj/ to /tʃ, dʒ/ in stressed environments, as in *tune*, *studio*, *endure*, etc.

Most of the variables have been selected on the basis of previous descriptions which mention these as areas where RP is known to display variation or undergo change. The exceptions are GOAT allophony and t-voicing, which are normally not mentioned in the RP literature, and which were included on the basis of my own observations of RP speech. However, none of the features have previously been empirically investigated in RP, and their current status in the accent was therefore largely unknown.

1.3 Why study RP?

As mentioned above, RP is probably the most extensively described accent of English. Corpus-based empirical work on the accent has, however, been lacking. The RP tradition is dominated by informal, impressionistic observations, or introspection. Jones writes in the first edition of his *English pronouncing dictionary* (1917):

The great majority of the facts recorded in this dictionary were ascertained by my personal observation. In the case of words with which I was unfamiliar, or about the pronunciation of which I was uncertain, I took every opportunity that presented itself of listening to the pronunciation of persons accustomed to use them. Only when I was unable to get an opportunity of hearing words actually used by persons familiar with them, did I resort to second-hand evidence. Even then I was careful to avoid accepting the evidence of any but those whom I had good reason to consider competent observers. (Jones 1917: vii)

The same note is made in the 13th edition, where he states that the pronunciation described in the dictionary “is in the main that which I use myself. I have, however, put my pronunciation in a secondary place in all cases where another form appears to me to be in more frequent use” (Jones 1967: xvii). The descriptive tradition from Daniel Jones has been continued by scholars such as Gimson (e.g. 1962, 1980) and Wells (e.g. 1982, 2000), who both provide exhaustive and detailed accounts of the accent.

Even though they do not refer to quantified data, there is no reason to suspect that the earlier descriptions are in any way inaccurate or imprecise. Indeed, these accounts are the foundation of what has come to be generally accepted and embraced as RP, and thus constitute the basis for all subsequent studies of the accent. Wells remedies some of the empirical lack in his *Longman pronunciation dictionary* (1990a, 2000), which includes data from opinion polls conducted among a large number of native British English speakers (cf. Wells 1999a). However, these are written surveys where people report on their pronunciation *preferences*, and thus not recordings of actual usage.

There is a small number of empirical studies of variation in RP conducted in the 1980s, such as Ball (1984) and Bauer (1984). These are, however, based on a very limited amount of data. The only large-scale corpus-based empirical study of current RP is Fabricius’

(2000) sociolinguistic investigation of t-glottalling among young RP-speaking Cambridge students. In addition, there have been a number of studies into language attitudes where RP is included as one of the varieties that are evaluated (cf. e.g. Giles et al. 1990, Garrett et al. 2003, Elyan et al. 1978).

The present study is then an attempt to fill some of the empirical gap by supplying quantitative corpus-based data on current RP usage. The RP accent has been somewhat neglected in the field of language variation and change, where the focus tends to be on non-standard varieties. There is, however, every reason to assume that RP incorporates the same variability and potential for change as any other living variety. Fabricius (2002b: 356) comments:

since it is obviously true that RP speakers are just as much a part of the British speech community as everyone else, the social characteristics and linguistic features displayed by these speakers provide a basis for empirical study.

While RP can be regarded as a speech variety just like any other, and can be approached as such, it is evident that the accent has a unique status, as a result of its codification. RP is the British English model for most pronunciation dictionaries and textbooks on phonetics, and the pronunciation norm for foreign learners of English in many parts of the world. The current social and linguistic status of RP is therefore of special relevance to language teachers and learners in an EFL setting. Moreover, in the field of accent studies, RP is widely used as a reference accent with which other (British) varieties are compared. All these aspects underline the need for updated descriptions of RP based on empirical data. The codification of RP carries with it the danger of perceiving the accent as static and uniform, and the phonetic properties of RP are thus often taken for granted. Wells emphasises the importance of keeping the descriptions of RP in line with current usage, and states that “it must not remain fossilized in the form codified by Daniel Jones almost a century ago” (Wells 1997b).

1.4 Why study broadcast speech?

People in Western countries probably hear more language from the media than they do directly from the lips of their fellow humans in conversation.

Bell (1991: 1)

The present work focuses on RP as spoken by television newsreaders. There are many reasons why broadcast speech is a useful and well-suited locus for observing current pronunciation. It offers several practical advantages over conversational speech, the most obvious of which is availability: broadcast speech is readily accessible and easy to collect. Also, the amount of speech available is virtually infinite. Moreover, recordings directly from television or radio are generally of a high technical quality and not marred by the background noise or interference that is typical of face-to-face recordings. Lastly, the problem of the ‘observer’s paradox’ (Labov 1972: 209) – how to observe the way people speak when they are not being observed – is not an issue, since broadcast speech is by nature intended to be “observed” by a large number of people, and the speaker’s performance will not be affected by the fact that someone is recording the broadcast.²

There is a long tradition for correlating RP with broadcasting, and with newscasts in particular (see further 2.6). Within the broadcast media, as Bell (1991: 1) points out, news is “the primary language genre” and “is seen by both media organizations and audiences as the focus of media content”. Newscasting is one of the few areas where RP still dominates, and therefore a valuable source for systematic observations of the accent. A number of descriptions of RP refer to the pronunciation of newscasters. Lewis (1975), for instance, bases his account of r-sandhi on “over ten years of careful systematic observation of ... fairly large numbers of speakers, in particular of scores of newsreaders employed in British national broadcasting” (37). Similarly, Brown (1990) describes a number of RP speech patterns, and most of her examples are taken from radio and television newscasts. Moreover, the British English variety described in pronunciation dictionaries is typically equated with the accent of radio and television newsreaders (cf. ODP: xi, EPD: v). Newscast speech is thus the basis for many statements about RP. However, the empirical foundations consist primarily of informal or random observations, and not systematic analyses of corpus data. The present study can

² The public nature of newscast speech in itself is, however, believed to have a fundamental influence on the level of speech-consciousness and, ultimately, on the linguistic choices of the presenters. These aspects are further discussed in 3.1.5 and 3.3.

contribute to a more firmly grounded empirical basis, by providing quantified data and detailed analyses of broadcast speech that has been systematically collected.

Apart from conveniently providing RP-speaking informants, broadcast media are a valuable source also in other respects for anyone conducting research into pronunciation. Broadcast media are perhaps the most dominating producers and presenters of language in our society (cf. Bell 1991: 1). Broadcast speech is the ultimate public language and is heard by mass audiences, and broadcasters are the ultimate language professionals. The media's linguistic choices largely reflect usages and evaluations of society as a whole, but may also influence general attitudes towards linguistic varieties (see further 2.6). Insight into the nature of broadcast speech should therefore be of intrinsic interest to language users and receivers in a society where broadcast media play a dominant social role.

Traditionally, the object of primary interest to sociolinguists has been the spontaneous, informal, *vernacular* speech, and broadcast language has been regarded as inferior to conversational data as a target of linguistic research (cf. Labov 1972: 211). James Milroy (2001b: 30) even argues that television and radio speech is not part of the speech community. The vernacular so eagerly sought by most sociolinguists has been described variously as the variety "in which the minimum attention is given to the monitoring of speech" (Labov 1972: 208) and as the low-status variety characteristic of a social group or a geographical area (cf. Milroy 1987: 58). It is claimed that the vernacular conveys "the best insight into the natural speech of a community" (Crystal 2003: 491), as it offers the best data for observing the structural characteristics of a variety or for investigating processes of linguistic change. The idea is that the vernacular is more regular and consistent because it is least influenced by notions of linguistic "correctness" and thus not exposed to potential hypercorrections in the direction of the standard (cf. Labov 1972: 248, Chambers and Trudgill 1998: 48). The vernacular is thus seen in contrast to standard or high-status speech, which is associated with formal styles. The vernacular has also been defined as "the language of locally based communities" (Eckert 2000: 17), and thus contrasted with supralocal speech. While the definitions of *vernacular* are not uniform, they share an underlying reference to non-standard (in the sense of publicly and institutionally unrecognised) varieties.

The subject of the present study, broadcast RP, would then seem to represent the exact opposite of the vernacular: a non-regional, prestigious variety, produced in a formal, public setting, with a high level of self-monitoring. It is however no less "natural" or "real" than localised speech produced in a casual, everyday context. While there is no doubt that data representing spontaneous conversational speech are useful for investigating many

sociolinguistic issues, the vernacular is ultimately an abstraction, and, as Milroy and Gordon point out, “the concept of an entirely natural speech event (or an entirely unnatural one) is untenable” (2003: 50). Wolfson (1997) argues that there is no single entity corresponding to the notion of natural speech, and that “if speech is felt to be appropriate to a situation and the goal, then it is natural in that context” (1997: 124). An examination of broadcast RP will provide a unique insight into speech intended for mass consumption, produced in a context where, presumably, the focus on “correctness” is great.

1.5 Research questions

The present study, then, is concerned with the following questions, which have both empirical and theoretical significance.

- What is the current status of variables 1-6 in the speech of RP-speaking news presenters?
- How close is current broadcast RP to the descriptions in the literature?
- Is RP approaching *Estuary English*?
- Are the pronunciation patterns similar across the three channels, or is there evidence of ‘station styles’?
- Are there systematic pronunciation differences between male and female speakers?
- How does attention level influence pronunciation patterns?
- How should RP be defined?
- How does RP relate to concepts such as ‘standard’ and ‘norm’?
- Why do so many newsreaders speak RP?

1.6 The structure of the dissertation

The dissertation consists of seven chapters which cover various aspects of the study. The introductory chapter presents the motivation and objectives of the study. Chapter 2 discusses various aspects of RP, including its historical origins and social connotations. The main part of this chapter is however devoted to the problem of defining the accent, which is of crucial importance for the validity of the results. It examines a range of possible defining criteria,

with a focus on how RP should be delimited as an object of study. Chapter 3 provides a more general theoretical background for the present investigation. It examines some key concepts within the field of variationist research, with an eye to how linguistic variation and change can be explained. This chapter also explores notions such as ‘standard’ and ‘norm’, and how they relate to RP. The last section of chapter 3 presents the results from a small-scale attitudinal survey conducted among a group of BBC news presenters. Chapter 4 looks at variation and change in RP, and gives a detailed presentation of the six variables that constitute the main focus of the study. Chapter 5 describes the methodological procedures employed in the collection and analysis of data, and discusses the empirical and theoretical implications of the methodological choices. The results of the analyses are presented and discussed in chapter 6. Chapter 7 provides a summary of the findings and assesses the results in light of the theoretical bases established in previous chapters and the objectives of the study.

2 Aspects of RP

2.0 Introduction

Received Pronunciation, or RP, is the linguistic term for the British English accent which has traditionally served as a prestige variety and as a pronunciation model in the teaching of English as a foreign language. Common layman's terms for RP include *The Queen's English*, *Public School Accent*, *Oxford English*, *BBC English*, *Standard English*, *talking without an accent*, *talking posh*, etc. These expressions all reflect important historical and social aspects of RP, as the accent of the Court and the upper classes, the accent of the educated, the accent used by presenters and newsreaders in the BBC, a codified accent that serves as a standard for foreign learners of English, an accent that conceals the regional background of the speaker, and an accent that for some is associated with arrogance and social pretension. All these aspects will be discussed in the present chapter, which includes an outline of the history of RP and the factors that led to its status as a non-regional norm, a presentation of its phonological ingredients and its social connotations, and a discussion of the problems involved in defining the variety.

2.1 The history of RP

Accent and Pronunciation must be diligently studied by the conversationalist. A person who uses vulgarisms will make but little way in good circles ... A proper accent gives importance to what you say, engages the respectful attendance to your hearer, and is your passport to new circles of acquaintance.

Anonymous (*Talking and Debating*, 1856)

The development of RP and its unique position in British society is closely linked to the rise of accent as a social signifier and the wish to establish a standard for spoken language.

The historical origins of an English speech standard are commonly traced back to the 16th century (cf. Gimson 1977, Mugglestone 1995, Nevalainen 2003) when prestige became attached to one type of pronunciation. According to Honey (1985) this development started as

early as the 15th century, with the emerging predominance of a variety which was “a fusion of South Central Midlands influences with existing London speech forms” (1985: 211). For political and economic reasons, it was the educated speech of the capital and the surrounding areas which emerged as the high-status variant. The fact that Britain’s central government, trade and fashion were mainly concentrated on the capital contributed to making the London accent widely understood throughout the country. Moreover, it was the pronunciation of the upper social ranks that provided the model for spoken language, which is in line with Haugen’s observation that “if a recognised élite already exists with a characteristic vernacular, its norm will almost inevitably prevail” (Haugen 1997: 349). The 16th century pronunciation norm was neither fixed nor codified, but it was ‘focalised’ in social and regional terms (cf. Nevalainen 2003: 135). It became a social norm associated with the upper classes in the southeast, and later in the whole of England.

During the 18th century, there was a growing preoccupation with spoken language and issues of correctness and purism, and with establishing a standard for “good” or “correct” speech. This trend was related to the increasing power and prosperity of the middle classes, whose members wished to erase all traces of their working-class origins in their speech. There existed therefore a large and highly receptive audience for dictionaries and manuals which showed people how to get rid of any “vulgarisms” and adopt a “correct” pronunciation of English (cf. MacMahon 1998: 383). In 1780 Thomas Sheridan published *A General Dictionary of the English Language*, where he encouraged the imitation of the speech patterns of “people of education at court” (quoted in Nevalainen 2003: 147). Increasingly, though, it was the speech of the learned and educated, rather than that of the Court, which was recommended as the standard. John Walker published *A Critical Pronouncing Dictionary and Expositor of the English Language* in 1791, and his norm was based on “good usage”: “those sounds ... which are the most generally received among the learned and polite, as well as the bulk of speakers” (quoted in MacMahon 1998: 387).

The emerging speech standard remained a social norm in the southeast of England, but did not become widespread until the late 19th century, as the notion of a completely non-localised standard accent became increasingly dominant. Before that, it was common for the members of the aristocracy to speak with the accents of their regions. The extensive geographical and social diffusion of the pronunciation standard was made possible primarily by the public school system, which was a nationwide network of residential schools for children of the upper and upper-middle classes. Honey (1985) dates the emergence of the new public school system to around 1870, and at about the same time, the term *Received*

Pronunciation was used for the first time to describe the standard speech form. It was commonly expected that all upper- and upper-middle class boys should be educated at a public school, and a new caste was created in British society: the ‘public school man’, whose primary hallmark was a specific accent. The public school system, and by extension the universities of Oxford and Cambridge, had an enormous influence in promoting RP and establishing it as the most prestigious spoken variety. Honey (1991: 17) writes that “it was, more than anything else, the emergence of an educated class that gave impetus to the development and spread of a standard accent”. The conformist practices of the public schools spread to all sections of the education industry, and had an eroding effect on local speech forms. An RP accent became the primary superficial marker of social standing, and it was no longer acceptable for members of the upper classes to speak with a regional accent. At the start of the 20th century the increasing influence and availability of RP had created a situation which, according to Honey (1985: 244) “is perhaps unique in Western Europe, whereby one variety of spoken language enjoyed overwhelming predominance in terms of general social acceptability”.

In the 1920s the dissemination and popularisation of RP was to become even wider with the introduction of sound broadcasting in 1922, followed by television in the 1930s. By using only RP speakers as announcers and newsreaders, the BBC underlined the social importance of the accent, and in the public mind RP became even closer linked with high status and intellectual competence (see further 2.6).

The antecedents of the term *Received Pronunciation* can be found in the work of Alexander Ellis (1869-89) who defines *received* (in the now obsolete sense ‘socially accepted’) pronunciation as “the educated pronunciation of the metropolis, of the court, the pulpit and the bar” (23). A few decades later Henry Cecil Wyld (1920: 2-3) introduced the term *Received Standard English* and described the accent as

the product of social conditions, and ... essentially a Class Dialect. Received Standard is spoken, within certain social boundaries, with an extraordinary degree of uniformity, all over the country. ... It has been suggested that perhaps the main factor in this singular degree of uniformity is the custom of sending youths from certain social strata to the great public schools. If we were to say that Received English at the present day is "Public School English", we should not be far wrong.

Daniel Jones, who played an important role in codifying RP and promoting the use of the term, originally labelled the accent *Public School Pronunciation*, because of its close association with the public schools. In the first edition of his *English pronouncing dictionary* (Jones 1917: viii) he writes:

The pronunciation represented in this book is that most usually heard in everyday speech in the families of Southern English persons whose men-folk have been educated at the great public boarding-schools. This pronunciation is also used by a considerable proportion of those who do not come from the South of England, but who have been educated at these schools.

In the third edition of the dictionary (1926) he changed the label to *Received Pronunciation*, which was to become the common term used by phoneticians. There have been scattered attempts at introducing new names to replace the now somewhat obsolete term *Received*. Roach and Hartman (Jones 1997) suggest *BBC English*, while Lewis (1972, 1985) prefers the term *General British*, as a parallel to General American. Other names that have been used include *English Standard Pronunciation* (Trim 1961, referred in Lewis 1985: 249), *Southern British Standard* (Wells and Colson 1971) and *Standard Southern British* (Shockey 2003). However, none of the alternative terms have caught on, and *Received Pronunciation* continues to be used, although, according to Lewis (1985: 251), “it remains a technical term of linguistics”.

The historical base of RP was educated southeastern English pronunciation as used by the upper classes. However, as Milroy (2001a: 26-27) points out, it is doubtful that the development of RP is just a simple continuation of the highest class accent:

the view ... that RP comes down in a straight line from earlier English courtly usage is somewhat over-simplified. ... There is little reason to suppose that we are dealing with the unilinear history of a continuous upper-class variety, as from a sociolinguistic point of view such a unilinear history is intrinsically unlikely. High prestige features can lose prestige over time, and low prestige features can gain prestige.

Furthermore, through the public school system, access to education and social advancement spread well into the middle classes, thus, “a middle-class, rather than an upper-class focus should perhaps be expected in early RP” (Milroy 2001a: 24-25). There is indeed evidence that

in the course of the 19th century RP adopted some features that had their origin in low-status varieties, such as the open back [ɑ:] in *ask*, *bath*, *fast*, etc. (see Muggleston 1995).

Up until the middle of the 20th century RP reigned supreme as the unrivalled English pronunciation standard. But in the decades after the Second World War Britain underwent radical social changes which also left their marks on the linguistic development and on the attitudes towards accent. Along with the general social changes, the role of RP also changed considerably. Between 1944 and 1966 the number of universities in Britain doubled and higher education became available to people from diverse social backgrounds. The increased democratisation in the educational system extended into the occupational and public life. Professional and academic careers became open to people from the lower social strata, who of course were non-RP speakers. Regional accent features “massively invaded the realms of the social élite” (Wotschke 1996: 221) and the hegemony of RP was broken. An educated speaker was no longer synonymous with an RP speaker, and RP was no longer the exclusive property of a narrow social class. In 1970 Gimson wrote:

The acceptance of the BBC accent, i.e. some form of RP, as a standard can no longer be said to be common amongst the younger people. The social structure of the country is much less rigid than it was forty years ago, and the young are particularly apt to reject authority of any kind. This general rejection includes the accent of the “Establishment”, i.e. RP. (Gimson 1970: 18-19)

Modern RP is still associated with education and social status, and “widely regarded as a model for correct pronunciation, particularly for educated formal speech” (Wells 2000: xiii). Jones’ *English pronouncing dictionary*, with RP as the model, came out in its 16th edition in 2003, and RP is still the British English accent taught to foreign learners of English. However, with the increasing democratisation of British society, RP has lost its former unique position, and non-RP accents are now heard in many contexts from which they were previously excluded. RP may now even be a disadvantage in many contexts, as it carries connotations of social exclusiveness and pretension (cf. 2.3). Hughes and Trudgill (1996: 9) describe the current situation in the following way:

It is sometimes said that nowadays there is not the same pressure as there once was to modify one’s speech in the direction of RP. Reference is made to the fact that announcers with non-RP accents are now to be heard on the BBC, that important posts

in industry and the civil service are held by non-RP speakers, and that some younger RP speakers have adopted, more or less deliberately, features of regional pronunciation. Perhaps the pressure is less, but it is still there.

2.2 Varieties of RP

RP, like all accents, changes over time and incorporates new features while others are lost. As a result, different subtypes of RP can be distinguished, which are more or less conservative or progressive, and which are spoken by different age groups. Gimson (1980) identifies three main types of RP: *conservative* RP, “used by the older generation and, traditionally, by certain professions or social groups”, *general* RP, “most commonly in use and typified by the pronunciation adopted by the BBC”, and *advanced* RP, “mainly used by young people of exclusive social groups” (1980: 91). The latter type presumably incorporates any innovations or new trends that may or may not become a permanent part of RP, and “may indicate the way in which the RP system is developing and be adopted in the future as general RP” (92). The distinction between general and advanced RP has proved not to be very fruitful and is abandoned in Cruttenden (1994) and (2001), where general RP is contrasted with *refined* RP, which is “commonly considered to be upper-class” and associated with “upper-class families and with professions which have traditionally recruited from such families” (2001: 80). Wells (1982: 279-283) makes a similar distinction between *Mainstream* RP and *Upper-crust* RP (U-RP). The latter is the more conservative and old-fashioned type of RP, popularly associated with an elderly Oxbridge don, an upper-class army officer, or the older members of the royal family. Mainstream RP corresponds to general RP and is the unmarked, neutral, modern type of RP, typically spoken by BBC newsreaders. Honey (1985) uses the terms *marked* and *unmarked* RP to make a similar distinction between the conservative and socially conspicuous RP and the modern and socially neutral variety.

In addition to the two main types, it is common to distinguish between RP and what Wells (1982) calls *Near-RP*: strongly modified regional accents which are close to Mainstream RP, but which include a few regional pronunciation features, and therefore do not fall completely within the boundaries of RP. Cruttenden (2001) uses the term *regional RP* to describe “the type of speech which is basically RP except for the presence of a few regional characteristics which go unnoticed even by other speakers of RP” (80).

Lastly, Wells (1982) distinguishes a fourth category of RP which he calls *adoptive RP* and which is the variety of RP “spoken by adults who did not speak RP as children” (283). The difference between native and adoptive RP is then the age at which a speaker acquires the accent. Phonologically, adoptive RP may be indistinguishable from Mainstream RP and will not be recognised as adoptive RP if a person learns to speak exactly like a native RP speaker. An adoptive RP speaker will, however, often retain traces of his native accent, or use RP only in certain situations, or use a very careful, deliberate type of RP.

The various types of RP are not clearly distinct accents with precisely enumerable lists of features, and there are no categorical boundaries between them. They are rather tendencies, and represent clusterings of features, “such clusterings varying from individual to individual” (Cruttenden 2001: 80). There are nevertheless features which are typical for the different varieties, and some of which are outlined below (cf. Wells 1982: 279-301).

Near-RP (or regional RP) can be any accent which is essentially RP, but which consistently or variably contains regional features that enable a phonetician to locate the speaker geographically. Exactly how many regionalisms a speaker must have before his accent is no longer near-RP, but a regional accent, is of course impossible to determine in empirical terms, in the same way that phoneticians disagree over where the boundary between near-RP and RP should be drawn. Upton et al. (2001), for instance, accept short front [a] in BATH-words as part of modern RP, whereas Wells (2000) considers this feature non-RP or near-RP.

U-RP (or refined RP) is characterised by realisations like [ɛʊ] or [oʊ] for /əʊ/ in GOAT words, [ɛ ~ eæ] for /æ/ in TRAP, a long /ɔ:/ in CLOTH, very open word-final /ɪ/ and /ə/, and no glottalling of /t/. The U-RP accent is typically accompanied by a special voice quality and manner of delivery. This type of speech is distinctly upper-class, and is today often considered affected and obsolete.

Adoptive RP speakers, according to Wells, typically lack control of the informal and allegro characteristics of RP. Native RP speakers make extensive use of assimilation, elision, /r/ sandhi and other features of connected speech, especially in informal contexts. Speech-conscious adoptive RP speakers tend to avoid such features, possibly because they regard them as “lazy” or “sloppy”, “and hence incompatible with their ‘best’ speech” (Wells 1982: 284).

The phonological characteristics of Mainstream RP, which is the object of the present study, are outlined in 2.5.

2.3 The social meaning of RP

It is impossible for an Englishman to open his mouth without making some other Englishman hate or despise him.

George Bernard Shaw (*Pygmalion*, 1916)

Accent is commonly linked with conceptions of social identity. The way a person speaks reflects aspects of his or her social background and is therefore an important indicator for the listener. Sociolinguistically speaking, RP is an unusual accent, in the sense that it is supra-regional and has enjoyed a unique social prestige. According to Wotschke (1996: 216) the most serious social gap has traditionally been that between RP and all other accents of English. There is no doubt that the legacy of the past can still be felt to some degree even today. Although its status has been reduced in the course of the last decades, RP is still a variety many speakers admire and modify towards in formal contexts. This means that RP is associated with certain social values.

There have been very few recent studies into the social aspects of modern RP. Most of what we know about attitudes towards RP stems from the research of the social psychologist Howard Giles and his associates, who carried out a number of studies on the social evaluation of accent in the 1970s and early 1980s, using primarily the so-called *matched-guise technique*. Such studies typically involve subjects listening to a number of speech samples and rating accents and their speakers on a number of semantic-differential scales. The results of these studies showed that RP was more prestigious than local accents and that RP speakers were deemed more competent than non-RP speakers. But while RP was associated with status and competence, the accent scored low for social attractiveness. RP speakers were generally perceived as educated, competent, ambitious and confident – all traits that are socio-economically related. Along the so-called Solidarity scale, however, which includes traits like trustworthiness, friendliness, generosity, sincerity and kindness, RP speakers scored low. Some listeners even rated the RP guises as arrogant and snobbish (for a detailed outline see Giles et al. 1990). Garrett et al. (2003) report similar findings from attitudinal studies conducted among a group of school teachers and pupils in Wales.

Although empirical research is scarce, there is evidence to suggest that the traditional prestige patterns are still largely valid and that RP even today ranks highly as the accent of education and authority, some twenty years after Giles' work. This is indicated for instance by

its extensive use among newsreaders and its dominance as a model for dictionaries and textbooks. There is nevertheless no doubt that the sociolinguistic situation of RP has changed considerably in recent years, largely as a result of changes in the social climate (cf. 2.1). It is a matter of general observation that the status of the accent has been dramatically reduced and that RP is no longer synonymous with success and prosperity. On the contrary, RP can now be a disadvantage in some social situations as its speakers may be perceived as posh or snobbish. An extreme example of this was seen in 2001, when the then new Speaker of Parliament, Michael Martin, sacked his secretary reportedly for being too posh and having a “la-di-dah” accent. This incident led Mary Killen to comment in a feature article in *The Guardian* that “the upper classes have come to expect this sort of thing in recent times. They are the last minority we are allowed to persecute with impunity” (Killen 2001).

There is plenty of anecdotal evidence of the falling status of RP and the frustration of RP speakers (cf. e.g. Ascherson 1994, Morrish 1999). Author India Knight (2001) complains about the “inverted snobbery” that RP speakers have to deal with:

If your accent is “posh” you are immediately viewed with hostile suspicion, the implication being you are probably some ghastly plummy nob, your very existence confirming the fact that there are still people who sneer down their long, well-bred noses at the plebs.

The result of these negative attitudes is often linguistic “downgrading” on the part of RP speakers, as a way to avoid the “establishment” connotations and sound more “down to earth” (cf. 3.1.3).

It seems, then, that the dislike of RP, and its association with arrogance and snobbishness, attested in previous attitudinal studies, is even more strongly felt today, and more outspoken. This negative aspect of RP is also reflected in popular culture: in many Hollywood films, the most evil and sinister, though often brilliant, villains are frequently played by British actors with RP accents.³

It should be added, though, that much of the dislike and criticism of the “accent of authority” is directed towards the conservative and socially conspicuous Upper-crust RP, often combined with idiosyncratic mannerisms and supra-segmental features (cf. Lewis 1985: 253). Modern, Mainstream RP is much more socially neutral and normally not subject to the

³ One of the more curious examples of “evil RP” can be seen in the sci-fi comedy *Galaxy Quest* (1999), where the main enemy, a malicious green bug-like alien from outer space, speaks perfect RP!

same dislike. Similarly, many of the comments about the crumbling prestige and impending death of RP primarily refer to U-RP, which is the conservative form spoken by older people and therefore destined to die out as a result of natural change.

Fabricius' recent study of t-glottalling in RP (Fabricius 2000) shows that the accent is widely used among young Cambridge undergraduates with public school background, a group which traditionally would be expected to use RP. My own study of broadcast speech shows that RP and near-RP dominate among news presenters. As a linguistic object, then, RP continues to exist, but its days of social and cultural hegemony are definitely over.

2.4 Defining RP

Everyone in Britain has a mental image of RP, even though they may not refer to it by that name and even though the image may not be very accurate.

Wells (1982: 279)

In an empirical study of speech variation, the first stage involves a definition of the object of study, and the selection of speakers to provide the phonological data. The decisions made at this stage are crucial to the validity of the results. As the present study aims to make valid statements about RP, it has to be made clear what is meant by the term *RP* and how the concept should be delimited. The need for an operational definition is fundamental when collecting data, to ensure consistency and reliability and avoid ambiguities. The present section gives an outline of the different ways in which RP has been defined by various linguists and discusses the problems involved with these definitions. It concludes with a presentation of the concept of non-localisability, which has served as the defining criterion in the present study of RP.

2.4.1 Introduction

RP is by far the most thoroughly described accent of English, and the model for many dictionaries and textbooks on phonetics. In spite of the large number of descriptions of RP, there exists no universal definition of the accent. Honey (1985: 241) talks of the “extreme divergence of the definitions of RP”, and according to Lewis (1985: 247) “no two British phoneticians are likely to agree on where the line between RP and non-RP is to be drawn”.

There are numerous descriptions of RP that list the phonological and phonetic features of the accent, but very few give the criteria for including a feature as part of RP. A number of sources discuss new trends and ongoing changes in RP (e.g. Wells 1994a, Wells 1997a, Taylor 1998, Upton 2004), but without explicitly stating which definition of RP forms the basis for the observations. The question of how RP should be defined is largely avoided or treated superficially, or the definitions are purely descriptive.

Defining a spoken variety is notoriously difficult. Every linguistic variety is to a certain extent an abstraction, a construct made by scholars for the purpose of description or study. The phonetic reality is never as systematic and clear-cut as a linguistic description necessarily must be. In real life there is a continuum between accents, and any partition of this continuum involves some degree of artificiality. Modern RP is particularly problematic to define, as the use of RP is not confined to one specific region or one identifiable social group. Moreover, the accent (like all living varieties) is constantly changing, and incorporates a considerable amount of variability. The notion 'RP' is not scientifically precise, and linguists disagree as to which features belong within the accent. There is general consensus that the phonological core of RP is identical with the segmental system found in the traditional descriptions (see 2.5). The problems arise when observers encounter variation and change. The codification of RP has led to the danger of perceiving the accent as static and invariable, and complicates the process of updating the descriptions. This fuzziness should, however, not prevent us from trying to circumscribe RP. As an object of study, RP has to be delimited. It is not a matter of finding the final, "true" definition, but the linguistic researcher has to operate with at least a valid working definition of the accent. The choice of definition will affect the decisions regarding which features fall within or outside RP. It is therefore important that the defining criteria are such that they can be used consistently.

Generally speaking, a linguistic variety can be defined either on the basis of attributes of the speakers (geographical origin, ethnicity, socio-economic class, etc.), or on the basis of linguistic features (phoneme inventory, lexical incidence, etc.), or with reference to evaluative criteria (the perceived status or prestige of the variety). The main concern is to ensure that the definition is operationable, and to avoid circularity.

Linguists writing about RP have operated with several different criteria for defining RP, the most important of which will be discussed below.

2.4.2 The BBC accent

Because of the close association of RP with broadcast speech, and with newsreaders in particular (cf. 2.6), *BBC English* has often been used as a synonym for the accent, and several linguists define RP by reference to “the form generally used by newsreaders of the BBC” (Gimson 1970: 88). In the latest edition of the *English pronouncing dictionary* the editors want to abandon the “archaic name *Received Pronunciation*” in favour of the term *BBC English*. They describe the British English pronunciation model (which is identical with the model most other linguists refer to as *RP*) as “the pronunciation of professional speakers employed by the BBC as newsreaders and announcers” (Jones 2003: v). RP is still the accent typically used by BBC newsreaders, and broadcast news is well suited as a source for observing RP in use, but it is problematic to define RP as the pronunciation of BBC newscasters. British radio and television companies, including the BBC, are much more permissive in their choice of newsreaders and announcers today than they were just a few decades ago. Several of the BBC newsreaders now have accents that clearly fall outside RP, including markedly regional varieties. RP can therefore no longer (if it ever could) be equated with newscast English, and other criteria are needed in order to narrow the scope of the accent.

2.4.3 An abstract norm

An important aspect of RP is its role as a model of pronunciation in the teaching of English as a foreign language, and as a reference point in accent research. One possible definition of RP views it as “no more than the codified version of English pronunciation” (Wells 1994a: 205), and reduces the accent to a theoretical construction – a standardised norm which is described in pronunciation dictionaries and textbooks, but which nobody really speaks. The majority of phoneticians, however, agree that we can identify a living speech variety that corresponds to the textbook descriptions of RP, and that some people have this variety as their native accent, others as a norm towards which they modify their speech. The problem is finding a definition which incorporates both these aspects.

The ambiguous use of the term RP is recognised by Fabricius (2000), who distinguishes between what she calls *constructed RP* (c-RP) and *native RP* (n-RP). The former refers to RP as a codified norm, the model described in pronunciation dictionaries,

while the latter refers to the native accent of a small group of people who have grown up within Great Britain. C-RP “has specific applications in areas where a standardised, non-variable pronunciation is required, most likely in formal situations such as certain broadcasting genres, while [n-RP] exhibits all of the variation we expect of naturally-occurring speech. The two are closely linked, but separate” (2000: 29-30). It is important to be aware of this ambiguity and to keep the two levels conceptually distinct (see further 3.2.2). The distinction does, however, not bring us any closer to solving the problem of delimiting the accent. N-RP is the basis for c-RP, since the codified version is (at least in principle) based on observations of “real” speech. N-RP is therefore the primary category, and the one that needs to be delimited for the purpose of empirical investigation. N-RP speakers cannot simply be identified as those who speak c-RP, or the whole concept becomes circular. Fabricius defines n-RP on the basis of the social background of the speakers, but such a definition is by no means straightforward (see 2.4.5 below).

An additional aspect here is the potential conflict between people’s mental image of RP and the actual linguistic usage. The special status of RP has led to a high level of consciousness about the accent in people’s minds, or at least about a “standard” or “high-status” speech variety (cf. the introductory quote from Wells). Such folk-linguistic notions can be very powerful, though they are sometimes far removed from the linguistic and social reality. The conceptions are based on cultural stereotypes and tend to adhere to the old monolithic image of RP, and they may affect the willingness to regard non-traditional forms as part of the RP accent.⁴ It is nevertheless important to keep in mind that these subjective notions are also an aspect of RP, although they may have no place in a scholarly definition of the accent.

2.4.4 Prestige

RP has traditionally been associated with prestige and status. Wells (1982: 115) describes RP as “the accent which enjoys the highest overt prestige in England”. The term *Received Pronunciation* itself reflects a social judgement as to what is “correct” or “acceptable”. The evolution of RP as a prestige standard was accompanied by negative attitudes towards local

⁴ Cf. the popular claim that “no one speaks RP anymore” (cf. Fabricius 2002b: 358). Even among linguists there are quite different conceptions about what RP is. When I have told British colleagues about the present project, a common reaction has been that (with the possible exception of Trevor McDonald) “newsreaders don’t speak RP anymore”.

accents and dialects, which by many were considered ugly or vulgar (cf. Honey 1988: 224). The prestige of RP is of course an extension of the social prestige of those who use the accent. As a result of its high status, RP has been codified and used as a model for comparison and imitation, and the codification is partly responsible for creating and maintaining a popular notion of a “correct” form of spoken language.

To the extent that RP is equated with the “correct” or “best” variety, it can be defined evaluatively. The social evaluation of RP is discussed in Giles et al. (1990), who state that RP can “be envisaged as an accretion of evaluations, an ‘ideal’ variety of English pronunciation, inherently conservative but predictably varying over time and space” (191). Similarly, Wells (1997a) includes the notion of an “ideal” accent as one possible criterion for defining RP: “We ask, what pronunciation is correct? What is beautiful, what is admired and imitated?” (1997a: 20).

If prestige is to be used as a defining criterion, it has to be able to be identified and measured in some way. Factors such as correctness, beauty and admiration are notoriously difficult to measure and operationalise for defining purposes. With regard to imitation, however, it is an empirical fact that a number of native speakers modify their accents towards RP in formal situations (cf. Ramsaran 1990b), which can be seen as an indication of overt prestige. However, the most common way of measuring the prestige of linguistic varieties is through direct or indirect elicitation of people’s evaluative judgements of speech, using the matched guise technique or similar methods (cf. Garrett et al. 2003). A relevant question is then how a prestige variety should be characterised. It is well known that RP is regarded very highly by some, but has negative connotations for others. The accent has typically been found to score highly for traits such as ambition, intelligence, education and wealth, but to be downgraded in terms of social attractiveness and solidarity (cf. 2.3). The notion of ‘prestige’ may therefore have to be defined in very specific terms as referring to a particular set of characteristics. But even then, RP may not necessarily or consistently be rated as the highest prestige variety. Newbrook (1999), for example, found that informants in West Wirral generally, but not always, rated RP forms as more prestigious than local forms.⁵ Coupland (1990) notes that the middle-class in Wales can be subdivided into two quite different groups based on their ideological orientation, and these groups have been found to differ markedly from each other in their reactions to RP (Bourhis and Giles 1976, referred in Garrett et al.

⁵ In those cases where overt prestige was attached to local variants rather than RP variants, this was sometimes done as a conscious rejection of RP as a norm and “in full knowledge of which forms are found in RP” (Newbrook 1999: 101). Often, however, RP was accepted as the source of norms, but there was uncertainty or ignorance as to the appropriate forms.

2003). Moreover, the rise of various ‘regional standards’ complicates the picture even further. Garrett et al. report that a regionally based candidate for ‘standard Welsh English’, while not reaching the same level of prestige as RP, is “relatively prestigious and has a superior overall evaluative profile to that usually found for RP” (2003: 143). Similarly, Milroy (1987) notes that researchers in certain parts of Britain have found it difficult to identify the “highest” prestige forms, or to operate with a single prestige norm, in speech communities where there are local high-prestige accents and sharp discontinuity with any supra-local norm (1987: 118-119).

A different kind of problem relates to the *operationalisation* of the concept of prestige for defining purposes. In attitudinal studies such as the ones referred to above, the linguistic varieties to be evaluated are predefined and selected in advance. If prestige is to be the defining criterion for RP, and prestige is determined in evaluative studies that require preselected speech samples, there is an obvious danger of circularity. The researcher will then need to resort to some other criteria when selecting (potential) RP speakers to be evaluated for perceived prestige.

Another consequence of using prestige as a criterion for identifying something as RP, is the risk of ending up with too broad categories. The linguistic researcher may need to distinguish between mainstream RP and conservative RP, or between RP and near-RP (cf. 2.2). Such a distinction is however normally not made by non-linguists, and these related varieties would probably receive the same evaluative judgements in terms of prestige. Other criteria would therefore be necessary in order to make such fine-grained differentiations.

2.4.5 Social definitions

One approach to defining RP could be to relate the accent to a social group. The earliest accounts of a *received pronunciation* described the accent of the educated members of the upper classes (cf. 2.1), and the term *received* originally referred to social acceptance. RP as a standard and as a marker of educated status was the result of certain social and political conditions in Britain in the period 1870-1950. The earliest definitions of RP therefore focus on the social aspect of the accent, and identify the speakers with reference to their class and education.

In present-day Britain, the conditions that gave rise to RP as a special social phenomenon, are no longer present. Up until the middle of the 20th century, RP remained the

accent of one particular social group. Since then, however, there have been radical changes in the structure of British society and in the attitudes towards accents. The class divisions are less strict, people are more mobile, and an increasing percentage of the population has access to higher education. Several of the criteria previously used to delimit RP are irrelevant today, as there is no longer a straightforward relationship between social class and education or profession. We now find RP speakers with a much wider spectrum of social backgrounds, and the upper classes no longer have a uniform style of pronunciation. RP is still linked in the public mind with the concept of 'educatedness', and the speakers are situated above the lowest end of the social scale, but RP is no longer the exclusive property of one identifiable class, and RP speakers therefore cannot be identified solely on the basis of their social background. This fact is recognised by many scholars (cf. e.g. Gimson 1984, Honey 1985). Ramsaran (1990b: 178) states that "since it is ... quite unrealistic to try to label the accent as belonging to a particular section of society, it is impossible actually to identify the accent ... in social terms". Gimson (1977: 157) underlines the need to "dissociate the definition of a standard from the speech of an easily identifiable and separate ruling class" and points out that "the upper classes no longer have a single, typical style of speech".

Honey (2000) claims that social class was never the most important defining characteristic of RP, and "superior social rank was not in itself a guarantee of the right accent". The correlation was rather with education: RP was not an exact synonym for the privileged and elite, "because it only covered those members of the elite who had had an appropriate education, and it also embraced those people not from an elite background who had managed to acquire this accent" (Honey 2000). Education is however not a reliable criterion for defining modern RP, as "it is no longer the case that all or even most educated people in England speak RP" (Wells 1997a: 20). This observation is also made by Gimson (1980: 91), who notes that "features of *regional* pronunciation, without any contamination from RP, will be found in highly educated and less educated speech".

In spite of the lack of a precise correspondence between social background and the use of RP, several phoneticians prefer a sociolinguistic definition of RP. John Wells, who is one of the leading authorities on RP today, defines the accent in social terms in e.g. Wells (1994a) and Wells (1999b). In the latter he describes RP as "the pronunciation of people at the upper end of the social scale – whatever that is at any given time". Such a definition is problematic, however, considering the fact that people from the upper classes no longer have a uniform

accent.⁶ The term RP would then cover an extremely wide range of pronunciations. The statement is modified in Wells (2000: xiii), where he writes that “the democratization undergone by English society during the second half of the twentieth century means that it is nowadays necessary to define RP in a rather broader way than was once customary”. The question also arises of how far down the social scale one should go when delimiting RP. As Wells (1997a: 19-20) points out, “[t]he proportion of the population regarded as upper-class is extremely small, and we clearly need to consider the upper-middle classes as well”.⁷ The problem of diversity of speech variants, however, remains.

Milroy (2001a) insists on including social factors as part of the definition of RP, and rejects a linguistically based definition. He points out that RP no longer has the social position and status it once had (and which was the basis for the term *received*), and concludes that, therefore, RP no longer exists. At the same time, however, he admits that the accent can still be heard in Britain. This means that the phonetic forms of RP do exist, but, Milroy maintains, “there is more than this involved in the definition of what RP is” (2001a: 15). I would draw the opposite conclusion and argue that RP exists, but can no longer be defined in social terms.

Fabricius (2000) employs a social definition of RP in her analysis of the accent: she uses two features of social background, namely social class and education, to define a social group as the basis for RP. Although social criteria are used to select the potential informants (young Cambridge students with public school background), the author nevertheless has to resort to phonological criteria in the final selection of speakers: “it was not possible to ignore linguistic criteria entirely. I established a phonemic definition of mainstream RP which I used as a check on the speakers chosen according to social and educational background” (2000: 78). From this it would seem, in my view, that phonological criteria are paramount and that social criteria strictly speaking are redundant.

This is not to say that social aspects are not relevant to RP. The accent undoubtedly has social connotations: it is closely linked with education and social status, and while it may not be exclusive to one particular class, it is typical of the upper and upper-middle classes. The point is that the social background of the speaker is in principle irrelevant when judging whether or not a speech sample represents RP.

⁶ Particularly in areas where a local or regional “standard” exists alongside RP (cf. Milroy 1987: 118-119).

⁷ The question of class and accent is further complicated by the fact that there is no straightforward way of defining social class. The criteria used to divide informants into social groups often refer to differences in occupational status, following the sociological tradition of dividing occupations into social classes. It is however not self-evident how to define e.g. ‘manual worker’, or where to draw the boundaries between ‘skilled’ and ‘partly-skilled’ occupations, etc. (cf. Hudson 1996: 151).

2.4.6 Phonological definitions

One approach to defining RP without reference to social class or education, is to refer solely to phonological criteria, and describe the phoneme system and its phonetic realisations and lexical incidence. Many writers stress that phonetic specification of RP is central to its definition. Gimson (1984: 46) points out that there is a phonological tradition of a standard, “a single phonological system which has been evolving in time” and that this “is the most reliable basis for our definition of present-day RP”. The phonological tradition can be traced through major works like Jones (1960), Wells (1982, 2000) and Cruttenden (2001), which all give detailed presentations of the phonological features of RP. The difficulties arise, as Ramsaran (1990b) points out, when phoneticians disagree as to which features belong within RP and where to draw the phonological boundaries around the accent. Any phonetician can set up a group of features which he identifies as ‘RP’ and find a speaker who fits his expectations. Different phoneticians who choose slightly different required features can then end up describing a number of different idiolects which they all refer to as RP (Ramsaran 1990b: 180). These problems are further complicated by the fact that RP, like all accents, is subject to change and variation. If, for example, a traditionally non-RP feature, such as the glottal replacement of /t/, is present in the speech of someone who is otherwise identified as an RP speaker, some will conclude that t-glottalling has now entered RP, whereas others will claim that the presence of a glottal stop automatically makes this a non-RP accent.⁸

All accents change over time, and at some point we have to include certain non-traditional features in the descriptions of RP, in order to find out how the accent is changing and ensure that RP does not remain stuck in the Jonesian version forever. A purely phonological description will soon become obsolete if it is not constantly updated. This brings us back to the problem of finding a group of speakers to represent the current usage. The selection of phonological features has to be based on a criterion that is objective, and flexible enough to allow for variation and change.

⁸ Other controversial features include vocalisation of non-prevocalic /l/ ([fɪʊm] *film*), the use of affricates /tʃ, dʒ/ in e.g. *Tuesday*, *reduce*, and fronting of the vowel of *goose* (from [u:] to [ʊ: ~ ʏ:]) (see further 4.2).

2.4.7 Non-localisability

RP has its origins in the southeast of England, as the pronunciation of educated speakers in and around the capital (cf. 2.1). Although RP has its historical roots in a specific region, and shares its main phonological basis with southeastern English, it is now non-regional, or supra-regional, in that it extends over the whole of Britain (or at least England), although only a small minority of people speak it. This spread of RP came about as a result of the nationwide diffusion of the educational system in the 19th century (cf. 2.1).

Non-localisability is an important characteristic of RP which most writers view as relevant when defining the accent. RP was early characterised as non-regional. Wyld (1920: 2) describes it as “not confined to any locality, nor associated in any one’s mind with any special geographical area”. Trudgill (2002: 172) takes it to be “a defining characteristic of the RP accent that, while it is clearly a variety that is associated with England, and to a certain extent also with the rest of the United Kingdom, it otherwise contains no regional features whatsoever”. Ramsaran argues that “if it is possible to identify ... a non-localisable accent and if the social definition of RP is outdated, then why not simply describe this ‘non-regional’ accent as the current version of RP?” (1990b: 179).

Non-localisability seems to be the decisive factor also in Fabricius (2000). The informants in her study were selected on the basis of their social background, (see 2.4.5 above), but the final selection was made with reference to localisability: speakers with “recognisable regional phonemic features” were deemed “not suitable for the ... study” (2000: 86). Three phonetic experts were asked to assess the localisability of the final speech samples, and “there was general agreement between the judges that it was difficult to place individual speakers in a specific regional category [and] the group was deemed overall to be representative of non-localisable speakers” (78-79). One of the phoneticians commented: “they are RP speakers because it is not possible to determine their geographical origin with any degree of delicacy” (79).

Wells (2000: xiii) also describes RP as non-localised, which he defines as “not associated with any particular city or region”. Furthermore, he mentions pronunciations which are “widespread in England among educated speakers, but ... nevertheless judged to fall outside RP”. These pronunciations are all localised forms.

There are, however, phoneticians who have a different view of non-localisability and who regard RP as an overall southeastern accent. Gimson (1984: 47) comments that “what has remained constant is RP’s regional base: its characteristic phonological features have always

been those of the south-eastern region of England”. Nolan (1999) rejects the notion of RP as non-localisable, which he characterises as the “common view which refuses to locate RP geographically, and instead views it as a non-regional prestige variety” (86). Nolan holds that RP forms a phonological continuum with local accents in the southeast and that there is no such link between RP and northern accents. He argues that there are parallels between the historical development of RP and the southeastern accents, and that they are undergoing similar changes. These observations are on the whole correct. Many of the features of RP are also found in accents of the southeast. When RP can be said to be non-regional, it is because the use of RP does not automatically locate the speaker in the southeast, and native RP-speakers are found in all regions of Britain. RP features that were originally southern are today neutral, in the sense of being geographically widespread. For example, the use of a rounded vowel in the word *cup*, immediately places the speaker in the north of England, while a speaker who uses the RP variant, an open central vowel, is not necessarily from the south. RP is thus a regionally unmarked variant, and in that sense non-localisable. The fact that the origin of a feature can be traced to a specific region is irrelevant. The historical origin of RP has always been placed in the southeast of England, although RP also contains features that had their origin in other regions (see e.g. Trudgill 2002: 175, Ramsaran 1990b, Wotschke 1996).

Nolan defines RP as “the long-established term for the prestige accent of South East England which also serves as a prestige norm in varying degrees elsewhere in Britain” (Nolan and Kerswill 1990: 316). According to Nolan it is prestige, not non-localisability which is the defining characteristic of RP. Many writers include prestige as a factor in their description of RP (cf. 2.4.4). Prestige and non-localisability are in many ways intrinsically linked in the British speech community. According to Lewis (1985: 244): “a fair degree of social prestige is associated with all speech which is completely non-regional”. The same comment was made by Sweet a century ago: “The best speakers of Standard English are those whose pronunciation ... least betray[s] their locality” (Sweet 1908: 7). In Britain, it has traditionally been a sign of social status to speak “without an accent”, that is without any regional features.⁹ This is still, to a large extent, the case, and probably part of the reason why RP has retained much of its prestige.

⁹ The status of non-regional speech apparently originates from the style of the old landed aristocrats, “who had a country house and a London house and whose personal connections were countrywide rather than local” (Ascherson 1994).

The rise of RP as a non-regional prestige variety can be seen as the continuation of a long tradition, namely the spread of London features to other parts of the country (see 2.1). In present-day Britain, with the breaking down of old class barriers this trend seems to be moving faster than ever before. Because of the traditional attraction of the capital and “the recent trend for people of working-class or lower-middle class origins to set the fashion in many areas” (Wells 1982: 118), it is primarily features of lower-class London speech which are spreading socially and geographically, and also making their way into RP. Reports of changes in RP attribute many of the “new” features to the diffusion of London speech (e.g. Wells 1997a, Taylor 1998). Wells (1994a) talks of the “Cockneyfication” of RP, and lists a number of features accepted into modern RP. Other writers predict the death of RP and the rise of a new standard form based on the accents of London and the surrounding areas – so-called *Estuary English* (see 3.1.3 and 4.2.3)

There are in other words two opposing ways of analysing the on-going linguistic trends: on the one hand, RP is viewed as a distinct variety which is in a constant process of change and modification partly through influence from other accents. The second view focuses on RP as a social phenomenon, and sees the decline of its exclusive social position as proof that RP is being phased out and replaced by a broader popularly based accent (a view supported by e.g. Milroy 2001b and Rosewarne 1994). RP as traditionally described is of course in decline, but to say that the accent is disappearing and being replaced is, I would maintain, to view RP as static, invariant and unchangeable.

What is then to be made of the recent reported changes “from below”? The idea that pronunciations which originate from the London area are making their way into RP is not problematic if RP is defined as a non-localisable accent: when features which were once regional become so widespread that they lose their exclusive local identity, they may eventually become part of RP (this is after all what happened to many of the traditional RP features). It is then not inconceivable that the future RP will be very similar to today’s London accent.

If we delimit RP as a non-localisable accent, the most important criterion for the inclusion of a new feature in RP must therefore be whether the feature is localisable or not. This implies, according to Trudgill (2002: 175), “that there will be features that for a period of time, while a change is taking place, may have an indeterminate status”. One way of verifying that a feature is non-regional is if potential RP speakers from different geographical regions show the same results with respect to a specific pronunciation feature. The feature in question must then be considered a widespread and non-localisable characteristic and thus acceptable

as part of RP. This provides a means by which the descriptions of RP can be kept updated and in accordance with current usage.

A definition based on non-localisability is also in line with the RP tradition. RP has always acquired widespread forms that originally were part of local accents. After the Second World War, with increasing social relaxation and a rising pride in local accents, the linguistic range of RP widened even more. Gimson (1980: 302) writes of a “wider-based RP” with the “admission into the permitted speech forms of certain variants until recently regarded as regional”. Many of the regionally-based post-War trends are now widely accepted as part of modern RP (cf. Wotschke 1996: 222-223). Non-localisability is however not an absolute value, and there will be disagreement as to how widespread a linguistic feature must be, both in terms of geography and usage levels, in order to be regarded as non-regional. The phonological boundaries of RP will therefore always be somewhat fuzzy in order to allow for the entrance of new features.

Defining RP as non-regional is not particularly controversial, and most writers refer to this element in their descriptions of the accent. But none of them use non-localisability as the only defining criterion. Even Trudgill, who includes it as a defining characteristic of RP (2002: 172), adds that non-regionality is “a necessary but not sufficient condition for a feature to be considered RP” (177). He gives as example the spreading of h-dropping, and states that “if all regions of England were to acquire /h/-dropping ... that would not make it an RP feature” (177). Wells (1994a: 199) also mentions h-dropping and g-dropping as widespread features characteristic of popular accents in several regions of Britain, but resisted by RP. These observations may be seen as an argument against a strict definition of RP as non-localisable. Trudgill does however not specify which other condition(s) would have to be fulfilled for a feature to be accepted as part of RP. Presumably the feature will have to be used by speakers of the upper or middle social strata, or at least not have working-class connotations, which, at present, h-dropping does. However, several of the features that have entered RP in recent decades have a similar history. T-glottalling, for example, has previously been strongly stigmatised, but has spread rapidly throughout the country and gradually lost its social stigma. The change in social status has occurred after, and possibly as a result of, the extensive geographical spread. It seems then that social diffusion follows geographical diffusion, and it is therefore not inconceivable that also h-dropping will become part of RP in the future if it continues to spread throughout Britain. At the moment, though, it seems to be found only with speakers who otherwise have regionally marked accents, and thus already are

excluded from RP. A non-localisable RP accent does then not incorporate all non-regional pronunciation features, but it excludes all regional features.

2.4.8 Summary

Recent social developments in Britain have made it increasingly difficult to define RP, and the term has become ever more imprecise. The narrow description of RP as the property of a single social class is no longer valid, and social background is therefore not a reliable measure for identifying RP speakers. At the same time, current phonological changes have led to a need to update the descriptions of RP. The question then arises of which criterion will best serve as a basis for defining the accent.

Which definition of RP is most appropriate, largely depends on the purpose the definition is to serve. All the aspects discussed above are of course relevant to RP and will have to be included in a descriptive definition which aims at providing a complete characterisation of the accent. An operational definition, which serves the purpose of delimiting RP as an object of study, needs to identify a valid criterion which can be used as objectively and consistently as possible. For the present study, where speaker selection is based solely on the informants' speech, the definition has to involve reference to phonological features.

The only aspect which has remained stable in RP, at least since the late 19th century, seems to be its lack of regional affiliation, and this can be used as a way of delimiting the accent. Non-localisability provides a criterion which is flexible in that it allows for changes in RP, and explains many of the ongoing tendencies in the accent. It secures a continuation of the RP tradition, and it can be tested objectively. While a definition based only on linguistic criteria is certainly not unproblematic, it has been deemed a valid working definition for the purposes of the present study (see further 5.1.1).

2.5 The phonological core of RP

The cultivated speaker employs a definite number of sounds which he utters with precision, distinctness, and in their proper places.

B. H. Smart (*A Practical Grammar of English Pronunciation*, 1809)

RP is typically identified with reference to its phonological characteristics, as an accent that includes certain pronunciation features and excludes others. There is a significant amount of variability within RP (cf. 4.2), and the phonetic borders around the accent are slightly fuzzy (cf. 2.4). There is nevertheless a core of phonological features that constitutes the essence of the RP accent, and which distinguishes RP from all other accents of English. This set of core features has constituted the main reference point in the selection of speakers for this study.

The basic phoneme system of RP has remained essentially the same since the accent was described by Daniel Jones almost a century ago. There is today broad consensus among linguists with regard to the principal phonological features of the RP accent. The disagreements mainly concern certain phonetic realisations, most of which involve consonant features (cf. 4.2).

The basic phoneme inventory of RP is outlined below, along with the most important phonetic and incidental specifications. Complete and detailed descriptions are available in e.g. Wells (1982, 2000) and Cruttenden (2001).

The phoneme system of RP includes 20¹⁰ contrastive vowels and 24 consonants. The vowel phonemes are the following¹¹:

/ɪ, e, æ, ɒ, ʌ, ʊ, ə, i:, u:, ɔ:, ɑ:, ɜ:, eɪ, aɪ, ɔɪ, əʊ, aʊ, iə, eə, ʊə¹²/.

The consonant inventory is the same as in the majority of other accents of English:

/p, b, t, d, k, g, tʃ, dʒ, f, v, θ, ð, s, z, ʃ, ʒ, h, m, n, ŋ, l, r, j, w/

Among the most important diagnostic features are the phonemic oppositions /ʌ/ - /ʊ/ in e.g. *putt* - *put*, and /ɒ/ - /ɔ:/ - /ɑ:/ in e.g. *cot* - *caught* - *cart*. With regard to lexical

¹⁰ Some writers acknowledge the weak vowels /ɪ/ and /ʊ/ as separate phonemes (see e.g. Wells 2000). These are the vowels found in unstressed syllables word-finally and prevocally, e.g. *city* ['sɪtɪ], *influence* ['ɪnfluəns]. Because of their limited distribution many will hesitate to assign phonemic status to these vowels, but rather analyse [i, u] as contextual allophones of /i:/, u:/ or /ɪ, ʊ/.

¹¹ The phoneme symbols used are the same as in Wells (2000) and Cruttenden (2001). See Upton et al. (2001) for alternative symbols.

¹² The diphthong /ʊə/ has a receding phonemic status, and is potentially replaced completely by /ɔ:/ with some speakers (see further 4.4.1).

incidence, /ɑː/ is used in words of the lexical sets¹³ BATH, PALM and START; the lexical set STRUT has /ʌ/, and LOT has /ɒ/. On the other hand, the use of /ʊə/ or /ɔː/ in CURE is optional, as is /ɪ/ vs. /iː/ in *happy*. With reference to phonetic specification, we can state that RP /iː/ and /uː/ are realised as relatively pure monophthongs, without diphthongisation, though the quality of /uː/ can vary from back [uː] to centralised or even fronted [ɯː ~ ʏː]. The diphthongs /eɪ/ and /ɔɪ/ have a close-mid and an open-mid starting point, respectively (never [æɪ], [oɪ]). /əʊ/ is normally realised with a central starting point, but is optionally pronounced with a back rounded starting point before non-prevocalic /l/. The most important phonotactic restriction in RP regards the distribution of /r/. RP is a non-rhotic accent, which means that /r/ only occurs prevocally. The accent further maintains the allophonic opposition between clear and dark /l/ in prevocalic versus non-prevocalic position. The phonemes /θ/ and /ð/ are realised as dental fricatives and never fronted to /f, v/ or realised as plosives.

2.6 RP and broadcasting

One hears the most appalling travesties of vowel
pronunciation. This is a matter on which broadcasting may be
of immense assistance.

John Reith (*Broadcast Over Britain*, 1924)

The link between RP and broadcasting goes back to the beginning of British radio broadcasting in the 1920s. The history of British broadcasting is to a large extent the history of the BBC. For many years the corporation held a monopoly on broadcasting in Britain, and its decisions and practices with regard to pronunciation were highly influential and contributed significantly to the strengthening and dissemination of RP as a standard.

John Reith was the first managing director of the BBC. He had “a Messianic vision of the importance of broadcasting” (Pointon 1988: 8), and aimed for the highest standards of excellence at all levels, including pronunciation. In 1924 Reith wrote:

We have made a special effort to secure in our various stations men who, in the presentation of programme items, the reading of news bulletins and so on, can be

¹³ For more about lexical sets see section 4.0.

relied upon to employ the correct pronunciation of the English tongue. (Reith 1924: 161).

In 1926 the Advisory Committee on Spoken English was formed, whose function it was to advise announcers on the pronunciation of contentious words. Among the members of the committee were Daniel Jones and Arthur Lloyd James, both linguists, who ensured that the extreme prescriptivism expressed by Reith was tempered. During the 1930s the BBC published a series of booklets, edited by Lloyd James, which included some recommended pronunciations for announcers. The first edition contained a few hundred words “of doubtful pronunciation” where two or more pronunciations were possible (e.g. *garage*, *idyll*, *iodine*, etc.). The Advisory Committee then turned its attention to proper names, and the following editions dealt only with place names and family names. After the Second World War the Advisory Committee was replaced by the Pronunciation Unit, which continues today as the Pronunciation Research Unit, and is part of BBC Information and Archives. The Unit today has a staff of three full-time linguists who research and advise on the pronunciation of names and foreign words to the entire BBC. They maintain a database of pronunciations available to all employees at the BBC. The aim of their advisory service is to ensure that the broadcasters are accurate and consistent in their pronunciation of, primarily, foreign names. The Unit does not advise on English common words or promote any particular accent.¹⁴

It has never been part of the BBC’s official stated policy that news presenters should speak RP (or any other accent). Nevertheless, there is no doubt that RP has been the dominating accent in news broadcasting and to a large extent still is. In the past RP was at least indirectly promoted. Lloyd James wrote in 1931:

There are varieties that are acceptable throughout the country, and others that are not. ... The BBC has no desire to accept or to dictate any standard of pronunciation other than the current usage of educated speakers. ... There is no standard dialect, but here, as in all communities, the educated speech of the capital starts with a heavy handicap in its favour. (Lloyd James 1931: 12-13)

At the time this was written, the accent of “educated speakers” was predominantly RP. Even though the BBC did not demand that the announcers and presenters should speak RP, they all did, at least in the era before the Second World War. According to Briggs (1985: 68) regional

¹⁴ I am grateful to Catherine Sangster and Lena Olausson at the BBC Pronunciation Research Unit for providing information about the current work of the Unit.

accents were, “in general frowned upon by the young BBC which prided itself on what it was doing to raise the standards of speech”. According to Abercrombie, however, the widespread use of RP was a result of the uniform social background of the presenters:

It has popularly been assumed that the BBC used to demand of its announcers that they speak RP, whose use the BBC promoted. All BBC announcers **did** speak RP, it is true, but in fact that was an accidental by-product of another policy: that BBC employees – administrators as well as announcers – should be of good social position, with appropriate interests and tastes. The BBC had an official whose business it was, by interview, to ensure this (he was a high-ranking ex-naval officer). The question of accent never arose; all suitable applicants naturally spoke RP. (Abercrombie 1992: 7)

A second factor which is often mentioned as a reason for the predominance of RP, is the assumption that RP was the most widely understood and tolerated accent in Britain. Whether this was, or indeed is, the case, has never been investigated or confirmed in any study. According to Quirk (1982), however, comprehensibility was probably not the main reason for using RP in newscasts: “it is much more likely that the point most favouring [RP] was its acceptability as the voice of prestige and authority” (1982: 5). The status of RP suited the BBC’s image and ‘media goals’ (cf. Leitner 1980). The latter are determined by the self-image of the medium, the expected audience, the topics, etc. The BBC, and especially the news broadcasts, aimed primarily to inform and educate, and was message-oriented and concerned with credibility – qualities traditionally associated with standard varieties (cf. Leitner 1980). Regional varieties are considered more addressee-oriented and associated with personal qualities. RP was therefore the “natural” accent for news broadcasting in the early BBC:

As a nonregional and prestige accent, which was used almost exclusively in all public and educational domains anyway, and which was considered the natural phonetic realization of texts written in standard English, it corresponded well with [the BBC’s media goals]. ... It could also be looked upon as appropriate in view of the high social responsibility that the BBC assumed. (Leitner 1980: 84-85)

There were a few attempts to introduce non-RP accents in serious broadcasting genres, but these regularly failed. During the Second World War, the BBC tried to use well-known personalities with local accents, such as the actor and Yorkshireman Wilfred Pickles, as newsreaders, but after massive complaints from listeners the experiment was abandoned (cf. McCrum et al. 1986: 30). Regional accents apparently did not convey the same authority as RP, as Mugglestone (1995: 324) comments:

Wilfred Pickles, employed to read the news during the Second World War on the premises that the Germans would find it impossible to imitate his Halifax accent, encountered a number of complaints as listeners criticized his voice. Without the accents of authority embedded in contemporary attitudes to RP, listeners claimed that they could not believe the news when conveyed in such Yorkshire tones, and that its integrity was compromised by such deviation.

It was only in the 1960s that non-RP accents entered the newscasts. This development coincided with general social changes which led to a greater tolerance of regional accents. At the same time, the broadcasting companies increasingly let the men behind the scenes participate in the news broadcasts, and these people, unlike the professional announcers, did not necessarily have an RP accent. Moreover, the newscasts generally became less formal and introduced a more colloquial style, a development which has continued up to the present time (cf. 3.1.5.4). With the growing influence of the lower classes, the BBC had to become more addressee-oriented, with focus on more personal involvement and less social distance. The BBC was, according to Leitner, “forced to accept that ‘the public’ demanded to be able to *identify with* the message that was being conveyed” (1980: 88). The linguistic consequence was the increased use of regional varieties in newsreading.

In spite of these changes, RP continued to dominate in “serious” news broadcasting. In 1971 G. M. Miller, former head of the BBC Pronunciation Unit and editor of the *BBC pronouncing dictionary of British names*, states:

Even today, when a much wider variety of voices is heard, the old style is still regarded as having an important place in broadcasting. The good announcer remains, as far as the BBC is concerned, the pleasant, unobtrusive speaker who does not distract attention from his subject matter by causing embarrassment, unwitting amusement, or resentment among intelligent listeners. ... In the BBC’s European and World Service English language broadcasts, understandably,

Southern English RP remains the accepted norm for all announcers, both in news bulletins and in programmes. (Miller 1971: v-vi)

As late as 1981 the BBC published a booklet with recommendations regarding the pronunciation of English vocabulary words. Although the editor states that “it can no longer be assumed that all broadcasters should speak R.P.” (Burchfield 1981: 6-7), the pronunciations recommended all coincide with the forms used in RP, with *Everyman’s English pronouncing dictionary* as the standard authority. In 1985, the then head of the BBC’s Pronunciation Unit, Graham Pointon, was even more explicit in appointing RP as the accent recommended for newsreaders. In the BBC2 documentary *Talking Proper* he declared:

We aim to recommend the pronunciation which will cause the least amount of offence to the greatest number of people. We use the EPD, *Everyman’s English Pronouncing Dictionary*, because we find this to be the most reliable. It’s based on the sort of English which used to be heard most frequently in public schools, and people who have received a high standard of education tend to speak with this accent. It’s known to phoneticians as Received Pronunciation, abbreviated to RP. The accent that we’re promoting is really that RP accent. I don’t think it matters that 95 per cent of the population doesn’t speak RP, because 95 per cent of the population will accept that sort of English as being the one which they want to hear from people giving authoritative statements like reading news bulletins or announcing programmes.

The *BBC English dictionary*, published in 1992, also equates the English used by its World Service presenters with RP: “The pronunciations in the BBC English Dictionary represent the standard accent of British English that is used by World Service newsreaders: ‘Received Pronunciation’ or ‘RP’” (Sinclair 1992: xxix).

The decline of RP in the broadcasting media has perhaps been most noticeable in the course of the last decade, as regional accents have entered ever more areas of television and radio broadcasting. In 1993 the then managing director of BBC Radio Liz Forgan decided to strengthen the use of regional accents on the air. She stated that the BBC was lagging a “little behind the sound of the nation, beginning to sound a bit antique”, and that the presentation style was not “energetic” and “vigorous” enough (quoted in Culf 1994). Abandoning RP was thus part of a process to modernise at least some sections of the BBC. According to

Ascherson (1994), this decision sparked protests from the audience similar to those that the BBC received during the Second World War, but, he adds, “the change had begun”.

Even though RP no longer dominates broadcasting in the way it used to, it is still the most common accent in newscasts, at least on the main channels that broadcast nationally and internationally. According to Wells (1982: 279) “RP is what everyone living in the United Kingdom hears constantly from radio and television announcers and newsreaders”. In his informal observations of BBC newsreaders, Lewis (1985) notes the predominance of RP: “Hardly any of the principal news-bulletin readers of recent years has had a very markedly non-[RP] accent. ... Virtually all the ones I have heard make considerable concessions, whether consciously or not, to [RP]; as is only natural in the situation” (252-253). As recently as 1996, Hughes and Trudgill stated that, if asked to point to a readily available example of RP, they “would probably suggest the speech of BBC newsreaders” (1996: 3). And in the latest edition of the *Cambridge English pronouncing dictionary* (2003) the editors write:

The model used for British English is what is referred to as *BBC English*; this is the pronunciation of professional speakers employed by the BBC as newsreaders and announcers on BBC1 and BBC2 television, the World Service and BBC Radio 3 and 4, as well as many commercial broadcasting organisations such as ITN.
(Jones 2003: v)

My own studies, based on observations of a selection of channels that broadcast nationally and internationally, suggest that RP is the accent that most newsreaders use or modify towards. There are presenters with regionally flavoured accents, but they all modify strongly in the direction of RP, and broad regional accents are absent. It should be specified that this observation pertains to newsreaders who present the “serious” news: political world news and current domestic affairs. Non-RP accents are much more common in other types of newscast, like sports news and business news, which in general also have a more informal style.

Two other recent studies of the speech of television and radio presenters observe that RP or near-RP is very prevalent in British newscasts. In his MA thesis, McKenzie (2004) found that RP was dominant among television news presenters from BBC1 and ITV, although the data were based on a very small number of speakers. In a study of the pronunciation of radio presenters and reporters from BBC Scotland (Wold 2003), the author found that many of the speakers “anglicised” their Scottish accent, and the modifications were always in the direction of RP.

News broadcasting has played, and to some extent still plays, an important role in exposing people to RP and strengthening its image as the accent of authority and competence. As a national broadcaster, the BBC became the main medium in which the general public encountered RP, and the accent became inextricably linked with the values and qualities of the broadcasting corporation. The exact nature and extent of the media's influence on people's linguistic behaviour and attitudes is largely unknown and difficult to measure, and the issue is virtually unresearched. In the early days of broadcasting, many people, both within and outside the BBC, were convinced that the new access to the accent of authority and cultivation would result in widespread imitation and adoption of the standard variety (cf. Quirk 1982: 6, Reith 1924: 161-162). This clearly did not happen. Sociolinguistic research has shown that linguistic features spread primarily through direct personal contact and face to face interaction (e.g. Trudgill 1986), and the media are accorded little or no influence on people's linguistic behaviour, at least as far as phonology is concerned. Among non-linguists, however, the view that the spread of standard varieties and the corresponding decline of regional varieties is largely due to the influence of the media, has been widely accepted without reservation (cf. Leitner 1983). In reaction to this simplified view, some linguists reject media influence entirely (e.g. Milroy and Milroy 1999). Others recognise at least the potential influence on language users. Quirk (1982: 18) "believe[s] in the positive influence of broadcasting services on the extended response to – and probably on the active use of – the language". In an international study of current attitudes towards varieties of English (Bayard et al. 2002), the researchers found that American English was clearly favoured over British English (near-RP), also among British teenagers. They contribute this result to the dominance of American English in broadcast media. Similarly, Williams and Kerswill (1999) suggest that exposure to southern accents through broadcast media may be a contributing factor in the spreading of these features among British teenagers in all parts of the country (cf. 3.1.3).

In his discussion of language on the radio, Leitner (1980) stresses the correlation between broadcast language and the rest of the speech community, in that the former is largely determined by the sociolinguistic structure of the community as a whole. The media's choices primarily reflect language attitudes in society in general, rather than create new attitudes (the BBC would not have used RP if the accent had not already enjoyed high prestige). But the media can strengthen and maintain these general attitudes. Bell (1983) claims that broadcast media play a multiple role – active as well as passive – in language standardisation:

First, in their choice of standards for their own use, broadcast media reflect the language evaluations of society at large. They then extend that standardization further by selecting certain forms and codifying them for use by announcers. And third, by the use of these forms and the standard language as a whole, the media disseminate the standard and further enhance its prestige. (Bell 1983: 29)

In today's broadcasting media the degree of codification is minimal, and there is no longer the same emphasis on a uniform standard variety. There can, however, be no doubt that broadcasting media have at least an indirect influence on our knowledge of and attitudes towards language. Exposure to accents from the media may not change the way people speak, but it instils in them a notion of, for instance, what "serious" language sounds like as opposed to "comical" language,¹⁵ or how people with authority speak. The media can thus participate in the assigning of social values to linguistic varieties. When RP is used in "serious" broadcasting like news or documentaries, the media contribute to enhancing the status of RP and strengthening its associations with authority and competence (see further 3.3.3). Similarly, the same media can enhance the status of non-RP accents by using them in contexts previously reserved for RP.

Media exposure certainly seems to be important as a way to promote minority languages or dialects. There are several examples of how language societies and organisations have campaigned for the increased use of minority varieties in radio and television as a means to enhance the status of these varieties and prevent their disappearance. The underlying assumption is then that the use of a variety in broadcasting media has an effect on the social role or status of that variety.¹⁶ If this is the case with non-standard or minority varieties, the same is presumably true also of standard varieties.

Lastly, there is one area in which broadcasting is of great importance and extremely influential, namely as a source of input to foreign learners of English. For people outside Britain, who are not exposed to English in their personal networks, access to authentic

¹⁵ Comedians and comical characters have traditionally spoken with regional dialects, a convention that can be traced back to Shakespeare and Dickens. Arthur Lloyd James in the BBC believed that regional accents had lost their potential authority precisely through association with buffoonery, and commented in *The Listener* in 1935: "You mustn't blame the BBC for killing dialect. The Lancashire comedian has killed the Lancashire dialect" (quoted in Quirk 1982: 6).

¹⁶ The Norwegian national broadcasting corporation (NRK) has separate radio and television programmes in Sámi, and claims in this way to "contribute through an active programme policy in consolidating and developing Sámi language and culture on a national, regional and county level. ... Thus NRK takes on a crucial role for the preservation and development of Sámi language, culture and identity" (<http://www.samiradio.org>).

English speech via television, radio and film is crucial, and contributes greatly to the learning process, not least with regard to pronunciation.

3 Theoretical aspects

3.0 Introduction

This chapter explores some important concepts in relation to the study of linguistic variability, and provides a theoretical framework for the present investigation of RP. The present study examines a set of variable features in RP, some of which represent ongoing linguistic change. It is, however, also concerned with RP as a social and normative phenomenon. This chapter discusses topics related to both these aspects.

Section 3.1 deals with key concepts in the field of language variation and change, some of which will be operationalised in the analysis of the variables, and all of which will serve as a theoretical basis for the interpretation of the results. In addition to the descriptive variationist approach, a comprehensive study of RP also needs to consider the evaluative and prescriptive aspects of the accent. RP has a unique sociolinguistic status, in that it has been codified and is widely used as a reference accent and a model for teaching. Terms such as *standard* and *norm* are therefore highly relevant in connection with RP, and these concepts will be clarified and discussed in the last sections of the chapter.

3.1 Language variation and change

The abiding paradox of language is that in the world of Speech the order of the day is “Advance!”, while in the world of print the order is “As you were!”

A. Lloyd James (*Broadcast English II*, 1930)

3.1.1 General

An important foundation for quantitative linguistic research is the notion that language is inherently variable and changeable. Variation within a linguistic system or a speech community can be stable, or it can be an indication of change in progress. Weinreich et al. (1968) state that “not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity” (1968: 188).

Variation is here taken to mean any phenomenon involving more than one type of realisation for a linguistic form. Variation can be found at intra-speaker level, within the speech of a single individual, and at inter-speaker level, between individuals. The former is related to factors in the linguistic context or in the speech situation, while the latter typically correlates with social categories such as age, gender, class, etc. One of the major contributions of the quantitative paradigm has been to provide evidence that linguistic variability is patterned, and that change progresses in an orderly manner.

A main goal of variationist research is to establish correlations between linguistic variables and other variables within and outside the linguistic context, and attempt to explain these. In cases where the variation reflects ongoing change, explanation can focus on one of two problems: why the change might start (the actuation problem) and how the change spreads (the transmission problem) (cf. McMahon 1994: 44). Much sociolinguistic research has been criticised for producing descriptions of covariation between linguistic and social data and presenting these as explanations (cf. Cameron 1990). Demonstrating a correlation between variables does not necessarily imply a causal relationship. An explanation should in some way account for the preconditions of an observed phenomenon. There are however several different kinds of explanation, with varying degrees of explanatory power, and there is not universal agreement as to what constitutes an explanation. Lass (1980) defines explanation in deductive-nomological terms as strictly causal and based on universal, predictive laws. Such explanations are unattainable in linguistics, and, according to Lass, language change therefore cannot be explained. According to McMahon (1994: 45), linguists should not strive for this type of explanation, but accept that within linguistics, explanations “can be probabilistic or statistical at best”. Linguistic phenomena may thus only be explained if explanation is defined in more low-key terms as “relief from puzzlement about some phenomenon” (1994: 45, quoting Greenberg 1979: 279).

Language change is usually explained by reference to *internal* and/or *external* dimensions. Internal, or language-inherent, explanations focus on structural aspects, and argue that change is motivated by (and constrained by) a strive for regularity, symmetry, economy and “naturalness”. External explanations hold that the origins of linguistic change are social and that change is the product of speaker activity in social and cultural contexts (cf. e.g. Milroy 1992). The external domain includes social categories such as class and gender, but also language contact, social psychology, as well as attitudinal and evaluative factors. It can sometimes be difficult to distinguish clearly between the internal and external domains, and many researchers have questioned the dichotomy. Some reject it as false, usually with

reference to the fact that language is an inherently social phenomenon and cannot at any stage be separated from its social functions (cf. Andersen 1989: 10). Others suggest that internal and external factors are closely related and may interact in a single change (e.g. Dorian 1993). The exact nature of the relationship between the two dimensions and their role in linguistic change is, however, not fully known. Woods (2001) shows that attempts to link internal and external factors to different types or stages of language change have had mixed results. Her own analysis of vowel shift in New Zealand English clearly suggests that the two dimensions interact and that independent reference to either fails to explain the shift. While certain external factors may motivate change, system-internal forces may determine which sounds will change and in which direction (cf. Labov 1994).

Related to the question of the causes of change is the issue of whether sound change is teleological or not, and whether speakers select variants consciously or unconsciously. Some will maintain that speech production is governed by articulatory considerations (new forms may arise from the tendency to reduce articulatory effort), and that speakers choose variants based on the demands of the speech situation (cf. e.g. Lindblom 1990). Others argue that sound change is accidental and serves no purpose. Ohala (1993), for example, claims that much of sound change results from misinterpretations on the part of the listener. A central sociolinguistic view of change is based on the notion that speakers select variants as signals of identity: “individual speakers choose among the available variants of all the available variables in order to locate themselves in a highly complex multi-dimensional social space” (Hudson 1996: 147).

As part of the explanation of language change, Labov (e.g. 1972, 1994) distinguishes between ‘changes from above’ and ‘changes from below’. *Above* and *below* refer to levels of social awareness as well as position in the societal hierarchy. Changes from above are usually borrowings from more prestigious varieties, and are introduced by the dominant social class, often with extensive public awareness. Changes from below appear first in the vernacular and are typically introduced by a lower class. They are systematic, internally motivated changes that take place below the level of social awareness – it is only when they are almost completed that people become aware of them (cf. Labov 1994: 78). If this dichotomy is applied to RP, it is evident that any changes in this accent must be characterised as changes from below, since RP is the “highest” variety in social terms. It can, however, be questioned whether these changes always necessarily operate below the level of conscious awareness. It seems evident that changes in RP can be externally motivated, as they can in other varieties, and that RP speakers can experience social pressure to adopt new forms. This raises complex

issues related to concepts like ‘prestige’ and ‘standard’, as it implies that non-RP features may at some level have higher status than RP forms, even though their origin is “low” in terms of the socio-economic status of the speakers (see further 3.1.3 and 3.2).

3.1.2 Lexical diffusion

Phonological variables differ considerably in the extent to which they influence the language system. Some variables represent general phonological processes and affect the pronunciation of thousands of words, while others involve a small class of items, or even just a single word. Moreover, some phonological processes seem to expand by gradually incorporating more words, while others apply to all eligible items at the same time. The present section deals with the implementation of sound change, with particular focus on the theory of lexical diffusion. This issue is especially relevant to the discussion of CURE lowering, yod coalescence and t-voicing, which all display irregular lexical distribution (see further chapter 6).

The traditional Neogrammarian view of phonological evolution posits that every sound change is regular and exceptionless, and is conditioned only by phonetic environment.¹⁷ It is phonetically gradual, advancing by imperceptible phonetic steps, and lexically abrupt, operating simultaneously in all relevant words (cf. Labov 1981, McMahon 1994: 48-49). However, a number of studies have reported findings that contradict the Neogrammarian model of regular phonological change. For example, Trudgill’s investigation of the variable (u) in East Anglia, and the Milroys’ study of the Belfast (ʌ) variable, reveal extensive lexical variation which does not seem to be rule-governed and which cannot be explained in terms of general phonological differences between the words (cf. Chambers and Trudgill 1998: 161, Milroy 1992: 158-159). These and other findings provide support for the model of lexical diffusion, which assumes that a sound change is phonetically abrupt and lexically gradual: it operates in discrete perceptible phonetic leaps, and spreads differentially and gradually through the relevant vocabulary.

Labov (1981, 1994) has shown that both the Neogrammarian and the lexical diffusionist view of sound change are valid, but that neither can be universally applied. He

¹⁷ The exceptionless hypothesis presupposes that the change is *mechanical*. There are thus two important exceptions to the regularity law: analogical change, “which involves conceptual relations that are not mechanical (or phonetic) in character” (Labov 1994: 422), and dialect borrowing, “which is generally considered to involve social relations of relative prestige that are not mechanical” (1994: 423). The causal claim at the heart of the Neogrammarian view is nevertheless that lexical exceptions do not develop system-internally.

proposes two distinct types of sound change with different patterns of implementation. Regular sound changes are phonetically gradual and produce continuous variants. They are conditioned only by the phonetic context, and spread as predicted by the Neogrammarian model. Diffusing changes are typically discrete and categorical, little sensitive to the phonetic surroundings, but may be lexically and grammatically conditioned. As regular changes are typically “unobservable” and occur at a non-contrastive level, they will take place below the speakers’ conscious awareness. Diffusing changes, on the other hand, often involve an easily distinguishable contrast between two elements and a high degree of social awareness. Among the regular changes are vowel shifts in place of articulation, consonant changes in manner of articulation, and deletion of glides and schwa. Lexical diffusion is typically found with changes that involve shortening and lengthening of segments, consonant changes in place of articulation, and deletion of obstruents. While Labov recognises that “both regular sound change and lexical diffusion are active and productive processes” (1994: 502), he maintains that changes of the Neogrammarian type are in the majority and seem to represent the more fundamental mechanism of change.

Several researchers have challenged Labov’s position, and offer a different approach to sound change. Bybee (2001, 2002) rejects Labov’s dichotomy of regular versus lexical diffusion changes, and suggests the possibility that all sound changes diffuse gradually through the lexicon, affecting some words before others. The patterns of diffusion are primarily related to word frequency, in that high-frequency words undergo change at a faster rate than less common words. Bybee views sound change as mainly, perhaps entirely, articulatorily motivated, involving either weakening or assimilation. Such reductive processes result from the automation of linguistic production and the tendency to minimise articulatory effort. Since repetition enhances automation, it follows that reductive changes will be more advanced in highly practiced items, i.e. in the most frequent words and phrases.

The observation that phonetically motivated processes tend to affect high-frequency words first has been attested in a number of studies and for a number of phenomena. Bybee (2002) shows that also phonetically gradual changes may exhibit lexical diffusion. The frequency effect illustrates the relation between the use of linguistic forms and their cognitive representation. According to Bybee, high-frequency items have stronger representations in that they are more easily accessed (2001: 6). Strength of representation in turn leads to reduction of form. Greenberg and Fosler-Lussier (referred in Shockey 2003: 15) link the reduction of frequent items to the fact that the brain seems to process frequent words more quickly than infrequent words, and that the frequent items therefore do not need to be fully

specified in order to be understood. Similarly, Dressler and Wodak (1982) demonstrate through psycholinguistic experiments that speakers pay less attention to frequent words and therefore tend to pronounce these more casually than infrequent words. Jurafsky et al. (2001) show that phonetic reduction is also linked to a word's *probability*, in that words which are strongly predictable from the neighbouring words are more likely to be reduced.

The word-frequency effect in lexical diffusion suggests a deep involvement of phonology with lexicon and grammar, and provides evidence that words and frequent phrases are the units of mental storage (cf. Bybee 2000, 2001). As particular tokens undergo on-line reduction, the stored representation changes. The greater the number of modified tokens, the faster the change will take place. High-frequency words will thus be affected by sound change more easily and rapidly than low-frequency words.

Phonetic reduction is also related to speech style and level of formality, in that different styles do not “allow” the same level of reduction. Bybee (Hooper 1976: 102) suggests that the source of reductive changes is found in casual speech, where the level of speech-attention is low and restrictions on reduction are few. This means that words that are used in informal settings will change at a faster rate than “formal” words. These words are of course likely to be the same as those that also have a high overall frequency, since the most frequent words are familiar, everyday words, and constitute the substance of casual speech. Low-frequency words occur less often in casual speech and are therefore less exposed to the reductive processes.

Hickey (1999) observes an interesting social aspect of the implementation of sound change. His study of the Dublin Vowel Shift reveals that what he calls *motivated participants* – “fashionable” metropolitan Dubliners who “vigorously reject a confining association with low-prestige Dublin” (1999: 268) – show Neogrammarian advance for the shift, while *detached participants* – “socially conscious urbanites from outside Dublin” (278) – show lexical diffusion. This difference, according to Hickey, stems from the lack of motivation on the part of the detached participants. They “do not grasp the motivation among their metropolitan counterparts actively involved in the shift” (279), but adopt the vowel shift as they encounter it in words of high salience and frequency.¹⁸ It seems, then, that a sound change can spread differently with different groups of speakers, depending on their level of motivation for, or involvement in, the change. A parallel can perhaps be drawn to native

¹⁸ It may be argued that this is a case of *dialect borrowing* rather than lexical diffusion.

versus adoptive RP speakers, where the latter may be more likely to exhibit lexical diffusion for changes which apply across the board for native RP speakers.

It is obvious, then, that at least some sound changes advance by lexical diffusion, though the view that all or even most changes do, seems somewhat extreme. Bybee has provided strong empirical evidence of the influence of language use on lexical representation and phonological structure, and of the importance of frequency. She has shown that word frequency can be a conditioning factor in phonological variation and change and should be taken into consideration in usage-based analyses. Her results are, however, based primarily on studies of typically reductive phenomena, such as /t, d/ deletion and schwa-deletion, and may therefore be somewhat biased. Frequency-based lexical diffusion seems evident for phonetically motivated reductive changes, but is not so evident for other types of changes. A number of vowel shifts (particularly chain shifts) seem to follow the regularity principle (cf. Labov 1994). However, Labov's dichotomy of regular versus lexical diffusion changes seems too simplistic, as some changes exhibit properties of both types (cf. Bybee 2002, McMahon 1994: 58-68).

3.1.3 Accent levelling

The present section discusses the effects of language contact, which is one of the external factors that may lead to linguistic change.

RP seems to be increasingly influenced by other varieties, and incorporating features that originate in non-standard accents (cf. 4.2.2). This development can be seen in light of a nationwide trend towards accent levelling. Levelling has been defined as "a process whereby differences between regional varieties are reduced, features which make varieties distinctive disappear, and new features emerge and are adopted by speakers over a wide geographical area" (Williams and Kerswill 1999: 149). A number of recent studies have found evidence of extensive levelling of accents and dialects in Britain, causing speakers to sound more similar.

The growing tendency towards levelling and convergence is a product of the social and demographic changes that have taken place in Britain over the last century. These changes have been accompanied by a great increase in geographical and social mobility, resulting in radical changes in people's social networks, and wider dialect contact. The traditional close-knit local communities, thought to contribute to the preservation of linguistic norms (Milroy and Milroy 1997), are breaking down, rendering people more receptive to

linguistic change. The effects of these social mechanisms have been a general abandonment of traditional rural dialects, and the replacement of local features by more widespread forms (cf. Kerswill 2003).

With regard to RP, the general access to education and increased social mobility has on the one hand led to greater access to the RP accent for people who previously were excluded from entering universities and skilled professions. Gimson wrote in 1964 that “RP has become less and less the property of a particular social class (...) and, with the increasing blurring of social boundaries in Britain, is adopted or aimed at by a more and more considerable section of the population” (1964: 132). One result of the social changes has therefore been increased standardisation, and a widening range of modified speech down the social scale (cf. Kerswill 2001). At the same time, the prejudice against regional varieties is disappearing, and people no longer have to abandon their local accent as they move up the social ladder. Non-standard accents are thus increasingly heard in public and formal contexts, where RP used to dominate. The levelling process does therefore not necessarily lead to modifications towards the standard. In the social climate of modern Britain, RP is no longer the obvious target of accent modification, and non-standard varieties are exerting an increasing influence on language variation and change. Several recent studies point to accent levelling as the main factor behind current changes in British English accents, and many of the changes seem to originate in non-standard southeastern varieties (cf. Foulkes and Docherty 1999a).

Accent levelling is taking place both at a national and a regional level. Some pronunciation features are being adopted by speakers in all parts of the UK, such as t-glottalling, TH-fronting and r-fronting. These features are characteristic of non-standard accents of the southeast, and are believed to be spreading from this area to the rest of the country (cf. Kerswill 2003, Williams and Kerswill 1999). For other features, the levelling is more regional in character, often centring around major cities. Fronting of the vowels in GOOSE and GOAT, for instance, is presently affecting all accents in the south (cf. Kerswill 2001), and in the north of England there are signs of levelling towards a mainstream northern regional variety, exemplified by the generalisation of pronunciations like [e:] and [o:] in FACE and GOAT (cf. Watt 2002). This effect is happening in many regions, “with people ending up with a regional accent rather than a very local dialect” (Kerswill 2001: 54).

The best known example of a variety which covers a large region, is the so-called *Estuary English*, which is said to be a new mainstream accent of the southeast, centred around the Thames estuary. The term was coined in 1984 by David Rosewarne, who defined it as “a

variety of modified regional speech”: “a mixture of non-regional and local south-eastern English pronunciation” (Rosewarne 1984). Estuary English appears to be the most influential variety in the southeast of England, and Estuary-like features are also spreading beyond this region to other parts of the country (cf. Cruttenden 2001: 81, Crystal 2003: 166). Estuary English is based on the popular accents of London and the Home Counties, and has been described as an intermediate variety between Cockney and RP (Rosewarne 1994). The generalisation and rapid spread of the accent can be seen as an illustration of the current levelling process (cf. Kerswill 2001).

While Estuary English is said to be a “new” phenomenon, it merely represents the continuation of an old process, namely the influence and spread of popular London features. What seems to be new is the scale and rapidity of the process (cf. Kerswill 2001). This development is linked to the social and attitudinal changes that have facilitated the spread of non-standard features (cf. 2.3). One of the variants that have spread most extensively across Britain in recent years is t-glottalling. The social and geographical diffusion of this feature illustrates how a sound change can move from vernacular towards a prestige standard. T-glottalling originates in non-standard speech and has traditionally been strongly stigmatised. It is characteristic of the lower-class London accent, which is identified as the source of the spread (cf. Wells 1982: 323). However, recent studies (e.g. Milroy et al. 1994) provide evidence of a radical change in the social status of the glottal stop, which has now become supra-regional and quite generalised in contemporary British English, and also entered RP (cf. Fabricius 2000).

The mechanisms underlying accent levelling and the spread of popular features seem to be a combination of geographical diffusion, speech accommodation and attitudinal factors. Geographical diffusion involves the gradual spread of a feature across geographical space, and is the basis for most dialectological accounts of change. Linguistic features usually fan out from a populous and economically and culturally dominant centre (cf. Kerswill 2003: 223). The “centre of gravity” and principal source of linguistic innovation in Britain today is London (Wells 1982: 301). Some features, however, seem to affect large areas simultaneously, and do not follow the traditional patterns of diffusion. This has been observed with the spread of TH-fronting in low-status urban varieties (cf. Kerswill 2003: 235-237). In such cases, social psychological factors are presumed to be the main mechanisms involved.

At the level of individual speakers, levelling is believed to happen mainly through the process of speech accommodation, whereby a speaker adjusts to his interlocutor. The accommodation theory was developed by the social psychologist Howard Giles to describe

how individuals modify their behaviour under the influence of and in response to others. In conversational situations the sender may adapt his accent to that of the receiver by reducing pronunciation differences, as a way of gaining the receiver's approval. The result of this process is accent convergence, where one speaker alters his accent to become more like the other, or, in cases of mutual convergence, both speakers modify their accents to become more similar to each other (cf. Giles and Powesland 1997). Such individual acts of accommodation may over time become permanent and, if the accommodation takes place on a sufficiently large scale, lead to language change (Trudgill 1986: 42).

The linguistic result of levelling is a reduction of forms. The features that normally disappear are those that are unusual or in a minority, while the forms with the widest geographical and social usage are retained (cf. Trudgill 1986: 98). Very local forms, or the most socially conspicuous forms, are therefore most likely to be lost. The levelling process may thus give rise to new varieties characterised by the absence of localised forms. Estuary English has been described as the result of accent convergence where the middle-class RP speaker accommodates “downwards” and the working-class local accent speaker moves “up-market” (cf. Rosewarne 1994). Estuary English is thus a “middle-of-the-road”, unmarked, variety that attracts speakers from both ends of the social spectrum. Traditional RP is socially marked, and those who wish to avoid the stigma of sounding “posh” will often accommodate. Author and RP-speaker India Knight admits to abandoning her native accent in the appropriate contexts: “Put me on Radio 4 and I speak normally. Stick me in a taxi and my natural accent completely disappears” (Knight 2001). Similarly, Crystal (1995) quotes a businessman commenting on the drawbacks of speaking RP: “if you were unlucky enough to have such an accent you would lower it. You would try to become more consumer friendly” (327).

It is generally assumed that the loss and adoption of accent features requires dialect contact and face-to-face interaction (e.g. Trudgill 1986). There is evidence, however, of externally motivated changes and levelling taking place apparently without any interpersonal contact. Recent studies have revealed the presence of southeastern pronunciation features in cities as remote from London as Glasgow and Hull (cf. Stuart-Smith 1999, Williams and Kerswill 1999). Moreover, these features are adopted primarily by working-class adolescents who have few opportunities for social or geographical mobility. It has been suggested that the broadcast media, and in particular television, may be a contributory factor in these changes (cf. Timmins and Stuart-Smith 2005). Young people throughout Britain are regularly exposed to non-standard London-flavoured accents through the media, and come to associate these

accents with youth culture (Williams and Kerswill 1999: 162). Accent features can thus be actively acquired because they are associated with certain values with which the speaker identifies. The process involved is then imitation rather than accommodation. Attitudinal factors thus play an important part in changes that are difficult to explain by traditional diffusion models. And the media may enhance attitudes and make people positively disposed towards a new feature.

With the increasing accent levelling, speakers tend to evaluate features more in terms of their territorial spread, and less in terms of class. Highly localised forms may be negatively evaluated as old-fashioned and parochial, while supra-local features, regardless of their social origin, are seen as modern and outward-looking. This awareness is then the motivation for adopting new features. Speakers will use the more widespread forms in order to signal a “modern” identity. At the same time, the incoming features should not signal too much disloyalty to local norms (cf. Watt 2002, Foulkes and Docherty 1999b: 14). Although RP is a non-regional variety, it has negative connotations for many as “posh” or “snobbish”. Features that originate in non-standard accents seem to have a broader appeal, and are therefore more likely to be widely adopted. In today’s Britain, then, the more broadly based popular varieties are increasingly taking over the function once reserved for RP, that of disguising origins. And RP itself is increasingly “exposed to dilution from regional characteristics” (Gimson 1964: 132). The levelling process can thus be seen to affect accents at all points of the social continuum. The overall result seems to be a parallel development, where the edges are being rubbed away from the most conspicuous varieties at both ends of the social scale. In RP, this is reflected in the gradual incorporation of features that until recently have been associated with non-standard speech (see further 4.2).

3.1.4 Gender

That women would be hopeless as broadcast announcers is the considered opinion of a number of officials of the BBC.

(British newspaper report, 1926)

One of the most important social correlations of speech variation is gender. Gender is also one of the main non-linguistic variables considered in the present study (the others are style and channel affiliation). The sample has been equally divided between male and female speakers, and the linguistic usages of the two groups will be compared and tested for statistical

significance. Gender is not expected to be relevant for all the variables, only for those that are believed to have sociolinguistic salience (such as t-voicing and yod coalescence). Gender as a social category is particularly interesting in a study of the speech of an otherwise very uniform group, such as the newsreaders in the present study. Any differences between the genders that are observed here can then be presumed to be the result primarily of this single factor. Gender may interact with other social variables (class, occupation, etc.) to the extent that women tend to have different jobs than men, lower incomes, different social networks, etc. In a study of newsreaders, the effect of these other variables is minimised, because all the speakers have the same occupation, and thus belong to the same socio-economic group. Moreover, they produce their speech under the same circumstances and in the same stylistic settings.

Gender as a social variable is based on the biological sex of the speakers. As such, gender is a relatively straightforward category, in that it is unproblematic to define, readily observable, and there is no doubt as to how the speakers should be categorised (in contrast to for example social class, which is notoriously difficult to define). And there are only two categories: male and female. As a factor in sociolinguistic analysis, however, gender is conceived as the social construction of sex (cf. Eckert 1997a, Cheshire 2002). Biological sex is the underlying basis for the differentiation of roles, norms and expectations that apply to men and women, and it is these social and cultural factors which constitute gender. The differences in patterns of linguistic variation between men and women are then a function of gender and only indirectly a function of sex. Speakers are typically classified in terms of their biological sex, but the findings are interpreted in terms of the social roles and self-images of men and women.

In a study of phonetic and phonological variation at inter-speaker level, gender is a relevant factor, because it has been found to be one of the most important non-linguistic influences on pronunciation. The remaining part of this section will briefly present the main sociolinguistic findings on language and gender, which form the basis for the expected gender differences in the present study. To what extent the gender observations for broadcast RP can be related to the general patterns found in previous studies, will be discussed in chapter 6 and section 7.3.

A number of sociolinguistic studies have found correlations between linguistic differences and the sex of the speaker. One striking pattern which has emerged in virtually all these studies is that, other things being equal, females use more high-status or standard variants, and fewer low-status forms, than males do. Hudson (1996) refers to this as the

Sex/Prestige Pattern, and characterises it as “one of the most robust findings of sociolinguistics” (193-194). The Sex/Prestige Pattern seems to apply to all age groups and classes, and most languages. Labov (1990) provides an extensive survey of examples from a number of studies which illustrate the sexual differentiation in linguistic usage. The general principle that women use more high-prestige forms than men needs to be somewhat refined, though. Labov (1990) limits the generalisation to *stable* sociolinguistic variables, i.e. variables that are not involved in linguistic change, and Hudson (1996: 194) adds the refinement that the variables must be socially stratified, with higher classes using the standard variants more often than the lower classes. Women then prefer features with overt prestige, while men may be more influenced by the covert prestige of low-status forms. Gender differences within English are mostly phonetic and phonological, and generally so small that most people are not consciously aware of them. The differences are quantitative, not qualitative, and represent statistical tendencies rather than absolute distinctions (cf. Trudgill 2000: 69-70).

Although gender differences are generally believed to occur only with linguistic variables that are stratified according to class, there is evidence of gender contrasts also with features that have not been found to be class-related, such as variables that represent phonological processes of reduction or lenition (cf. 3.1.5.3). Some research has shown that women tend to favour fully articulated “canonic” forms and avoid forms that show the effects of casual speech processes. Zue and Laferriere (1979) observed this pattern in their study of medial /t/ in American English. They found that women more frequently articulated the “full” aspirated [t] and had significantly longer phonetic durations than men, who produced shorter consonants with fewer distinctive features (1979: 1047-48). This tendency may in some way be linked to notions of correctness, as full, phonetically explicit forms often tend to be regarded as more “correct” than reduced forms (cf. 3.1.5.3).

There is no single or universally accepted explanation for the different linguistic behaviour of men and women, but several interpretations have been proposed (cf. discussions in e.g. Holmes 1992: ch. 7, Hudson 1996: 195-199, Trudgill 2000: ch. 4, Labov 2001: 275-279, Chambers 2002: ch. 3). It has been suggested that because of the traditionally different social roles of men and women, men derive their status from their work, while women tend to assert their status through symbols, such as linguistic forms. According to the *network* explanation, men use more low-status forms (which equal local forms) because they are more tightly integrated into the local community, since they tend to work outside the house more

than women.¹⁹ Another explanation is based on the notion that working-class speech is associated with “roughness” and masculinity, which may lead men to be more favourably disposed to non-standard forms than women, who prefer the “sophisticated” middle-class forms.²⁰ Moreover, women’s traditionally greater role in the socialisation of children may lead them to be more sensitive to social norms of “correct” behaviour, and more aware of the social connotations of speech.²¹

In the case of RP, evidence suggests that there may be more social benefits for women who use this accent than for men. Elyan et al. (1978) found that females who use an RP accent rather than a regional accent were evaluated more favourably in terms of both masculine and feminine attributes. The researchers conducted a study using the matched-guise technique to determine evaluative reactions to female speech in RP versus the Lancashire accent. The results revealed that listeners perceived the RP-speaking women to have higher self-esteem, be more fluent, intelligent, self-confident, adventurous, independent, strong *and* feminine than the regional speakers. In addition, the RP women were judged as more likely to have a well-paid, prestigious job and an egalitarian relationship with their spouse. The RP-speaking females were also perceived as possessing some less favourable masculine traits, such as aggressiveness and egotism, but at the same time they were rated higher on the femininity traits, such as gentleness and sensitivity. RP-speaking women are thus seen as more masculine as well as more feminine than women with regional accents. Elyan et al. argue that the female RP accent is “the voice of perceived androgyny”, and suggest the following explanation for why RP-speaking women are perceived in this fashion:

the RP accent is seen as the voice of sophistication and social power, and the voice of those who are in a position to exert some control or influence over their social environments. In the case of women then, such attributes may be seen to allow them a considerable opportunity both of accentuating their femininity by use of expensive fineries while at the same time succeeding in male-dominated pursuits. (Elyan et al. 1978: 130)

¹⁹ These explanations seem less relevant today, with jobs being much less gender-segregated. Newsreading, for instance, used to be reserved for men, while today there are just as many female as male presenters.

²⁰ Self-evaluation tests have shown that women *are* indeed more sensitive to prestige norms, as they significantly *over-report*, i.e. claim to use more standard forms than they actually do. Men, on the other hand, tend to under-report, and thus aim at non-standard speech (cf. Coates 2004: 62-63).

²¹ Although there is no single explanation for the Sex/Prestige Pattern, it seems clear that it is not an inherent, necessary, or biologically determined pattern, but a culturally based phenomenon. This is evident from the fact that there are counter-examples which illustrate that the situation is not unchangeable (cf. e.g. Chambers and Trudgill 1998: 86, Mathisen 1999: 121-122).

RP-accented females thus seem to be perceived favourably on a greater variety of evaluative dimensions than RP-speaking males, and this positive evaluation may provide another reason why women use more standard forms than men.

In stable situations, then, women tend to be more conservative and prefer the features with overt prestige. In unstable situations, women are often in the lead of linguistic changes that are in the direction of the prestige variety, while lower-class men typically lead changes that introduce or revitalise vernacular forms (Chambers and Trudgill 1998: 84). The female lead in changes that are standardising is expressed in Labov's Principle 3: "In linguistic change from above, women adopt prestige forms at a higher rate than men" (Labov 2001: 274). These changes either involve the introduction of a new feature from outside the speech community or the re-distribution of prestige forms within the community. They take place at a relatively high level of social consciousness, and show increased frequency in formal styles (cf. 3.1.1). Since changes from above are similar to stable sociolinguistic variables, it is to be expected that the gender patterns are similar for the two types. However, most linguistic changes are below the level of social awareness, and women seem to be leading many of these changes too. According to Labov (2001) "[w]omen have been found to be in advance of men in most of the linguistic changes in progress" and "the cases where men are in the lead form a small minority" (280, 284). Women are thus on the one hand more conservative than men, as they use more standard forms, but on the other hand more progressive, as they adopt new innovative forms more quickly. The overall pattern is nevertheless that women are in the forefront of sound changes that are not stigmatised.

Several variationist studies have identified women as the main linguistic innovators (cf. e.g. Labov 1990, Watt and Milroy 1999). Young women have for example been found to lead in the diffusion of t-glottalling in various speech communities (cf. Mathisen 1999, Holmes 1995, Milroy et al. 1994, Mees and Collins 1999)²², even though t-glottalling is a vernacular form which traditionally is strongly stigmatised. These findings are in apparent contrast to the Sex/Prestige Pattern and the typical gender differentiation. There is however strong evidence that t-glottalling has lost much of its social stigma and is developing as a new prestige norm, as its status has changed from a local vernacular variant to a generalised non-regional variant. The female use of t-glottalling may then be seen as an adjustment to what has become a prestige form (cf. Mathisen 1999: 114). Alternatively, as Milroy et al. (1994) suggest, it is the female preference for a feature that brings about a reversal of the evaluation

²² Fabricius (2000), however, found no significant gender difference in the use of t-glottalling with her young RP-speaking informants.

of that feature. According to this interpretation, gender is prior to class, and females *create* the prestige forms, in the sense that the forms females prefer become overtly prestigious in the community (1994: 26).

Social patterns of linguistic variation and change have traditionally been interpreted primarily in relation to class, which has been seen as the most important social factor. There are however several studies that have found greater differences between the genders than between classes (cf. Milroy et al. 1994, Watt and Milroy 1999, Watt 2002, Coates 2004: ch. 4), which suggests that, in many communities, gender may be the primary marker and the strongest single conditioning factor for the distribution of linguistic features. Several recent studies have also challenged the traditional class-based dichotomy between standard and non-standard, and instead described linguistic variables in terms of their territorial spread (cf. 3.1.3). Speakers seem increasingly to recognise variants as local versus supra-local, and males and females show markedly different patterns of usage, with males preferring the more localised variants and females using the more widespread and unmarked variants (cf. Milroy et al. 1994, Watt and Milroy 1999, Stoddart et al. 1999). This, according to Foulkes and Docherty (1999b), “does not mean that females necessarily orient themselves towards the standard more than males, but rather that they are more susceptible to influences from any kind of non-local forms” (16). Gender differentiation may then be best accessed in terms of local versus supra-local forms, and even gain in generality, as suggested by Watt and Milroy (1999: 43): “The correct generalisation appears to be that men tend to retain localised forms (which may or may not be stigmatised) and women adopt supra-local forms (which may or may not be prestigious)”.

Woods (1997) adds a further dimension to the discussion by linking women’s preference for supra-local forms to a general tendency for women to be more cooperative and listener-oriented than men. This tendency is attested in several studies where women have been found to accommodate more to the needs of the addressee and to organise interaction with the aim of providing support and showing solidarity (cf. e.g. Holmes 1986, Coates 1998). Women’s speech, then, according to Woods, “is characterised by precisely those features which come into play when speakers of different dialects interact” (1997: 117), and these strategies may lead women to innovate changes that are motivated in contexts of dialect diversity and contact.

In the case of RP, which is both overtly prestigious *and* non-regional, a distinction between the traditional notion of standardness and the concept of territorial spread is less relevant to the gender patterns, since all the features available to the speakers are by definition

supra-local (cf. 2.4.7). Any gender differentiation within modern RP will presumably be found mainly in the use of the most recent features, which have entered the accent from below (cf. 4.2.2).

3.1.5 Style

Stylistic variation is one of the main factors of quantitative linguistic research. The stylistic dimension involves intra-speaker differences, and considers how changes in the situation affect speech. The present section first discusses the meaning of the term *style* and how the concept is relevant for the study of broadcast RP. It further outlines some of the main stylistic findings that have been produced by quantitative studies, and the principal models of interpretation. Section 3.1.5.4 discusses stylistic aspects of broadcast speech in general and newscasts in particular (see also 5.1.3).

3.1.5.1 Stylistic variation

As a technical linguistic term, *style* is somewhat imprecise and has an ambiguous reference. Some researchers refer to style as an independent variable which affects and explains linguistic variation. It is more common, however, to consider style as the linguistic outcome of situational changes, so that style is itself one axis of sociolinguistic variation, and is correlated with certain extralinguistic variables (cf. Bell 1984: 146). A ‘formal style’ is then speech which is characterised by a high level of standard or prestige forms. The use of such forms tends to increase in formal *situations*, and there is thus a close, but not necessary, correlation between context and linguistic form. It is crucial to distinguish between the linguistic code and the extralinguistic factors which potentially affect the code, and use *style* to refer to only one of them. In the present study all speech variation is within a single variety, RP, which represents the ultimate level of high-status or standard forms. The amount of linguistic variation is therefore limited, and can hardly be said to represent different speech styles. In the discussion of intra-speaker variation (see chapters 5 and 6), the term *style* will therefore be used to refer to context, situation or circumstances, independently of linguistic usage. Variations in these factors then may or may not affect the informants’ speech patterns.

The newsreading situation is relatively static, and there is very little stylistic variation in the presentation of broadcast news. Occasionally, however, the presenter interviews a subject, either in the studio or via satellite, and this situational shift is hypothesised to affect

the use of certain pronunciation variants. The main intra-speaker aspect will therefore be the comparison between presentation style and interview style (see further 5.1.3). Furthermore, there are some differences between the three channels that can be described as stylistic, such as the amount of commercial breaks and the number of presenters in the studio (cf. 5.1.2), which may influence the linguistic choices of the newsreaders. However, these factors cannot be measured at intra-speaker level, and will therefore not be analysed as stylistic variation.

Variationists have traditionally considered style-shifting to involve shifts in the usage levels for features that are arranged along a standard – non-standard continuum. The basic pattern revealed in most studies is that speakers of all classes increase the level of high-status forms in more formal contexts, while vernacular forms dominate in casual familiar contexts (cf. e.g. Trudgill 2000: ch. 5). Intra-speaker stylistic variation can thus be said to reflect inter-speaker social variation: features that are more common in the lower classes will also be more common in informal styles, while speakers shift towards the norms of the upper classes in formal situations. This close association of linguistic forms with social context presupposes a shared evaluation of linguistic variants in the community. According to Bell (1984, 1997), stylistic differentiation *derives from* social differentiation, in that any linguistic variable which shows style shift is always subject to social evaluation in the speech community, and no linguistic feature varies on the stylistic dimension only. However, situational variation is also found with features that reflect “natural” phonological processes and that are not socially stratified. For example, a decrease in the level of situational formality will often involve an increase in the amount of phonetic reduction (see further 3.1.5.3).

3.1.5.2 Explaining stylistic variation

Attention to speech

There has been some debate about how best to analyse and explain stylistic variation (cf. Milroy 1987: ch. 8, Schilling-Estes 2002). According to Labov (e.g. 1972), the main factor responsible for situational changes in speech is the amount of attention speakers pay to their own speech. In informal contexts people are less conscious about their speech, while the level of awareness and speech-monitoring increases in formal situations. As the amount of attention increases, so does the use of standard features. Speech is then ranked along a linear continuum from casual to careful, with reading style, where attention is directly focused on speech, as the most careful. The Attention to Speech approach has been criticised on a number of grounds. It has been pointed out that speech-consciousness is difficult to quantify, and that empirical

foundation for the attention variable is lacking (cf. Bell 1984: 148). Moreover, there is no necessary correlation between the use of standard variants and the degree of self-monitoring. It is quite possible for speakers to direct attention to producing any type of speech, not just high-status forms. Milroy, for example, observes that her Belfast informers do not consistently use more standard features when reading (Milroy 1987: ch. 8). It seems evident, however, that attention has some influence on the application of certain phonological processes that are articulatorily motivated (see further 3.1.5.3 below).

Audience design

An alternative approach to interpreting style-shifting is the Audience Design model, proposed by Bell (e.g. 1984). According to Bell, people design their speech for their hearers, and stylistic variation is essentially speakers' response to their audience. Audience Design is closely related to the Speech Accommodation theory, which states that speakers tend to adapt their speech towards that of their receivers, in order to win their approval (cf. 3.1.3 above). Bell's main piece of evidence for the Audience Design framework is his study of the speech of radio newscasts in New Zealand (Bell 1982, 1984). He found that single presenters who read the news on two or more different stations changed their pronunciation in a systematic and consistent way as they moved between the stations. When presenting on the most "serious" channel (coded YA), which broadcasts national news bulletins, the use of standard forms increased, while there was a systematic shift towards more non-standard forms when they presented on the "lighter" commercial and local stations. Audience design is the only tenable explanation for such individual speaker shifts. All the other variables likely to influence style shift, such as topic, studio setting and degree of attention, were the same. Only the (actual and target) audiences were different, with the YA listeners tending to be older, better educated and have professional careers (Bell 1982).

Similar evidence of addressee accommodation is reported by Coupland in his study of the speech of an assistant in a Welsh travel agency. He found that the assistant shifted her speech patterns to match the differences between the clients she talked to (Coupland 1980). In situations of personal interaction, speakers can thus respond directly to the personal traits and linguistic patterns of their interlocutor. What is unique about broadcasting is that it involves a notional audience that is absent from the speech situation. The values and characteristics of the addressees are unknown to the communicator, and can only be assumed.

Bell recognises that other situational factors may influence speech style, but argues for the primacy of the audience variable. Style shifts seemingly related to non-personal factors

such as topic or setting are subordinate to and derive from audience design, in the sense that speakers associate particular topics or settings with certain types of addressees (cf. Bell 1984: 181).

Both audience design and attention to speech seem to affect speakers' use of linguistic features, but in different ways. Just as the style shifts Bell observed with radio newscasters are unrelated to variation in attention, differences between spontaneous speech and reading or between reading a text and reading a word list (attested in numerous sociolinguistic studies), are difficult to explain in terms of audience design.

3.1.5.3 “Natural” phonological processes

As indicated above, there seems to be some correlation between attention and speech style, but not necessarily between attention and standardness. Attention-based differences are more likely to reflect differences in phonetic explicitness or degree of phonetic reduction (cf. Dressler and Wodak 1982, Lindblom 1990). Variation in attention may therefore affect the use of “natural” phonological processes such as assimilation, elision, weakening and epenthesis, which often occur in fluent continuous speech. These connected speech processes are “natural” in the sense of being phonetically motivated: “[t]hey are likely to be explicable in terms of vocal tract characteristics and of the motor control mechanism” (Kerswill 1985: 2). Connected speech features generally enhance articulatory ease but reduce perceptual distinctness. According to Shockey (2003), most connected speech phonology is subconscious and therefore not affected by stylistic variation. Other linguists have found evidence that the phonetic motivation interacts with extra-linguistic factors. The correlation between phonetic explicitness and attention to speech has been investigated by Dressler and his colleagues (see Dressler and Wodak 1982). Through a number of experiments they demonstrate that less attention results in more phonological reduction, and vice versa. The application of connected speech processes thus seems to be directly related to attention to articulation, which again is determined by situation or level of formality.

Connected speech processes are often referred to as *allegro* or *fast speech* rules. According to Shockey, however, “[r]esults are not yet conclusive about whether increase in speech rate increases the amount of phonological reduction” (2003: 11). There is some empirical evidence of increased reduction in fast speech (cf. Fosler-Lussier and Morgan 1999), but this effect seems to be subordinate to speaker attention. Studies have shown that speakers are able to control phonetic explicitness at all rates, and avoid reduction in fast speech by focusing on articulation (cf. Kerswill 1985, Dressler and Wodak 1982). This

indicates that connected speech processes are not just automatic subconscious mechanisms, but that they have the potential of being manipulated.

In spite of their physiological foundation, many connected speech processes are variety-specific, and they may be involved in linguistic change (cf. Dressler and Wodak 1982, Kerswill 1987). Connected speech processes therefore have the potential of being sociolinguistically salient. Nolan and Kerswill (1990), for example, found evidence of sociolinguistic differentiation for certain connected speech processes with speakers in Cambridge. Moreover, phonetic explicitness is popularly associated with “correctness”, while low-status or non-standard speech is often characterised as “sloppy” or “lazy”. Kroch (1978) also links resistance to phonetic reduction to notions of social prestige. He claims that low-status varieties tend to be articulatorily more economical than the prestige variety. Speakers of high-status varieties, according to Kroch, tend to resist changes that reflect processes of phonetic conditioning and articulatory simplification as a way to mark themselves off from speakers of lower class varieties.

Careful articulation is thus highly evaluated, and may therefore by many be considered appropriate in formal contexts, in much the same way as “traditional” high-status features. Pronunciation variables which represent natural phonological processes are then likely to be influenced by situational factors which affect the degree of articulatory care, and they are hence susceptible to stylistic variation.

Among the features analysed in the present study, intrusive [r] and t-voicing are both phonetically motivated processes that also show evidence of social salience. Intrusive [r] is inserted to avoid a vowel hiatus, but is widely regarded as incorrect or slovenly, while t-voicing is a weakening process that enhances articulatory ease, but is also socially stratified (see further 4.4.3-4.4.4).

3.1.5.4 Stylistic aspects of broadcast speech

In broadcast speech, the stylistic aspect is relevant on many levels. Newscasting is one of the most formal broadcast genres, both in terms of setting and topic. The presenters talk directly into the camera, on very serious topics. The speech is mostly scripted, and the amount of speech-consciousness and self-monitoring is presumably very high. These characteristics are likely to have a profound effect on the linguistic choices of the newsreaders (cf. 3.3). For instance, there is a high focus on clarity and comprehension, and the usage level for phonological processes of the reduction or lenition type may therefore be low, as these processes tend to worsen perceptibility.

All interaction probably contains some pressure to accommodate or seek the addressee's approval. This factor is magnified in public situations such as broadcasting, which addresses an enormous audience. The prime need of mass communicators is to win and hold the audience in order to maintain their audience size or market share. Audience design therefore pervades all aspects of broadcasting. According to Bell (1984), the language of newscasts is suited to the audience to such a degree that individual speech differences are minimised, and newsreaders tend to converge on a common "corporate style". Stations with similar audiences thus share consistently similar speech styles.

Broadcasting is one-way communication, and thus lacks the direct feedback of face-to-face interaction. The adjustment of newscasters to their target addressees, demonstrated in Bell's study, can be said to be a type of speech accommodation, but does not necessarily entail linguistic convergence. The most important factor determining the linguistic choices of the presenters seems to be shared evaluations of linguistic variants. Bell notes that the newscasters on the YA station in New Zealand tend to use RP features, and thus largely diverge from their audience's actual speech. However, since RP is regarded as a prestige model by the YA audience as well as by the newscasters, it is interpreted as positive accommodation. Such style shift towards an external reference model is referred to as *referee design* (Bell 1984: 186-196). Referees are "third persons not physically present at an interaction, but possessing such salience for a speaker that they influence speech even in their absence" (1984: 186). In broadcasting situations the referee variety is known or assumed to hold prestige for the audience as well, and referee design is thus a form of indirect audience design. This mechanism is believed to be an important factor contributing to the extensive use of RP on the three news channels investigated here (see further 3.3).

A discussion of newscast speech raises another important issue related to style, namely the nature of read or scripted speech. News presenting involves mainly scripted and carefully prepared speech, very different from the spontaneous vernacular language produced in everyday conversations. The processing and production of words in conversation is a very different activity from reading aloud written words. Milroy (1987: 173-178) argues that speech and reading are not directly comparable and should be analysed as separate areas of behaviour. One of the fundamental differences between the two styles has to do with the awareness of spelling. When reading, speakers may be more influenced by the orthographic representation of the words, and use the spelling as a guide to pronunciation. The immediate presence of the written forms is then likely to reduce the amount of phonetic simplification

and reduction. According to Shockey (2003), most stylistic variation that involves connected speech features or phonological reduction is found between scripted and unscripted speech:

Degree of formality seems to have little effect on unscripted speech: one finds the same types and nearly the same number of reductions in formal English as one does in casual speech. ... The impression that formal speech is less phonologically reduced than casual speech is probably based on the fact that much of (if not most) formal speech is scripted rather than spontaneous.
(Shockey 2003: 17)

The patterns of phonological variation that one finds in scripted speech may then be somewhat different from what is expected based on descriptions of conversational speech.

Newsreading is then predominantly scripted speech, where the performance aspect is foregrounded. It has a mainly ‘transactional’ function, and the premium on clarity of communication is therefore high. Newscasting is thus a rather formal format, which is reflected in the language used. At the same time, however, newscasting also illustrates how norms of formality have changed. Today, a much wider range of linguistic forms are “allowed” in formal contexts, and the public manner of speech has become more relaxed and colloquial over the past decades. According to Quirk (1982: 8) newsreading is perhaps the genre where the style of presentation has shifted most radically. This shift is congruent with changes in the audience’s attitudes and tastes. Quirk talks of “the new free-wheeling, breezy style of presentation” (1982: 13) and the increased focus on immediacy and authenticity, which do not combine well with a too careful or elocutionary reading style. There has thus been a noticeable increase in the level of informality. The role of the presenters is more one of *telling* us the news rather than just reading it, as if they were addressing the audience personally. Brown (1990: 6) characterises the “modern” newsreading style in the following way:

even BBC newsreaders ... [now] read the news for the most part in a casual, relatively informal style, just like a member of the public might. ... Generally ... they speak as naturally as if they were talking to someone they know, rather than to a huge, anonymous audience which might include listeners for whom English is not the first language.

The stylistic changes in newscasting have considerable linguistic consequences, the most important of which is the entrance of non-RP accents (cf. 2.6). Moreover, there is evidence of increased tolerance towards RP features that previously would be considered inappropriate in newscasting, such as intrusive /r/ (cf. Burchfield 1981).

3.2 The concept of ‘standard’

People who talk about standardising English, and especially those who accuse the B.B.C. of aiming at a dead level of uniformity in the speech of their Announcers, should think twice before they speak. ... There will never be a standard “æ” or a standard “t” or a standard “s” or a standard English.

A. Lloyd James (*Broadcast English II*, 1930)

This section discusses the concept of linguistic standard in relation to RP. It presents an outline of various definitions of the term *standard* as used in the field of linguistics and with reference to English in particular. ‘Standard’ is often linked to evaluative notions such as correctness and prestige, but more formal definitions focus on invariance of form as the most important characteristic of standard varieties. In light of the various criteria for defining a standard, I will discuss to what extent there exists a spoken standard in Britain and whether RP is such a standard.

3.2.1 A spoken standard?

RP is often referred to as a standard, or as the British English standard accent. The term *standard* is however often left undefined, and the concept tends to be taken for granted. There is also disagreement as to whether there can be such a thing as a spoken standard. O’Donnell and Todd (1991), for example, state that “it is important to realise that RP is not a standard pronunciation; there is, in fact, no such standard” (39-40). Honey (1985), on the other hand, holds that “we cannot escape from [RP’s] significance as a standard” (257) and that we should not deny “the existence of a ‘standard’ version of anything simply because in its purest form it is difficult to define or measure” (256).

The term *standard* is used with reference to both written and spoken language, but a *standard language* is normally understood as a written language (cf. Deumert and

Vandenbussche 2003b: 3-4). In the written mode the term refers to the fixity of spelling, lexicon and grammar, codified as a set of established rules. A spoken standard may then be established based on the norms of the written model. Phonology, on the other hand, is more resistant to standardisation, because of the extremely variable and changing nature of pronunciation. In England a nationwide standardisation of the written language began as early as the fifteenth century with fixation of spelling and morphology. Codifications of an English pronunciation model did not take place until the late nineteenth century (see Nevalainen 2003). The problems associated with standardisation of phonology is reflected in the constant new editions of the English pronouncing dictionaries.

Language standardisation is always a linguistic process of variation reduction (cf. Deumert and Vandenbussche 2003b: 3). The process is however motivated by social or political interests which shape the direction of the standardisation. Haugen (1997) sets up a four-step model of language standardisation:

1. selection of norm
2. codification of form
3. elaboration of function
4. acceptance by the community

Language standardisation typically begins with the selection of a variety as the basis of the standard. The chosen variety, if it does not already enjoy prestige, will gain prestige, along with the people who use it. The second stage is the codification process, whereby the rules and conventions of the variety are explicitly and officially formulated in manuals, dictionaries and grammars. Where codification aims at minimal variation in form, elaboration refers to those activities which are aimed at maximising the functional reach of the standard variety. Finally, the standard must have a body of users. The diffusion and acceptance of the norm by the relevant population is an essential part of the standardisation process.

Haugen's analysis of the factors typically involved in standardisation has been widely accepted by other linguists (see e.g. Trudgill 1999), and the process is typical for how a national standard language develops. One controversy in the English standard debate is the question of whether or not there is such a thing as a spoken standard of English, and moreover, if RP is such a standard. Even though most scholars define standard varieties as primarily written varieties, there is no doubt that RP shares many of the features associated with standard languages, and that it has undergone at least some aspects of standardisation.

Historically, RP was, at least partly, the result of a process of standardisation. In the eighteenth and nineteenth centuries there was a growing preoccupation in Britain with linguistic correctness and “good” versus “bad” speech. This preoccupation extended to phonological features, and by the end of the eighteenth century the process of imposing one particular form of speech as a “standard” was under way. As with the written language, it was the accent of the social and educational elite in the southeast that was selected and regarded as the “best” English. The concept of a spoken standard developed throughout the nineteenth century, and some linguists even saw it as a potentially uniform mode of speech which could unite the nation. Even if uniformity was never achieved, the pronunciation of the educated was taken as a standard to be emulated, and a new sense of the term *standard* emerged, signifying a level of excellence to which speakers should strive (cf. Crowley 2003).

From 1870 onwards the conformist practices of the new public school system and the prestige universities promoted the use of the “standard” accent, which was explicitly taught and prescribed. In the early twentieth century the accent was labelled *RP* and underwent codification, and Daniel Jones published his pronouncing dictionaries. Jones recognised the inherent variability of spoken language and was wary of setting up a standard for how people ought to speak. Nevertheless, his pedagogical purposes forced him to select one single variety as the basis for his model, and he was instrumental in the construction of a particular accent as the “standard” to be met (Crowley 2003: 146). *RP* gradually lost its regional identity and dominated the public spoken arena well into the twentieth century, not least in broadcasting. Only in the 1960s did regional accents make their way into the BBC, and at the same time the public schools and the Church of England relaxed its “requirement of a prescribed and relatively homogeneous accent” (Milroy 1999: 188).

In the past, then, *RP* enjoyed many of the privileges normally reserved for written standard varieties: it was selected as a prestigious model of correctness, it was codified and explicitly taught, it spread geographically and became non-regional, and it received strong institutional support. The present-day situation is somewhat different, and attitudes and ideologies have changed. Crowley (2003: 259) sums up the current scholarly situation in the following way: “there is a lot of confusion, little consensus and a good deal of scepticism towards the idea of a ‘spoken Standard English’”.

Even though *RP* no longer enjoys the status it had in the nineteenth and twentieth centuries, the accent is still frequently referred to as a standard accent in the linguistic literature. ‘Standard’ is however often linked to the notion of ‘prestige’ or ‘correctness’, and standard forms equated with those forms typically used by speakers from the highest social

class, and used with more frequency by all speakers in formal styles. When RP is called a standard, then, it is normally in the sense of “the highest prestige variety”. Wells (1982: 34) characterises RP as a standard accent, which he defines as

one which, at a given time and place, is generally considered correct: it is held up as a model of how one ought to speak, it is encouraged in the classroom, it is widely regarded as the most desirable accent for a person in a high-status profession to have.

He further states that RP “enjoy[s] overt prestige. People agree that this is the correct way of – speaking [...]. Such an accent is a de facto standard” (104). Similarly, Upton et al. (2001: xii) describe RP as “a standard of pronunciation which is generally considered correct and is also used as a model for the teaching of English to foreigners”. Cruttenden also describes RP as an “implicitly accepted standard of pronunciation” and “the result of a social judgement ... as to what is ‘correct’” (2001: 78-79). The prestige attached to a standard variety is always overt. So-called non-standard varieties are then accents or dialects with little or no overt prestige, but which may enjoy covert prestige.

The traditional correlation with standard varieties and prestige is complicated by the fact that the evaluative distinction between ‘standard’ and ‘non-standard’ may be changing. The traditional dichotomy, which was based on social class, is being challenged by the emergence of new prestige varieties with “low” origins. Tore Kristiansen (2001) has shown that among young Danes, the notion of ‘Standard Danish’ is increasingly moving in the direction of traditionally low-status Copenhagen speech, and that they operate with two language “standards”, each of which is prestigious in its own way. The traditionally high-status Copenhagen accent is a standard for the domain of education and business, while “low” Copenhagen is a standard in the modern broadcast media and popular culture. In the evaluative judgements of the young Danes, the traditional distinction between Status and Solidarity (cf. Chambers 2002: 254) does not seem to apply. Instead, Kristiansen suggests an alternative division, where prestige and excellence are perceived in terms of ‘Superiority’ and ‘Dynamism’. His parameters cut across the traditional dichotomy and combine the traits in new ways. The Superiority scale incorporates traditional Status traits like *ambitious* and *gifted* along with the traditional Solidarity traits *trustworthy* and *pleasant*. These qualities are judged as characteristic of the “higher” voices. The “lower” accents, on the other hand, score high on the Dynamism scale, which combine the traditional Solidarity qualities *interesting* and

straightforward with the Status traits *independent* and *efficient*. If we apply these comparative scales to the situation in Britain (cf. 2.3), it seems that RP retains prestige in terms of Superiority, but is facing increasing competition from “dynamic” non-standard varieties. Kristiansen’s study also illustrates the need for a specification of what is understood by terms like *standard* and *prestige* in discussions of accent evaluation (cf. 2.4.4).

According to James Milroy (2001a) the categories of ‘standard’ and ‘prestige’ should be kept separate, and he rejects the evaluative definitions of the term *standard*. Even though a standard variety has prestige, high prestige is not necessarily definitive of what constitutes a standard. Milroy instead defines a standard variety as characterised by uniformity in structure, and standardisation as a process which suppresses optional variability. According to this strict definition, there cannot in practical usage be such a thing as a fully standardised variety, as total uniformity is never achieved in practice. Uniformity is rather seen as the *goal* of standardisation as a process (2001a: 534). A standard variety thus becomes an idea in the mind, “a clearly delimited, perfectly uniform and perfectly stable variety – a variety that is never perfectly and consistently realized in spoken usage” (543). This idealised, fictive standard is a result of what Milroy and Milroy (1999) call the “standard language ideology”, which dominates standard language cultures (such as Britain) and which has the effect of developing among speakers a consciousness of a “correct” or “canonical” form of language. For analytical purposes it is important to keep the two aspects of the standard separated. Descriptions of the structural properties of a variety give us information about linguistic usage. This can be very different from people’s notions and perceptions about the standard, which are shaped by beliefs and ideologies.

Most linguists recognise that standardisation involves variation reduction, but they do not necessarily, as the Milroys, reject the notion of a standard language as a real linguistic product of standardisation (cf. e.g. Deumert and Vandebussche 2003b). Holmes (1992) simply describes standard varieties as codified varieties (83). Codification is of course a means to achieve uniformity and stability, so codified varieties may have a higher degree of uniformity than other varieties. However, a variety need not necessarily have total uniformity in order to be referred to as a standard.

In Britain the issue of linguistic standards is complicated by confusions around the phrase *Standard English*. Just as with the term *standard*, there is no general consensus as to what constitutes Standard English, or even if such an entity exists. The consciousness of the special status of RP as a prestigious accent and its traditional association with the “best English” complicates the picture even further.

Trudgill (1999) defines Standard English as a dialect, distinguished linguistically from other dialects of English by its grammatical forms. It is the variety of English normally used in writing, and it has nothing to do with pronunciation; Standard English can in principle be spoken with any accent. He further characterises Standard English as a purely social dialect associated with the education system and spoken by “educated people” at the top of the social scale. It is a non-regional dialect, even if its origins are in the southeast of England.

Most linguists today follow Trudgill’s characterisation of Standard English and will hesitate to use the term *Standard English* with reference to RP or any other accent. There is however no doubt that many laypeople operate with a concept of a correct or standard accent of English, and that this is identical with RP.²³ Although RP cannot be equated with Standard English itself, there are striking similarities between the two, and it is not unreasonable to say that RP to some extent is the phonological equivalent of the Standard English dialect: a non-regional, codified social variety associated with educated speakers. RP does, however, not have the same status as Standard English, which is an international norm for written English.

Linguists should, then, be very careful with using the term *standard* about modern RP, at least without specifying how they define the term. It can be a useful concept when we need to distinguish RP from so-called non-standard (i.e. vernacular) varieties. The problem is that *standard* is a loaded term with prescriptive connotations, which potentially signals that RP is in some way superior to other accents. RP could instead be referred to in more neutral terms as a codified variety or a reference accent.

3.2.2 C-RP vs. n-RP

In a discussion of RP and standards, it is important to keep in mind that RP exists at two different “levels”, namely as a codified model in dictionaries and textbooks, and as a living accent spoken by a certain percentage of native English speakers. Fabricius (2000) refers to these two “variants” of RP as c-RP (constructed RP) and n-RP (native RP) respectively (see 2.4.3). The codified version of RP is the closest we come to an English speech standard. It is a “frozen” representation of a variety which is constantly changing. Codification promotes stability and uniformity and provides a model for reference and teaching purposes.

²³ Or increasingly, perhaps, so-called Estuary English, as illustrated by the title of Paul Coggle’s book from 1993: *Do you speak Estuary? The new Standard English*.

In her discussion of standard languages Bartsch (1987: 258) makes a similar distinction between a prescribed standard and actual usage. She points out that from a normative point of view the standard is considered an invariant, fixed point, whereas from an empirical point of view the standard is a range which includes variation of form. The two can be reconciled by seeing the empirical range as structured from an imagined point of reference. The prescriptive standard has an empirical reality in that it plays a role as the ultimate model towards which speakers are oriented. It is identified by linguists, but it is more a construct or something postulated than something real.

Smith (1996: 66) also tries to reconcile real usage with the normative standard in his description of RP. He regards RP as standardised or *focused* rather than standard or fixed. In apparent contrast to this, though, he states that RP is a “magnet” of pronunciation towards which speakers tend to varying degrees, but which, he claims, nobody reaches. He therefore concludes that RP is a mere abstraction, a “centripetal norm” which nobody speaks in its “pure” form. His description of RP as focused rather than fixed is in line with the linguistic reality. His conclusion, on the other hand, which implies a refusal to recognise RP as a living variety, is somewhat extreme. Rather than seeing RP as an abstract and unattainable entity (which means that there are no RP speakers, only near-RP speakers), the lack of fixity is a result of the fact that RP incorporates variation and is in a constant state of change, so that the borders around the accent are fuzzy (cf. 2.4).

It is possible to unite the two “levels” of RP without necessarily discarding the notion of RP as a “real” variety. C-RP is merely an attempt to represent n-RP visually in a systematic way. The dictionary descriptions of RP are necessarily simplified, but they have always incorporated some variability, and thereby at times reflected ongoing changes. A dictionary variety will, however, inevitably lag somewhat behind the current usage. In principle, though, c-RP is based on observations of n-RP. Real speakers are the models for c-RP, and not the other way around. Rather than speakers striving to reach the standard of c-RP (ordinary people do not read pronunciation dictionaries), the descriptions of c-RP strive to keep updated and in line with the current usages in n-RP. Consequently, n-RP is the “primary” version of the accent (see also 3.3.1).

3.2.3 Summary

The term *standard*, as we have seen, is used in two main senses: ‘uniformity’ and ‘level of excellence’. The former refers to internal structure and is a purely linguistic criterion. The latter is a social, speaker-based and evaluative category. A *standard variety* can refer to a concrete linguistic variety characterised by a high degree of uniformity in structure, which has undergone codification, and which normally enjoys high prestige. Alternatively, a standard variety is seen as a fictional notion in the minds of speakers, a “social myth constructed for ideological purposes” (Bex and Watts 1999b: 9). In the written mode, *standard* refers to the fixity of spelling and grammar. With reference to the spoken mode, standard language is, in the words of Smith (1996: 65) “an extremely complex and notoriously loaded term”.

It is problematic to talk of *standard spoken English* if by that we mean a set of highly codified rules of pronunciation and a fixed, uniform mode of speech. Such a thing does not exist. There exists, however, a partly standardised and codified accent, namely RP, which serves as a reference accent in the phonological literature, and as a model for teaching English pronunciation to foreigners. Although RP is partly standardised, it is not to be equated with Standard English, which is a matter of spelling and grammar. To the extent that *standard* is used synonymously with ‘overtly prestigious’, RP can arguably be characterised as a standard variety. However, such an evaluative definition assumes that all the members of the speech community agree in the social evaluation of speech forms, which is not necessarily the case. The term *standard* further carries with it evaluative notions of correctness which are reminiscent of the old days of linguistic hegemony, and should therefore be used with caution. It is important to keep in mind, however, that while there may be no standard in linguistic terms, *ideologies* of a standard can persist, and largely do in the case of RP.

Crystal (2003: 431) combines several of the elements discussed above in his definition of standard as “a prestige variety of language used within a speech community, providing a unified means of communication and thus an institutionalised norm”. Crystal links *standard* with the concept of ‘norm’, a term which is highly relevant in an analysis of RP, and which will be discussed in detail in section 3.3.

3.3 The concept of ‘norm’

The present study investigates the speech of a group of news presenters with strikingly similar pronunciation. Whenever such a regularity in linguistic behaviour can be observed within a community, a relevant question is which factors influence the speakers and determine their linguistic choices. An important term in this context is the *norm*. This chapter discusses the norm concept in relation to spoken language in general and to RP in particular. The first sections present some definitions of the term *norm* along with a discussion of norms as a social and psychological phenomenon. In section 3.3.3 the discussion of norms is related to the study of broadcast RP. I discuss the reasons why RP is so prevalent in newscasts, and present the results from a small-scale study of norms conducted among a group of BBC World news presenters.

3.3.1 What are norms?

The term *norm* is often used in much the same way as *standard*, to refer to a prestigious linguistic variety, and RP is commonly described as a norm (see e.g. Wells 1982 and Crystal 2003). *Norm* is however also used to refer to the social “forces” underlying the linguistic behaviour within a group.

Oxford English Dictionary (<http://dictionary.oed.com>) defines norm as “a standard or a pattern of social behaviour that is accepted in or expected of a group”. A linguistic norm, according to Crystal (2003), refers to “a standard practice in speech or writing” (319). A norm thus represents a form of consensus which regulates linguistic usage. Labov (1972: 120-121) refers to norms as the primary criterion in his definition of the speech community:

The speech community is not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms; these norms may be observed in overt types of evaluative behaviour, and by the uniformity of abstract patterns of variation which are invariant in respect to particular levels of usage.

According to Vannebo (1980) there are three common uses of the term *norm*, all of which are closely related and partly overlap. Firstly, *norm* is used to describe a certain behavioural pattern accepted within a given community. Secondly, the term refers to

evaluations or judgements as to what is considered correct or incorrect, good or bad, beautiful or ugly, etc. Thirdly, norms can be seen as behavioural rules which constitute social pressure, and the breaking of which can lead to social sanctions. Bartsch (1987: 4) describes norms as “the social reality of the correctness notions”. A norm implies an expected regularity, but, she argues, it is the element of sanction or criticism against deviations which constitutes the normative force and distinguishes a norm from a mere custom or convention (1987: 166).

Sundby (1974: 19) sees the norm as a *relational* concept: a norm is always a norm *for somebody*. In this perspective, a uniform behavioural pattern within a group can be seen as an indication of the existence of certain norms within the group, but the behaviour itself is not the norm. The norm is rather what leads to the conformity. The norms that regulate written language are primarily explicit, prescribed norms (or rules), established and codified by an authority with normative competence. Spoken language, on the other hand, is normally regulated by so-called *internalised* norms. These norms are established by way of an inner motivation, in that the individual identifies with or has a psychological affiliation with a social group and adopts their values and behaviour as his own. This process may be conscious or unconscious.

Spoken usage is then not regulated in the same way as written language. In the case of RP, however, we can to a certain extent talk of prescribed norms in relation to speech: the RP of pronunciation dictionaries (c-RP) can be seen as a prescribed norm for how this variety is to be pronounced, with descriptions of individual sounds and their lexical distribution. However, this norm differs from the typical prescribed norms that regulate written language, in that it does not have the same regulatory status, and native speakers do not consult pronunciation dictionaries as a guideline for how they should speak in the same way as they might consult spelling dictionaries and grammars. C-RP is a guideline primarily for linguists and foreign learners of English. To the extent that RP is the variety that native speakers modify towards and adopt in formal situations, it is not as the result of a prescribed norm. This means that adoptive RP speakers must be exposed to the accent in the community, through the mass media, or both.

3.3.2 Studying norms

If we consider norms in relation to spoken language, a uniform speech pattern within a group is indicative of the existence of a speech norm. But where exactly is the norm situated, and

can it be observed directly? In a linguistic context what we can observe are the systematic patterns of usage within speech communities. This regularity of behaviour is a collective phenomenon, which indicates that the norms are situated above the individual. However, the process of internalisation so characteristic of speech norms, takes place in the mind of the individual. Sandøy (2003) holds that norms exist within the individuals as part of their conception of the community. Norms are then defined as that which the individual *believes* is expected by the community (Sandøy 2003: 172). When these beliefs are shared by a large number of people, the result will be seen as a collective pattern. This means that the norms cannot be observed directly; we can only observe what we assume is their linguistic outcome. The norms can however be studied indirectly, by observing people's behaviour, by interviewing speakers and asking them about their attitudes towards language, or by means of the so-called *matched guise technique*, which is used to elicit people's evaluative reactions to various accents (cf. 2.3).

Who or what, then, contributes to the formation of the norms? Among the most important factors are our perception of what is common or normal, other people's reactions to our behaviour, and expressed attitudes and stereotypical views about language (cf. Sandøy 2003: 173-174). If we disregard prescribed norms, which are promoted by overt and institutional enforcement, norms are thus often subconscious mechanisms that are psychologically motivated and shaped by our social surroundings.

According to Mæhlum (2003: 88) the primary 'norm agents' are family, friends, educational institutions, and to a certain extent the mass media, which all influence the individual in different ways. Not all members of the community contribute equally to the formation of the norms; some members are stronger and more influential than others because of their status or role. In many cases a speaker relates to different, often competing norms, which he may follow depending on the social situation. Romaine and Reid (1976) distinguish two types of sociolinguistic norm: 1) the community norm, which is the norm of the community to which the speaker belongs; and 2) the social norm, which refers to the wider social acceptability of linguistic variants, of which speakers are explicitly aware.²⁴ A speaker who identifies with a group outside his native community or has aspirations to "get on in the world" may develop negative attitudes towards his native accent or dialect and adopt standard variants. Conversely, in order to mark a distance to the wider community or solidarity with his local community, a speaker may increase his use of conservative or broad variants in his

²⁴ In the case of native RP speakers the two will presumably converge, as adherence to either norm will lead to the use of RP.

native variety (cf. Chambers 2002: 234). As far as RP is concerned, there is no doubt that newsreaders are among the most important norm agents. Their speech is heard by large numbers of people on a regular basis, which means that their pronunciations are likely to be strongly represented in people's minds. News presenters further have a status as "serious" and professional communicators, which is likely to add to their influence and prestige. Newsreaders are indeed often referred to as typical representatives of RP speech, which means that there is an awareness about their pronunciation. Newsreaders do not directly affect the way people talk in their everyday life, but they may influence people's conception of "standard" pronunciation or "serious" speech (cf. 2.6).

Language attitudes reflect the social and cultural structures in society as a whole. In order to understand the function of a linguistic norm we have to consider the social and cultural values associated with various linguistic variants. Knowledge of these values is part of the individual's motivation for internalising certain norms. Kerswill (1996: 181) refers to this knowledge as *sociolinguistic competence*, and defines it as "a person's ability to recognize language varieties within the community, to evaluate them socially, and to exploit them in the communication of social meanings". By using a certain variety the speaker signals adherence to the group who regularly uses this variety and an affiliation with their value system. Language can thus be seen as one of the most important factors in shaping our identity. In the case of RP, there is no doubt that the social qualities associated with the accent are part of what motivates speakers to use it. It is difficult to determine to what extent the adoption of a linguistic variety is a conscious process, and just how freely an individual uses his sociolinguistic competence. The impact of social structures on people's values and behaviour has been studied extensively by cultural sociologists like Pierre Bourdieu. The latter will argue that the core values of the dominant culture are manifested in every detail of an individual's behaviour and that the incorporation of social structures in people's everyday life is an unconscious and unreflected process (cf. Bourdieu 1991).

3.3.3 A broadcast norm?

If we consider the group of television news presenters as a community, there is no doubt that a certain uniformity of linguistic behaviour can be observed, in that the great majority speak

RP or near-RP.²⁵ The consensus on pronunciation among newsreaders is also observed by Roach and Hartman, among others (see Jones 1997: v). This similarity in speech behaviour is indicative of a common norm (cf. 3.3.1). Some of the presenters are probably native RP speakers who have had this accent since early childhood. In view of the small percentage of native RP speakers in Britain (roughly estimated at 3-5 per cent by Trudgill in 1974) it is statistically highly unlikely that this applies to all the RP-speaking presenters. Information about the background of various presenters further indicates that some of them are adoptive RP speakers. They are no doubt educated speakers, but in today's Britain being educated is no longer the same as having an RP accent (cf. 2.4.5).

As a community, newsreaders are particularly interesting because they have such a uniform accent without ever interacting with one another. They do not form a social community outside the context of work, and they are rarely at work at the same time. This means that the possible mutual influence they have on each other is only indirect, and that the norm originates from elsewhere. The broadcasting companies have no expressed requirement that instructs the presenters to use RP (cf. 2.6). In other words, there is no explicit prescribed norm at work that demands the use of any particular accent. The widespread use of RP/near-RP is therefore presumably the product of an internalised norm, since presenters adopt or modify towards the RP accent without explicit external pressure. In this context the evaluative component of the norm concept is highly relevant: in light of the formal nature of the newscast genre and the status of RP, it is reasonable to assume that evaluative judgements of RP as the "best" accent, and the fear of being criticised or perceived as uneducated or incompetent may lead to an indirect or self-imposed pressure to use RP.

In order to test the hypothesis that the widespread use of RP is the result of an internalised norm, I carried out a small-scale sociolinguistic study among some of the newsreaders from BBC World in September 2004. The use of RP is particularly prevalent among the news presenters on this channel: with the exception of a couple of non-British newsreaders, all the presenters speak RP or something very close to it. This group was therefore chosen as the informants for the study, as all the members seem to share a common norm. Moreover, the BBC is the only broadcasting company with its own pronunciation unit. I interviewed three RP-speaking news presenters and the staff at the BBC Pronunciation Research Unit. The aim was to gain some insight into the presenters' attitudes towards their

²⁵ This conclusion is based on observations of newsreaders on a selection of British television channels that broadcast nationally and internationally. Regional accents are more common on other channels, especially local ones, and RP is certainly not as hegemonic as it was some decades ago (see further 2.6).

own pronunciation and identify some of the factors which influence their linguistic choices and lead to the extensive use of RP.

During the interviews I tried to uncover the informants' level of consciousness with regard to accent and whether they thought any accents would be inappropriate for presenting the world news. Among other things I asked them if they had changed their accent since they started as presenters, if they had received comments about their accent from colleagues or viewers, and whether they avoided certain pronunciation features. The results of the interviews to a large extent confirmed the theory of an internalised norm. Accent is not a topic that is discussed among the staff at the BBC, and the company has no stated policy regarding accent. They are very concerned with diction and clarity and with pronouncing foreign names correctly. The BBC Pronunciation Research Unit spends most of its time researching and advising the BBC on how to pronounce words and names in foreign languages. Accent, on the other hand, is, at least officially, not an issue.

I spoke with three presenters who all had an RP accent: one female presenter (informant A), and two male presenters (B and C).²⁶ Informants A and B had acquired the accent as adults; one grew up in India, the other is from the Midlands. The third informant, who is from London, did not answer the question of whether his accent was adopted or modified in any way. They all said that they had made no conscious decision to speak with a certain accent when they started working as presenters, and that they were not instructed to talk in a "BBC way". Informant C said early in the interview:

I don't think I sat down and said "I'm joining the BBC, I've got to talk in a certain way".

But later on he suggested that there may after all be a subconscious mechanism influencing his linguistic behaviour:

Before I go on air ... I suppose there is a subconscious signal which says to me "I've got to talk in a certain way".

Informant A claimed that the uniformity of accent among newsreaders is a result of the BBC's hiring policy:

People tend to speak in a very RP sort of way. But that's to do with the BBC's hiring, it's not really to do with the individuals adjusting the way they speak.

²⁶ Informant C is also among the 30 speakers in the sample, while A and B are not.

Informant B also suggested that the widespread use of RP is a management issue, but at the same time he admitted that during his time as a newsreader, his accent had “evolved” away from his native Midlands accent. According to the head of the Pronunciation Research Unit, the BBC does not deliberately hire (or avoid hiring) presenters with certain accents.

All the informants maintained that accent was not important in newscasting, as long as the audience could understand what was being said. They recognised the fact that there are no broad regional British accents on BBC World when it was pointed out to them, but they all agreed that a regional variety would be perfectly acceptable for presenting the world news. The interviews revealed, however, that negative attitudes towards localised accents do exist, partly among the presenters themselves, partly among the audience. Informant A told me about how some of the non-British presenters with non-RP accents had received e-mails from viewers complaining about their pronunciation. Informant B had a negative view of his native Midlands accent, which he described as “hard” and “flat”, and he admitted that he had struggled to “get rid of” the northern [ʊ]:

I remember very early on in my career when I first started reading the news, that I was always having a difficulty with the letter “u”. Because up in the Midlands it would come out as a very hard [ʊ], [bʊs], instead of [bʌs]. And I do remember for quite some time having a struggle getting rid of that.

The same presenter also admitted that some regional accents would not be suitable for presenting the world news, but only because they would be unclear or hard to understand. BBC World has a large audience for whom English is a foreign language. In that context there may be a general feeling that non-regional forms (i.e. RP) are more appropriate with regard to comprehension, although this is not a stated policy (cf. 2.6). There was certainly an awareness that BBC World differs somewhat from the domestic channels, as expressed by informant B:

I think the BBC does pride itself on people speaking well and correctly as far as possible. Certainly on BBC World. On the national channels it is different, because they do use people from all over the UK, and those people do speak in their own regional way, and they might not actually speak “correctly”. [...] I think it probably doesn’t matter on the domestic channels [...] but on BBC World, no, it’s certainly different for us.

There is then a difference between BBC World and the domestic channels in the use of regional accents, but this difference is not based on any stated language policy. The general impression from the interviews is that the motivation to use RP primarily comes from the presenters themselves, possibly as a result of an indirect pressure or unspoken expectations,

and presumably based on the social values associated with RP and the constant exposure to the accent from “serious” broadcasters. It seems, in other words, that the presenters have internalised a norm. The head of the Pronunciation Unit suggested that if presenters use RP or near-RP, it is their own choice:

Maybe the conformity you observe is the result of people’s own decisions, and those are fed by what they see around them or what they think they have to conform to. [...] Maybe their perception of what an announcer sounds like informs what they come to sound like.

The long tradition of using RP in newscasts may have instilled in them a notion of how a newsreader should sound, and a belief about how they are expected to talk. Informant C seemed to confirm this theory when asked about the uniform accent of the BBC World presenters:

I would say that it’s blatant... plagiarism is the wrong word, it’s blatant piggybacking. You absorb the morning news programmes and you hear a pronunciation, and it sticks. It sticks somewhere in the software, and with me it sticks long term. So it’s a passive rather than an active relationship.

The traditional association between RP and news broadcasting is of course a result of the prestige and status that RP has enjoyed in Britain and the social values assigned to the accent. A number of studies have shown that RP scores highly for qualities like authority and competence (cf. Giles et al. 1990). The accent has typically been used by people who are educated, powerful and influential. Many of these are public figures whose voices are heard in the mass media. Furthermore, there is a strong correlation between accent and social situation. A shift towards the standard in more formal contexts is a very general finding of much sociolinguistic work (cf. 3.1.5). The more formal the situation, therefore, the stronger the expectation to use a standard variety. By using a non-standard accent the speaker risks not being taken seriously (cf. Mæhlum 2003: 95-96). The news presenters’ motivation for using RP is then twofold: the wish to come across as educated, “serious” and competent, and the fact that they are in a public and rather formal setting – a context which traditionally favours the use of RP. In the newsreading situation, these factors may be further enhanced by expectations from the audience: according to Bell (1983), the language of broadcast news is often seen as the embodiment of standard speech, and “listeners have cast broadcasting in the role of guardian of the standard language” (1983: 38). When asked about whether they felt any “responsibility” to maintain a certain “standard” of pronunciation, one of the presenters denied that notion completely, while the other two recognised that this might play a part, but

mainly at a subconscious level. Informant B suggested that pronunciation is part of the overall image of the presenter:

I think all the way through, certainly since day one when I came across here, pronunciation... I mean I was guided towards the Pronunciation Unit, and that is just a superb reference point for us. So that was important. Looking smart was another aspect to it. So, there were certain parts of getting the presenter on air which was crucial, and still is, to BBC World.

In Britain, “having no accent” has been associated with status and education. To speak without an accent means to have no regional pronunciation features, i.e. to speak with an RP accent. The BBC World informants were definitely aware of the status of non-regional speech. When asked to label their own accent, two of them said that it was non-localisable, and that other people could not tell which part of Britain they were from. The third presenter labelled her accent as “a middle-class educated person’s accent”, but also linked class with presence or absence of regional features:

Accent variation is more pronounced the more working-class you are. If you try and position yourself as being part of the intellectual class you get rid of the accent altogether.

While RP is associated with authority and competence, regional accents typically score highly for personal qualities like friendliness and sincerity (cf. 2.3). Informants B and C expressed an awareness that their accent could be perceived as posh by some, and that it would not be appropriate on local radio and television channels. Informant C said:

I don’t think I would be employed on Radio Newcastle or BBC Newcastle or Tyne Tees because I’d be seen as a snotty southerner. And it would turn a lot of people off: “oh, he’s not one of ours”.

In a local setting an RP accent might create a distance to the viewers or listeners. BBC World, however, has a different audience and a different “image”, where information and journalistic integrity are central (cf. 5.1.2). They regard themselves as disseminators of “high quality” information for an educated audience, and accent may, subconsciously, be part of this image. It is noteworthy that on all the news channels I have investigated there is a marked difference between the “serious” world news, where RP dominates, and the sports news, where the tone is informal and colloquial, and regional accents are much more common. It would seem, therefore, that the subject matter and the assumed background of the audience affects the language of the presenters. This can be seen as a manifestation of Bell’s notion of

audience design (cf. 3.1.5), which means that the presenters use the accent that they believe, or assume, will suit their audience. It may diverge from the audience's own speech, but it is a divergence on which the communicator and audience agree. This suggests that the social evaluation of RP is shared by the larger speech community.

In light of the "serious" image of BBC World, it is not unreasonable to assume that the presenters use RP because they wish to be associated with the social values assigned to RP as the accent of the "educated". Alternatively, the presenters believe that they are expected to speak RP, and this creates an inner motivation to adopt the accent, if they do not already have it. This awareness or acceptance of what is considered "appropriate" behaviour is to a large extent an unconscious mechanism, and can be seen as part of the presenters' sociolinguistic competence.

3.3.4 Summary

In a linguistic context norms are interesting to the extent that they affect and regulate linguistic behaviour. Norms are often seen as a collective phenomenon which influences the individuals. It is however impossible for a norm to exist independently of the individuals who internalise it. When a norm becomes internalised, the individual adopts the value system and behaviour of a particular group. Linguistic behaviour is then seen as the realisation of, or approximation to, a norm, which in its turn results from a particular set of attitudes. This mechanism can result in a society's negative attitudes towards a variety being adopted even by those who use that variety. An example of this was seen with the presenter from the Midlands who expressed negative attitudes towards his native accent, and who had adopted an RP accent.

Norms have been defined as that which the individual believes is expected from the community. This may lead to a sort of "self-policing", where the individual lives up to (real or imagined) expectations. Interviews with newsreaders and staff from BBC World suggested that the widespread use of (adoptive) RP is largely the result of an internalised norm. There is no expressed demand or requirement for newsreaders to speak RP. Nevertheless, the accent hovers in the background as the variety most presenters use, adopt or modify towards. Two of the presenters admitted to using an RP accent that was not their native variety, and one of them revealed that he had modified his accent after he started working as a newsreader. The fact that RP is so widely used in newscasts suggests that it still enjoys a certain status as the

accent of the educated, and therefore may be perceived as adding credibility to “serious” broadcasting such as newsreading.

4 The variables

4.0 Introduction

The present chapter gives a detailed presentation of the six pronunciation variables that constitute the main focus of the study. The descriptions outlined here will provide the background for the discussion of the analyses and results in chapter 6.

The first section offers a brief introduction to the concept of the linguistic variable, while section 4.2 gives an overview of the variation and change that can be found in RP. The main bulk of this chapter is section 4.4, which provides a detailed account of each of the phonological variables studied here.

Throughout the discussion there will be references to so-called *lexical sets*. The concept of lexical set involves using a keyword to represent a vowel phoneme and all the words that are realised with this vowel (e.g. DRESS represents all the words that have /e/ in RP). Keywords for lexical sets are written in small capitals. The idea of standard lexical sets was introduced by Wells (1982), who established a number of sets based on the vowel correspondences between RP and General American. The advantage of using lexical sets is that they enable us to refer concisely to large groups of words and to the vowel they share. Thus GOAT refers to words like *boat, home, goal*, etc., as well as to the vowel phoneme these words have (/əʊ/ in RP).

4.1 The linguistic variable

One of the main achievements of sociolinguistics has been to show that linguistic variation and change is not completely random or free, but is constrained by both social and linguistic factors. This insight was achieved primarily as a result of the invention of the linguistic variable as an analytical tool for handling linguistic variability. The linguistic variable can be defined as “a linguistic unit with two or more variants involved in covariation with other social and/or linguistic variables” (Chambers and Trudgill 1998: 50). Crucially, the variant realisations of the same variable do not have different referential meanings. Such linguistic

units which are subject to social or stylistic variation are also most susceptible to change in the long term, and the study of linguistic variables is therefore important to the understanding of linguistic evolution (cf. Labov 2001).

Linguistic variables at the level of pronunciation are generally referred to as phonological variables. Hudson (1996: 170), however, distinguishes between phonological and phonetic variables, according to the level at which the variants differ. The former involve variation where the same lexical item has alternative phonological structures, and the variants represent different phonemes or presence/absence of phonemes, e.g. /ʊə/ vs. /ɔ:/ in CURE words, or /r/ vs. Ø in *car, heart*, etc. Phonetic variables involve allophonic variation, where one phoneme has alternative phonetic realisations, e.g. [əʊ] vs. [pʊ] for /əʊ/ in *goal, hold*, or [t] vs. [ʔ] for /t/ in *bottle, network*. The present study investigates variables at the phonological as well as the phonetic level, but will refer to both types as phonological variables.

Linguistic variables are usually symbolised graphically by the use of parentheses, so that e.g. (t) means ‘variable t’, which may have the variants [t] and [ʔ] in intervocalic positions. Consonantal variables are generally easier to deal with than vowel variables. Consonant variables usually have discrete variants that are easily distinguished, while vocalic variants are most often continuous, i.e. ranged on a phonetic continuum of vowel quality, with no obvious boundaries between the variants (see further 5.2.2).

In the study of linguistic variation, the concept of the linguistic variable allows for quantitative research and comparative studies. By examining how often variants are used and how their use correlates with the social background of the speakers, we can explore the nature of linguistic variation, and compare speakers with each other, or compare texts produced by the same speaker in different situations.

Many linguistic variables have been found to correlate closely with social characteristics such as age, gender, class and social situation, such that the different variants have different social “values”, and these variables have been the main focus of sociolinguistic research. The sociolinguistic view is that speakers consciously or unconsciously choose among different pronunciation variants in order to place themselves in a complex multi-dimensional social space, and the purpose of sociolinguistic research is then to test hypotheses about relations between linguistic and social variables (cf. Hudson 1996: 147). The social values of speech variants can contribute to explaining much of the linguistic variation that can be observed in all speech communities. The class differentiation in a linguistic variable, for

example, seems to be directly linked to stylistic variation: the variants used by the upper classes are ascribed the highest status and prestige, and in formal situations speakers of all classes tend to increase their use of higher-status variants, so that these variants also become the “formal” variants. Variables which are subject to social differentiation as well as stylistic variation are referred to as *markers*, while variables that do not exhibit systematic stylistic variation and convey little social import are referred to as *indicators* (Chambers and Trudgill 1998: 72). Speakers are generally more sensitive to and aware of variables that are markers. These variables often represent on-going change, and they may be subject to overt stigmatisation, either because they involve a phonological contrast or a divergence between pronunciation and orthography. A typical marker in English is h-dropping in lexical words like *house* /aʊs/, *ham* /æm/, which has low social prestige and is generally avoided in formal style (cf. Cruttenden 2001: 192).

In the same way that social factors can determine linguistic variability, many phonological variables are also conditioned by linguistic factors. Unlike the accommodatory allophones of phonemes, the variants of variables are not fully predictable by the linguistic environment, but may be more likely to occur in certain phonological, morphological, lexical or syntactic contexts. The study of variability thus often combines linguistic and non-linguistic elements, since both function as *variable constraints* and participate in determining linguistic variability (cf. Chambers and Trudgill 1998: 130). The establishment of a linguistic variable thus amounts to an empirical assertion of covariation, either within or without the linguistic system (cf. Labov 1972: 162).

4.2 Variability in RP

Men of undoubted education and intelligence differ in pronunciation from one another, from pronouncing dictionaries, and from my own habits.

Alexander J. Ellis (*Early English Pronunciation*, 1874)

The basis for the present study is the assumption that RP comprises variation at several levels of its realisation. In the following this variability will be looked at in some detail. Section 4.2.1 discusses variability in RP in general and the different types of variability found within the accent. Section 4.2.2 outlines the changes that can be observed in RP, while 4.2.3 compares RP with so-called Estuary English.

RP is often considered a homogeneous, monolithic accent with standard forms that have one realisation only. This view is partly a result of the fact that RP has been codified and presented in a “frozen” form in pronunciation dictionaries and textbooks, and partly a product of what Milroy and Milroy (1999) call the “standard language ideology”: people have a notion of a correct, canonical type of pronunciation where only one form can be correct, and any deviations are considered “incorrect” or “undesirable”. As RP in many ways functions as a standard variety (cf. 3.3), many people consider the accent invariable and uniform. Spencer noted fifty years ago that “it is implicit in all that has been written about RP that RP speakers represent in some sense a ‘homogeneous speech community’” (Spencer 1957: 22). However, RP is, and always has been, subject to variation and change, as are all living varieties of language. Daniel Jones recognised several types of variability in RP and included a large number of alternative pronunciations in his *Pronouncing Dictionary* from 1917. In *An outline of English phonetics* he writes:

Although those who use RP have much in common in their speech, it must not be thought that RP is absolutely uniform. Quite a number of variations are to be found in it. For instance, the qualities of sounds used in some words vary from speaker to speaker. And in the case of some words two distinct pronunciations must both be considered as belonging to RP. (Jones 1960: 13)

In recent years many scholars have discussed the variability found in RP and the changes that are taking place within the accent. Wells (1982) presents a comprehensive overview of the different types of variability found in RP. Recent changes in RP are discussed in e.g. Gimson (1984), Wells (1994a, 1997a), Nolan (1999) and Upton (2004). These descriptions are, however, not based on quantitative corpus-based investigations. Anne Fabricius is one of the few linguists who have conducted empirical sociolinguistic research into the variability of contemporary RP (see Fabricius 2000, 2002a). She has, however, concentrated on one variable, namely t-glottalling.

4.2.1 Types of variability

We can identify four types of variability in RP: systemic variability, realisational variability, connected speech variability, and lexical-incidental variability (cf. Wells 1982: 285-297, Hughes and Trudgill 1996: 36-37). Systemic variability refers to differences that affect the

phoneme system. Until some decades ago, many RP speakers had an additional phoneme /ɔə/ in words like *court* and *soar* which contrasted with the monophthong /ɔ:/ found in *caught* and *saw*. This phonemic contrast is however not found in modern RP. In current RP, some speakers have an additional long /æ:/ which potentially contrasts with /æ/. The two sounds can be heard in near-minimal pairs such as *bad* [bæ:d] vs. *pad* [pæd] and *glad* [glæ:d] vs. *lad* [læd]. Wells (1982: 289) stresses, however, that many RP speakers have no such contrast, and even those who do may have it only in strongly stressed environments. The only systemic variable that may directly affect the current RP phoneme system is the variation between /ʊə/ and /ɔ:/ in words of the lexical set CURE. The monophthong /ɔ:/ seems to be gradually taking over for the centring diphthong /ʊə/ in words like *sure*, *poor*, *tour*, etc., to the extent that some RP speakers may not have /ʊə/ as part of the phoneme inventory. This phenomenon is referred to as CURE lowering (Wells 1982), and has led several phoneticians to predict the disappearance of /ʊə/ from the RP vowel system (see further 4.4.1).

Realisational variability is found at a lower level of analysis, and refers to varying phonetic realisations of one and the same phoneme.²⁷ In the case of the GOAT vowel /əʊ/, for instance, some speakers vary between a central starting point [əʊ ~ øʊ] and a back rounded starting point [ɒʊ ~ ɔʊ] used before non-prevocalic /l/, whereas other speakers have the same quality in all environments (see 4.4.2). T-voicing (see 4.4.4) and t-glottalling are also examples of realisational variables, where the phoneme /t/ in certain positions may be realised as a voiced tap [ɾ] or a glottal stop [ʔ], in addition to the traditional realisation as an alveolar plosive [t]. The quality of the /u:/ phoneme varies considerably in modern RP, with realisations ranging from a back [u:] to a central or even front [ɯ: ~ ʏ:], with the centralised variant becoming increasingly common in all phonetic environments. Similarly, the final vowel in words like *city*, *happy*, *very*, etc. varies in quality from a traditional close-mid [ɪ] to the increasingly preferred close front [i]. Wells (1982: 257-258) refers to this phenomenon as *Happy Tensing* (see also Lewis 1990).²⁸

Variation in phoneme system and phoneme realisation often represents differences between individual speakers or between groups of speakers (e.g. different generations), and

²⁷ This type of variability does not include accommodatory allophonic alternation, such as a dental articulation of /t/ before /θ/ in e.g. *eight*, or a contracted /t/ before /r/ in e.g. *train*.

²⁸ *Happy Tensing* may alternatively be analysed as a lexical-incidental variable, where the words affected have alternative phonemic structures. Within this analysis [i] is regarded as a realisation of the phoneme /i:/ (cf. Hughes and Trudgill 1996: 43, Cruttenden 2001) or is established as a separate phoneme /i/ that replaces unstressed /ɪ/ word-finally and prevocalically (cf. Wells 2000).

typically reflects on-going changes in the accent. The third type of variability is only found in continuous speech and depends as much on speed of delivery and level of formality as on individual differences. This variability comprises what Nolan and Kerswill (1990) refer to as Connected Speech Processes and define as “a variety of reduction and simplification processes” (296). The most common connected speech processes are assimilation, elision and liaison, which typically affect features of morpheme and word junctures (cf. Cruttenden 2001: 278). Assimilation is a type of coarticulation that mainly affects a phoneme’s place of articulation: e.g. *good boy* /gʊd bɔɪ/ → /gʊb bɔɪ/, *ten cats* /ten kæts/ → /teŋ kæts/, *did you* /dɪd ju/ → /dɪdʒu/. Elision is the omission of a phoneme that would be present in a careful citation form: e.g. elision of /t/ in *next cab* /neks kæb/ or /h/ in *tell him* /tel ɪm/. In some cases a whole syllable is elided, or even several syllables, as when *camera* is pronounced /kæmɹə/, or *literary* becomes /lɪtri/. Wells (2000: 165) refers to this process as Compression. Smoothing (see 4.4.5) is a subtype of Compression, where a prevocalic diphthong loses its second element and is reduced to a monophthong, e.g. *science* /saɪəns/ → [sæns]. The last main type of connected speech process is liaison, which refers to a “transition between sounds, where a sound is introduced at the end of a word if the following syllable has no onset” (Crystal 2003: 269). The only regular type of liaison in English is r-liaison, or r-sandhi, which involves the use of linking and intrusive /r/: *far away* /fɑːr əweɪ/, *media event* /miːdiər ɪvent/ (see further 4.4.3). Connected speech processes are largely governed by rules and principles of articulation, and their outcome is determined by the phonetic surroundings. They are commonly assumed to be much more frequent in casual and fast speech, although there are findings that suggest a weaker correlation with style and pace for some of the variables (cf. Cruttenden 2001: 293-294, Shockey 2003). Except for this possible correlation with stylistic factors, connected speech processes are not believed to vary significantly from speaker to speaker (Wells 1982: 286).

The last type of variability found in RP is the lexical-incidental variation. Unlike the phonetically conditioned processes presented above, which systematically affect a large number of items, lexical-incidental variation involves lexicalised alternation in individual words or classes of words. Many words have two or more alternative pronunciations in RP, with different speakers having different phonemic representations for the same word, sometimes combined with different stress patterns. Among the items that alternate in this way are *again* /əgeɪn, əgeɪn/, *often* /ɒfn, ɒftən/, *zebra* /zebrə, ziːbrə/, *issue* /ɪʃju, ɪʃu/, and *kilometre* /'kɪləmi:tə, kɪ'lɒmɪtə/. In these cases both forms fall within Mainstream RP, but

one alternative may be more usual than the other, or one may be considered more conservative than the other.

There have been a few studies investigating the preferences of native speakers with many of these variable words (see Wells 1999a). The results of these opinion polls are included in Wells (2000). Lexical-incidental variability typically affects individual words in a relatively unsystematic way. There are however cases where alternative phonemic structures are found systematically in large numbers of words, or classes of words. In RP there are several words that fluctuate between /ɑ:/ and /æ/, e.g. *plastic, plaque, transitive, translate*, etc. Similarly, there is considerable fluctuation between /ɪ/ and /ə/ in large groups of words that end in *-ate, -ible, -ily, -ity, -ness*, etc., or that begin with *be-, de-, re-*, etc. (*delicate, sensible, happily, sanity, kindness, believe, depend, repair*, etc.). In all these cases the pronunciation with /ɪ/ is the more conservative (cf. Wells 1982: 296). The alternation between /tj, dj/ and /tʃ, dʒ/ in *Tuesday, duke, reduce*, etc., is among the more recent lexical-incidental variables in RP (cf. 4.4.6).

4.2.2 Changes in RP

Many pronunciations, which thirty or forty years ago were confined to the vulgar, are gradually gaining ground; and if something be not done to stop this growing evil ... the English is likely to become a mere jargon, which every one may pronounce as he pleases.

Thomas Sheridan (*A General Dictionary of the English Language*, 1780)

Linguistic variability is in many cases indicative of linguistic change (cf. 3.1.1). Several of the variable features mentioned above reflect new trends or on-going developments in RP, where new pronunciations replace old ones. Some features have been documented in RP since the days of Daniel Jones, others are recent or have become increasingly common in the course of the past few decades. This section gives a brief survey of the most important changes that have taken place or are taking place in RP, in addition to features that have only recently started to make their way into Mainstream RP. A few changes affect the phoneme system of RP, but the majority are changes at the realisational level.

The changes in RP can be arranged into three categories, grouped according to how well-established they are, i.e. how many RP speakers have embraced them. The first group (A) incorporates changes that are firmly established with most speakers and widely accepted as part of Mainstream RP. Group B includes changes that are still in progress, and where there

is a great deal of variation among current RP speakers. The third group (C) contains the most recent changes and pronunciations that are on the verge of being accepted into RP. These new features indicate the direction of the future development in RP. There is not complete agreement among linguists as to just how well-established some of the changes are, and which ones should be considered part of RP. The lists below are based primarily on descriptions in Wells (1997a, 1994a), Cruttenden (2001), Upton (2004) and Ramsaran (1990b).

A. The following changes are almost complete, in the sense that the “new” pronunciations are now typical of the large majority of Mainstream RP speakers (the old pronunciations are heard with many U-RP speakers):

1. The distinction between /ɔə/ and /ɔ:/ is lost, making *floor* and *flaw* homophones /flɔ:/.
2. Words of the lexical set CLOTH are pronounced with /ɒ/ rather than /ɔ:/, e.g. *off* /ɒf/.
3. The TRAP vowel /æ/ is lowered and realised as an open [a].
4. The GOAT vowel /əʊ/ has changed in quality from [oʊ] to [əʊ].
5. Yod coalescence in unstressed syllables, e.g. *culture* /'kʌltʃə/, *soldier* /'səʊldʒə/.
6. /r/ is realised as a post-alveolar approximant in all positions, whereas formerly a tapped [ɾ] was usual in intervocalic positions.

B. The following changes are well-established with many RP speakers, but are still very much in progress:

1. The SQUARE vowel /eə/ is realised as a monophthong [ɛ:].
2. The CURE vowel /ʊə/ is disappearing and being replaced by /ɔ:/.
3. The PRICE vowel /aɪ/ is realised with a back-centralised starting point: [ʌɪ].
4. The GOOSE vowel /u:/ is fronted to [ʊ:].
5. The DRESS vowel /e/ is lowered to [ɛ].
6. Final /ɪ/ is realised as [i] in e.g. *happy*, *very*.
7. Unstressed /ɪ/ and /ʊ/ are replaced by /ə/, in e.g. *possible* /'pɒsəbl/, *executive* /ɪg'zɛkjətɪv/.
8. Syllable-final /t/ is replaced by a glottal stop [ʔ] in pre-consonantal positions, e.g. *network* [neʔwɜ:k].

C. The following changes are the most recent innovations. They are heard with some RP speakers, but are not (yet) universally accepted as part of Mainstream RP:

1. Yod coalescence in stressed syllables, e.g. *tune* /tʃu:n/, *reduce* /rɪ'dʒu:s/.
2. Glottalling of syllable-final /t/ before a stressed vowel or a pause, e.g. *but Ann did it* [bət̚ 'æn dɪd ɪʔ].
3. Vocalisation of non-prevocalic /l/ to [o ~ ɤ], e.g. *film* [fɪɔm].
4. R-fronting: the realisation of /r/ as an approximant [ʊ] or [ʉ].

As can be seen from this overview, many of the changes affecting the RP accent are phonetic, often involving vowel quality. Most of these changes are therefore not reflected in the majority of pronunciation dictionaries, which only give phonemic representations, using the traditional phoneme symbols. Upton et al. (2001) is to my knowledge the only pronunciation dictionary that employs “new” symbols to reflect many of the changes that have taken place in the RP vowels. They use the symbol [a] to represent the more open TRAP vowel, [ɛ] for DRESS, [ɛ:] as the symbol for SQUARE, and [AI] for the vowel in PRICE.

As regards the social connotations and general acceptability of the features in A-C, those pronunciations that are most firmly established tend to be regarded as the most “correct”, while the innovations (which often come from non-standard accents) tend to be stigmatised. Because of this difference in status, the conservative pronunciation variants will often be used in formal situations, and the less established variants in casual contexts (cf. Cruttenden 2001: 79). The most controversial RP features are then found in groups B and C. The “controversy” refers to the fact that there is a considerable amount of disagreement among linguists regarding the acceptability of these features as part of RP. Bauer (1985), for example, includes the precondition that the RP speakers in his survey should have [ɪ] rather than [i] as the final vowel in *happy*, whereas Wells (1982) recognises [i] as a possibility in RP. Similarly, Wells (1997a) assigns l-vocalisation and stressed yod coalescence to changes affecting RP in the late twentieth century, while Maidment (1994) excludes both features from RP.

Most of the changes that have taken place in RP seem to originate from the London accent, or Estuary English (see 4.2.3). Reports of changes in RP attribute many of the “new” features to the diffusion of London speech (cf. Gimson 1964, Wells 1994a, 1997a, Taylor 1998), and Wells predicts that this tendency is likely to continue (1997a: 25). This is not surprising in view of the fact that the London region has always been the most influential area

and the main source of new fashions, in pronunciation as in all other domains. Features from southeastern accents, and particularly Popular London English (Wells 1982: 302), are spreading to all parts of Britain, entering the accents of the largest English cities first (cf. Foulkes and Docherty 1999b: 11). This is not a new phenomenon, but with the increased mobility and democratisation in modern Britain, the trend seems to be moving faster than ever before (cf. 3.1.3). Wells (1994a) talks of the “Cockneyfication” of RP, and Taylor (1998) claims that linguistic “downgrading” is a growing trend with middle- and upper-middle class speakers. Features such as t-glottalling, yod coalescence, l-vocalisation, *happy* tensing and the new vowel qualities in PRICE and GOOSE are all changes that are associated with the London accent (cf. Wells 1982: 301-331). Many of these features, while maybe originating from Cockney, are however not exclusive to London speech. They are now shared by several other varieties spoken in major cities in Britain, and have entered RP after having become widespread and non-regional (cf. 2.4). Other Cockney features that are entering popular accents throughout Britain include h-dropping ([æm] for *ham*), g-dropping ([draɪvɪn] for *driving*), and TH fronting ([fɪn] for *thin*), but these pronunciations are (as yet) resisted by RP (cf. Wells 1994a, Kerswill 2001).

Although the most considerable influence on RP (and other varieties) is the accent of the London area, there are exceptions to this trend. Some of the changes in RP do not conform to London or southeastern English, and can plausibly be attributed to other sources, or to internal factors. The lowering of the TRAP vowel from [æ] to an open [a] is actually a change towards a quality found in northern accents, and away from Cockney, which has [ɛ ~ æ] in this lexical set. Similarly, there may be a tendency away from the traditional fully back [ɑ:] in BATH towards a centralised quality, especially before nasals (*chance, sample*) (cf. Upton 2004: 223). This is also a feature more associated with northern accents. Upton (2004) further mentions a possible change in the RP STRUT vowel, which is traditionally pronounced as a centralised and relatively open [ʌ]. This vowel is however in danger of being confused with the lowered TRAP vowel [a], and there now seems to be “an increasing appearance ... of an innovation in which [ʌ] is raised and retracted from the centralized, towards (though not to) a half-close ... position” (2004: 223), resulting in a quality somewhere between the northern [u] and the southern [ʌ]. Such a development would also be a move away from Cockney, which has a very open fronted [a ~ ɶ] in STRUT (Wells 1982: 305). Likewise, the decline of unstressed /ɪ/, which is increasingly replaced by /ə/ in endings like *-ace, -ate, -ed, -ily, -ness*, etc., cannot be attributed to influence from Cockney. The source of this change is rather to be

found in provincial England and perhaps also North America, Australia and Ireland (Wells 1994a: 201).

4.2.3 RP versus Estuary English

Many of the recent trends and on-going changes in RP have been characterised as features of Estuary English (EE), and can be seen in parallel with the rise of the EE phenomenon (cf. Wells 1997a). EE is a supposedly new accent which phonetically is said to be situated somewhere between RP and Cockney. It is spoken by the lower middle classes in the Greater London and Home Counties area and is spreading rapidly, both geographically and socially (cf. 3.1.3).

The descriptions of Estuary English are extremely varied and sometimes contradictory. Cruttenden (2001: 81) equates EE with London Regional RP (i.e. near-RP with some localised London features), whereas Davenport and Hannahs (2005: 34) describe it as “working-class and lower middle-class speech in South Eastern England”. Estuary English has on the one hand been hailed as the new standard English that will replace RP (cf. Coggle 1993), and on the other been dismissed as hype and unfounded claims (cf. Maidment 1994). Wells (1998b) defines EE as “standard English spoken with an accent that includes features localisable in the southeast of England”, a very broad definition, which, while distancing the accent from non-standard Cockney, covers a wide range of pronunciations. A similar view is adopted by Przedlacka (2002), who, in her study of EE, concludes that the variety covers a number of distinct accents in the Home Counties area, and is “not a single and definable variety” (2002: 97). Kerswill (1994) regards EE as part of a general trend of accent levelling in the southeast of England (cf. 3.1.3). Estuary-like features are, however, spreading beyond the southeast, to other parts of the country (along with non-Estuary features like TH-fronting and h-dropping), which, according to Crystal (2003: 166) indicates that we are dealing with broader issues of language change.

While linguists disagree as to whether or not EE is a separate variety, most of them agree that it is not new. The phenomenon of EE is a continuation of a trend that goes back several centuries, namely the tendency for features of popular London speech to spread geographically and socially. With the shrinking class differences, increased social mobility and new patterns of settlement in modern Britain, this trend is more noticeable today. During the last 50 years large numbers of Londoners have moved out of the capital, bringing their

accent with them, and providing a strong London presence all over the southeast of England, including a strong linguistic influence (cf. Coggle 1993: 24). The increasing dominance of EE in the media may also contribute to the popularity of the accent (cf. Crystal 1995: 327). According to Coggle (1993) EE is now spoken across a very wide social spectrum, and appeals to a large number of people because it is perceived as modern, informal, socially neutral, approachable and high on “street-cred” (85).

The phonetic features generally referred to as characteristic of EE include t-glottalling, l-vocalisation, *happy* tensing, yod coalescence, GOAT allophony, diphthong shift in FACE, PRICE, GOAT and MOTH (towards [æɪ, aɪ, ʌʊ, æʊ]), and r-fronting (cf. Wells 1994b, 1997b, 2000: xiii, Cruttenden 2001: 88, Coggle 1993). Many of these features are also among those said to be entering modern RP (cf. 4.2.2). Broader Cockney features, like TH-fronting, h-dropping and monophthongisation of /aʊ/ (giving [mæ:f] for *mouth*), are said to fall outside EE (cf. Cruttenden 2001). There is thus a significant phonetic overlap between EE and modern or informal RP, and an accent that by some is referred to as EE, may by others be described as modern RP.²⁹ The claims that EE will replace “fuddy-duddy” old RP and become the new standard accent and model for general imitation (cf. Rosewarne 1994, Coggle 1993) are therefore partly justified. Several of the phonetic characteristics of EE are among those features (originating in London) that are currently spreading throughout the whole of Britain and losing their localisability. If RP is defined as a non-localisable accent (cf. 2.4), EE represents a good indication of where RP is heading. Wells is quite explicit about the direct link between the two, and states that many of the EE features “are being, or are likely to be, gradually incorporated into RP” (1997a: 25) and that the descriptions of RP should be updated by “gradually incorporating one or two of the changes typical of EE” (1997b). Wells is, however, somewhat inconsistent in his descriptions of EE and RP, in that features like t-glottalling, l-vocalisation, GOAT allophony and yod coalescence are sometimes characterised as EE features that fall outside RP (e.g. Wells 1994b, 1998b, 2000), and sometimes included as part of modern RP (e.g. Wells 1994a, 1997a).

Nolan (1999), on the other hand, cautions against assuming a direct and one-sided influence from EE on RP. He argues that there may be common internal pressures that promote similar changes in these (and other) varieties, and that if there is influence, it is likely to go both ways.

²⁹ Even Tony Blair’s accent has been characterised as Estuary English (cf. Rosewarne 1998, White 1999).

4.3 Choice of variables

The present study of RP investigates six phonological variables:

7. CURE lowering: /ʊə/ > /ɔ:/ in *poor, sure, pure*, etc.
8. GOAT allophony: [əʊ] > [ɒʊ] before [t] in *goal, cold*, etc.
9. R-sandhi: /r/ vs. Ø in *far away, law and order*, etc.
10. T-voicing: [t] > [t̥] in *city, little, a lot of*, etc.
11. Smoothing: [aɪə, aʊə] vs. [aə] in *fire, power*, etc.
12. Yod coalescence: /tj, dj/ > /tʃ, dʒ/ in *tune, duke*, etc.

The variables represent systemic, realisational, and incidental variation, as well as connected speech processes. Variables 1, 2 and 5 concern vowel features while 3, 4 and 6 are consonantal variables. Two of the variables, r-sandhi and smoothing, are familiar as variable features of RP, and have long been present in the accent. The others represent on-going changes or recent trends in RP. CURE lowering is a change in slow progress that has been going on for at least a century. Yod coalescence is a relatively new development that is beginning to be accepted as part of current RP. GOAT allophony and t-voicing are generally not included in traditional descriptions of RP (e.g. Wells 1982, Cruttenden 2001) or in recent accounts of ongoing trends in the accent (e.g. Upton 2004, Wells 1994a). Observations suggest, however, that both these features are (increasingly) common in speech that can be identified as RP.

Three of the variables, smoothing, r-sandhi and yod coalescence, each incorporates two subvariables: (aɪə) and (aʊə), linking and intrusive (r), (tj) and (dj) respectively. For each variable, two variants have been identified, representing the conservative versus the “new” pronunciation, or presence versus absence in the case of r-sandhi. A systematic investigation of the six variables will provide insight into the relative frequency of the different variants and their correlation with linguistic and non-linguistic factors. The analysis may further reveal any significant covariation of variables that represent related phenomena.

There has been very little or no corpus-based research on variables 1-6 with data from RP speakers, at least in recent years. The variables have therefore been chosen as the present object of study partly to fill an empirical gap in the field of accent studies. The selection of

variables is further based on the fact that they represent areas where RP exhibits considerable variability. This knowledge is derived partly from previous descriptions of RP (notably Wells 1982) and partly from my own observations. Moreover, the variables represent a wide range of phonetic and phonological variation, and thus cover a large part of the variability currently found in RP.

Most of the variables are expected to correlate with linguistic as well as stylistic and social factors. The hypotheses about the indexical values of the variables and likely directions of change are based primarily on existing descriptions of RP, of which there are many. The majority of these are reference works and textbooks on phonetics, as well as pronunciation dictionaries. These will serve as the main basis of comparison and point of reference in the description and analysis of variables. The six variables are, however, not unique to RP, but can be found in several other accents of English, some of which may have directly influenced the RP accent. It is therefore at times relevant to refer to descriptions and studies of varieties other than RP.

4.4 The variables studied

The present section gives a more detailed description of the phonological variables that constitute the focus of study in the present investigation. The presentations include a phonetic description of the variables and their respective variants, an overview of relevant previous research, a discussion of the variables' linguistic and social correlations and expected patterns of variation, as well as a discussion of whether the variables represent on-going changes in RP.

4.4.1 Cure lowering

Words of the lexical set CURE are in RP traditionally pronounced with /ʊə/, a centring diphthong with a back-central close(-mid) starting point. There is however an increasing tendency among RP speakers to use the back mid rounded monophthong /ɔ:/ of FORCE/NORTH/THOUGHT in these words, thus creating homophonous pairs like *sure – shore*, *poor – paw*. Wells (1982: 237) refers to this development as CURE Lowering, as the /ʊə/ of

CURE is lowered and monophthongised to [ɔ:], i.e. /ɔ:/. The term *lowering* may be taken to imply that the /ʊə/ phoneme is undergoing a qualitative change and gradually becoming phonetically identical with /ɔ:/. This does not seem to be the case, however; with those speakers who have both phonemes, the two sounds remain phonetically distinct. It is perhaps more correct to say that /ʊə/ is being replaced by the “lower” monophthong /ɔ:/ in an increasing number of words, thereby losing some of its phonemic ground. Previous explanations of CURE lowering, however, seem to suggest a phonetically gradual change whereby /ʊə/ and /ɔ:/ approach in quality. Gimson (1964) relates the increasing replacement of /ʊə/ with /ɔ:/ to a qualitative change in the /ɔ:/ phoneme:

The quality of /ɔ:/ appears, under London influence, to be considerably closer, [ɔ:], than at the beginning of the century. This closing of /ɔ:/ brings it into increasing qualitative proximity to /ʊə/, so that the levelling of earlier /ʊə/, /ɔə/ and /ɔ:/ is now common. Moreover, /ʊ/ is now in closer qualitative proximity to /ɔ:/ = [ɔ:] than to /u:/ = [u]. (Gimson 1964: 135)

Other researchers seem, like Wells, to assign CURE lowering to an inherent change in the phonetic quality of /ʊə/. Cruttenden (2001) holds that the lowering of the diphthong is directly relatable to the merger of earlier /ɔə/ with /ɔ:/ as “the first element of /ʊə/ can be lowered considerably without risk of confusion” (134). Ball (1984: 40) suggests a similar phonetic answer to the disappearance of /ʊə/:

Diphthongs, being long, will tend to monophthongize to long vowels. ... Phonetically, the first element of /ʊə/ is opener than /u:/, and the glide to [ə] takes it even further away. With the absence of a stable /ɔə/ ... it again appears straightforward for the diphthong to lower, and then monophthongize to /ɔ:/.

What all these explanations seem to miss, however, is the point that CURE lowering is a phonetically abrupt process, and that /ʊə/ for most speakers remains as a separate sound, qualitatively distinct from /ɔ:/.³⁰ What appears to be happening is a lexically gradual change, whereby an increasing number of CURE words transfer to the FORCE set. The increasing

³⁰ The first element of /ʊə/ is often realised with a central, rather than back, quality, especially after /j/. This can be seen in relation to the general fronting of /u:/ and /ʊ/ (cf. 4.2.2). The tendency to centralise after /j/ also applies to /u:/ (cf. Wells 1982: 148).

convergence of /ʊə/ and /ɔ:/ thus seems to be a case of what Labov refers to as *merger by transfer*: “a unidirectional process in which words are transferred gradually from one phonemic category to another” and which, as a rule, “is not consistent with a result that shows an intermediate phonetic form” (Labov 1994: 321).

CURE lowering may be regarded as part of a more general trend, whereby RP centring diphthongs are being replaced by monophthongs. The centring diphthongs have a low functional load and seem to be particularly vulnerable and prone to monophthongisation. /ɔə/ has already disappeared from the phoneme system of RP, having merged with /ɔ:/, and /eə/ of SQUARE is increasingly being substituted by a monophthong [ɛ:] (cf. 4.2.2). With the lowering of /æ/ towards [a], there is room in the front area for a new monophthong, which secures phonological status for the SQUARE vowel. In the back area, on the other hand, there is less room for new vowels,³¹ and /ʊə/ of CURE can then either transfer to /u:/ or to /ɔ:/, alternatively to the central /ɜ:/.³² When /ʊə/ is moving to /ɔ:/ and not to the other phonemes, this can be explained by reference to Element Phonology (cf. Harris 1994). According to this theory, the monophthongisation of a diphthong involves a fusion of the elements within the vocalic compound, so that some properties of both the vowel components of the diphthong will be preserved. A transfer of /ʊə/ to /u:/ would then entail keeping the element of [ʊ] but nothing of [ə], and a lowering to /ɜ:/ would preserve the [ə] quality, but not [ʊ]. In order to retain the backness and roundedness of [ʊ] as well as the relative openness of [ə], the closest monophthongal alternative is /ɔ:/.

The phenomenon of CURE lowering is discussed in many writings on RP, and phoneticians have noted that the merging of /ʊə/ and /ɔ:/ is on the increase (cf. e.g. Wells 1982: 287). Gimson (1964: 133) refers to the increasing coalescence of the two sounds as one of the most notable changes in the traditional RP vowel system. CURE lowering is, however, not a new development. The variation /ʊə/ ~ /ɔ:/ in CURE words is recorded by Daniel Jones in the first edition of his *Pronouncing Dictionary* (Jones 1917). He gives [ʊə] as the main

³¹ The number of distinctions that can be maintained along any one phonetic dimension is limited. For vowel height, it is argued that three levels are the most that can be maintained in the back, and four in the front (cf. Labov 1994: 328).

³² There are a few CURE words that have an alternative pronunciation with /ɜ:/, such as *during*, *jury*, *lure*, *lurid* (cf. Wells 2000).

variant, and [ɔ:] as an optional, though less frequent, pronunciation (along with [oə] and [ɔə]) in CURE.³³

Despite the decline of /ʊə/ and predictions about its disappearance, the sound still remains a part of the phoneme system of RP, as most RP speakers will usually have at least some words with /ʊə/ (cf. Wells 1982: 287). Ramsaran (1990b: 181) argues that “it seems not unreasonable to include /ʊə/ as having phonemic status”, on the grounds that the diphthong remains contrastive in at least a few minimal pairs, such as *lure* /lʊə/ – *law* /lɔ:/, *cruel* /krʊəl/ – *crawl* /krɔ:l/. Both lexical specification and the phonetic environment seem to play a part in the variability of CURE lowering. The replacement of /ʊə/ by /ɔ:/ is especially prevalent in common and monosyllabic words such as *poor*, *sure*, *moor*, *tour*, *your*, etc., while the diphthong seems to survive in rare and exotic words like *gourd*, *dour*, *Ruhr* (cf. Ramsaran 1990b: 181, Gimson 1984: 49). Furthermore, according to Wells (1982), /ʊə/ tends to be retained in words where the vowel is preceded by a consonant plus /j/, e.g. *cure*, *pure*, *furious*, *bureau*, etc. Where the /j/ is word-initial, however, as in *your*, *Europe*, /ɔ:/ is frequently heard. This means, Wells argues, that “there is a certain tendency to make [ʊə] a positional allophone of /ɔ:/, restricted to the environment Cj__” (Wells 1982: 287-288).

It would seem, from the observations referred to above, that the use of /ʊə/ vs. /ɔ:/ is largely a case of lexical variability. However, most phoneticians consider it likely that the phoneme inventory of RP will be affected as a result of CURE lowering, and that /ɔ:/ will completely take over the lexical distribution of /ʊə/. Gimson (1984: 49) claims that the /ʊə/ phoneme is “very much in danger” of disappearing, in the same way that /ɔə/ has been lost, and states that “the frequency of occurrence and functional load of /ʊə/ [is] extremely low”. According to Brown (1990: 38) the phoneme “is rapidly disappearing in the speech of younger RP speakers as it merges with /ɔ:/”. She even goes as far as saying that “/ʊə/ for all practical purposes does not exist in the speech of many younger RP speakers”. Bauer (1994: 126) and Nolan (1999: 79) both predict a phonological change in current RP with the transfer of the CURE set from /ʊə/ to /ɔ:/, and Wells states that the phonemic status of [ʊə] is “beginning to become dubious” (1982: 288).

As CURE lowering is a sound change in progress, a certain correlation with age is to be expected. In an opinion poll survey (referred to in Wells 1990b) Wells found that for *poor*, the

³³ In an analysis of an excerpt from a BBC news bulletin from 1936, Leitner records [ɔ^ə] in the word *Europe*, a pronunciation which deviates from the contemporary recommendations of the BBC’s Advisory Committee (Leitner 1980: 95-96)

pronunciation /pɔ:/ was preferred over /pʊə/ by the youngest respondents by a ratio of 4:1, whereas for the oldest group the ratio was 1:3. In a similar investigation, using ten young students as informants, Ball found that all subjects used /ɔ:/ in *poor*, and seven of the ten had a monophthong in *tour*. For *cure*, however, only one speaker used /ɔ:/, and one had [ɔə], while the rest used /ʊə/ (Ball 1984: 42). A more recent poll preference survey (see Wells 1999a, 2000) shows the same systematic difference between old and young speakers in the use of /ʊə/ versus /ɔ:/ in words like *lure*, *poor*, *sure* and *yours*, and a similar age grading is observed in Tollfree (1999: 169) The trend towards CURE lowering is further reported to be more advanced in local urban accents (Wells 1984: 67)³⁴ – a pattern which is typical of linguistic changes in progress.

The observations of phonological variation in CURE words suggest that this variability is highly complex. It is partly optional: the lowering sometimes applies, sometimes not, often randomly; many words fluctuate between the two vowels, often with one and the same speaker (cf. Wells 1982: 65). It is partly socially determined, as the choice of phoneme seems to correlate with the age of the speaker. The variation can also partly be described in terms of lexical specification, as some words are more likely to undergo lowering than others. Lastly, the variation partly depends on the phonetic environment, as the lowering is reportedly less likely to occur after a consonant + /j/.

As the centring diphthong /ʊə/ is the most conservative variant in CURE words, it would seem reasonable to assume that this pronunciation is the most prestigious,³⁵ compared to /ɔ:/, as innovations tend to be more stigmatised. With regard to the CURE variable, however, this is not necessarily the case, as the two variants [ʊə] and [ɔ:] seem to be fairly equal in status. This observation is supported by Tollfree's findings for CURE in South East London English, which revealed that social class was not an influencing factor in the distribution of /ʊə/ vs. /ɔ:/ (Tollfree 1999: 169). Ball (1984) notes a slight gender difference, in that females use /ʊə/ more frequently than the males, which seems to confirm the conservative status of /ʊə/. This result is however based on a very small sample of ten informants reading four CURE words, and thus not very conclusive. Ball also states that "it is open to argument as to which of [ʊə] or [ɔ:] is the most prestigious" (1984: 43). The apparent equality in the social status of the two vowels may be linked to the fact that neither of the two variants represents a new

³⁴ In their investigation of the Derby accent, Docherty and Foulkes (1999: 50) report that the realisation of CURE is "overwhelmingly [jɔ:]".

³⁵ Prestige is here understood in terms of 'Superiority' (cf. 3.2.1).

sound in RP: /ʊə/ and /ɔ:/ are both well-established phonemes in RP. Moreover, the process of CURE lowering has been going on for such a long time that both vowels are now common in the lexical set CURE.

4.4.2 GOAT allophony

GOAT allophony refers to an allophonic variation in the phonetic quality of the GOAT vowel. In RP, words of the lexical set GOAT (*goat, soap, hold, home, know*, etc.) are traditionally (at least from the early 20th century) pronounced with a back-closing diphthong with a mid-central unrounded starting point: [əʊ] (Gimson 1962: 128, Wells 1982: 146, Cruttenden 2001: 135). Increasingly, however, many speakers have a different vowel quality, with a back rounded starting point, [ɒʊ], before non-prevocalic /l/, in words like *roll, cold, shoulder*, etc. For these speakers [ɒʊ] seems to be a contextual allophone of /əʊ/ found in the environment before dark /l/: *home* [həʊm] but *hole* [hɒʊtʃ], *goat* [gəʊt] but *goal* [gɒʊtʃ]. This phenomenon has been referred to as *GOAT allophony* (cf. Maidment 1994).

The back rounded [ɒʊ] variant in pre-/l/ GOAT items is common in London English, where the realisational variability has resulted in a phoneme split. What was once a matter of allophonic variation, with [ɒʊ] before non-prevocalic /l/ and [ʌʊ] in all other environments, has ended in a vocalic split and an additional phoneme /ɒʊ/ used in the subset *GOAL*. Through a process of morphological regularisation London speakers have extended the [ɒʊ] variant to words where the /l/ is morpheme-final but prevocalic, yielding minimal pairs such as *wholly* [hɒʊli] vs. *holy* [hʌʊli]. L-vocalisation further leads to more minimal pairs such as *soul* [sɒʊ] vs. *so* [sʌʊ], *bowl* [bɒʊ] vs. *bow* [bʌʊ] (cf. Wells 1982: 312-313). The split in the GOAT set is “well-established in all kinds of London-flavoured accent”, according to Wells (1982: 313). Cruttenden mentions GOAT allophony as a feature found in Cockney as well as in London Regional RP (2001: 136), and it has also been described as typical of Estuary English (cf. Wells 1994b, Maidment 1994). Tollfree (1999) concludes from her empirical survey that the GOAT split “appears to be categorical in South East London for all speakers (including near-RP speakers)” (167).

GOAT allophony (or split) is however not confined to the London vernacular. There is evidence of at least variable allophony in many regional urban accents in England (see Foulkes and Docherty 1999a). Wells (1984: 64) states that *goat-goal* variation is found in

many parts of the country, but that it is only in London that the *goal* vowel is phonemically distinct from *goat*. GOAT allophony is generally not recognised as a feature of RP (cf. Maidment 1994), and there have been no previous studies of this vowel variable with speakers of non-regional RP. Wells (1997a) is the only description that mentions GOAT allophony in connection with RP:

Increasingly in RP words such as *fold*, *goal* are said with a back rounded diphthong with a starting point comparable to the [ɒ] of *lot*. This diphthong is found only before dark /l/, [ɫ], or the vowel that develops from it. (Wells 1997a: 25)

In LPD (Wells 2000) the use of a back rounded variant [ɒʊ] in words like *cold* is included as a possibility in the relevant dictionary entries, although it is described as a feature of near-RP (2000: xvi, xx). In other sources, however, Wells characterises GOAT allophony as a feature that falls outside RP. In Wells (1982) he states that “RP characteristically lacks the pre-/l/ allophone [ɒʊ] of many other accents” (237). In Wells (1994b) GOAT allophony is mentioned as one of the areas where Estuary English agrees with Cockney but differs from RP. Moreover, GOAT allophony is not included in the discussions of recent changes and new trends in RP in e.g. Wells (1994a) and (1998a). Linguists other than Wells on the whole do not mention GOAT allophony as an option in RP, and dictionaries like ODP (2001) and EPD (2003) do not include [ɒʊ] as a variant at all.

My own observations of RP in the media suggest that [ɒʊ] before non-prevocalic /l/ is a common feature among speakers of modern non-regional RP. In comparison with the standard descriptive works, the presence of GOAT allophony in RP undoubtedly indicates an ongoing change in the accent. It seems reasonable to infer that GOAT allophony is among the features originating from the capital, considering its prevalence in London and the surrounding areas. GOAT allophony is however not associated with any social stigmatisation, unlike many other London variants. The feature appears to be evaluatively fairly neutral, which probably stems from the fact that it is used by speakers of both higher and lower classes, from the “maximally to minimally broad” varieties (Tollfree 1999: 164).

In the analysis of GOAT allophony among the speakers in the present study, I will investigate to what extent there is a systematic qualitative contrast in the realisation of the GOAT vowel in words with and without a following dark /l/, and whether there is evidence of a phoneme split.

4.4.3 R-sandhi

One of the cardinal sins of utterance is the insertion of an R where none exists in the spelling. All speakers are ... guilty of this at one time or another. ... It is not easy to avoid using it ... in conversational utterance, but the idea of it elsewhere must not be entertained.

J. Clifford Turner (*Voice and Speech in the Theatre*, 1950)

R-sandhi comprises the phenomena of linking and intrusive /r/, which involve the insertion of /r/ in cases of vowel hiatus, e.g. *for/r/ ever*, *China/r/ and Japan*. RP speakers are said to make extensive use of linking and intrusive /r/ in the appropriate places to help words and syllables run together more smoothly. Wells (1982) and others refer to this process as *r sandhi*.³⁶ An alternative name for the same phenomenon is *r liaison*, which is used by e.g. Cruttenden (2001) and Giegerich (1992).

In RP, as in other non-rhotic accents, /r/ only occurs prevocally,³⁷ and a word said in isolation never ends in /r/. In connected speech, however, a word-final /r/ may be pronounced if the next word begins with a vowel sound. When the spelling includes an *r*, as in *better off*, *far away*, the inserted /r/ is known as *linking /r/*. In such cases, the r-sound corresponds to an historical /r/, still reflected in the orthography. By analogy, many RP speakers link adjacent vowels with an /r/ even when there is no *r* in the spelling, e.g. *Anna/r/ and I*, *law/r/ and order*. When a phonetic /r/ occurs in such unhistorical environments, it is referred to as *intrusive /r/*, as the /r/ “intrudes” into contexts where etymologically it does not belong. Intrusive /r/ is most often heard across word boundaries, as in the examples above. It can, however, also operate word-medially before a suffix beginning with a vowel, e.g. *withdraw/r/al*, *gnaw/r/ing*. The effect of r-sandhi is to avoid a vowel hiatus and maintain an articulatory flow.

Sandhi /r/ occurs only after open or mid vowels and before a word or a syllable beginning with a vowel sound.³⁸ Linking /r/ has the widest distribution and is found after words or syllables that end in one of the vowels /ə, ɜ:, ɑ:, ɔ:, ɪə, eə, uə/. Intrusive /r/ has a more restrictive distribution than linking /r/ simply because there are much fewer words with these endings that do not contain an *r* in the spelling and thus are potential candidates for /r/ intrusion. The most common word-final vowel preceding intrusive /r/ is no doubt /ə/. There is also a considerable number of words that end in /ɔ:/. Examples with /ɑ:/ are few, but include

³⁶ The term *sandhi* comes from a Sanskrit word meaning ‘joining’ (Crystal 2003: 405).

³⁷ /r/ can also potentially occur before a syllabic consonant in e.g. *barrel* [bær̩].

³⁸ Sandhi /r/ is also possible after a consonant in cases where the /r/ is syllabic, in e.g. *better off* [bet̩r̩ ɒf].

e.g. *spa, bra, cinema*. Intrusive /r/ potentially occurs after /ɪə/ only in the word *idea*. The remaining vowels do not occur word-finally except in a negligible amount of rare or foreign words (e.g. *milieu* /mi:lʒɜ:/). Intrusive /r/ is thus primarily found after /ə, ɔ:/ and to a certain extent /ɑ:/.

As mentioned above, the only present-day distinction between linking /r/ and intrusive /r/ is the occurrence of an *r* in the spelling. Phonetically, they can be said to represent the same phenomenon. Wells (1994a: 202) claims that the unetymological intrusive /r/ represents the generalisation of the linking /r/ principle to cases which are phonetically analogous but historically and orthographically different. He states that “intrusive /r/ arises essentially from the natural tendency to give identical treatment to words with identical endings” (Wells 1982: 223). Similarly, Foster (1968: 251) finds that the motive which obliges speakers to use intrusive /r/ is analogy with linking /r/: since *sore* /sɔ:/ has a prevocalic variant /sɔ:r/, it is reasonable to give *saw* a parallel prevocalic variant with /r/.

The use of r-sandhi has commonly been explained either as a rule of deletion or as a rule of insertion. Historically, RP (and other accents) became non-rhotic through a sound change deleting prepausal and preconsonantal /r/ (cf. Wells 1982: 218-222). Deletion theories thus assume that the historical /r/ is present in the speakers’ underlying lexical forms, and that the /r/ surfaces in prevocalic environments but is deleted before a consonant or a pause. The insertion theory, on the other hand, proposes that /r/ is inserted after a word-final non-high vowel when another vowel follows (see further McMahon et al. 1994). The insertion model is the one that best serves to explain intrusive /r/ and therefore seems to be preferred by most linguists (e.g. Johansson 1973, Wells 1982, Brown 1988, Foulkes 1997). The view that r-sandhi results from an insertion rule is supported by the fact that many non-rhotic English speakers insert /r/ in foreign words and expressions (e.g. *viva/r/ España*). Wells argues that since “these instances can only reflect an insertion rule, it is reasonable to conclude that all sandhi /r/ in contemporary RP and other non-rhotic accents reflects the same insertion rule” (1982: 226).

The use of r-sandhi is optional, and hence variable, in non-rhotic speech. It is said to be very prevalent with native RP speakers, and has been described as one of the most characteristic features of connected speech found in the accent (Wells 1982: 286). Linking /r/ is however generally thought to be more frequent than intrusive /r/. Daniel Jones discusses linking and intrusive /r/ in the introduction to his dictionary from 1917 and comments that the latter is considerably less common than the former (Jones 1917: xvii). In the dictionary itself

potential linking /r/ is marked with an asterisk after the relevant entry words, while potential intrusive /r/s are not indicated. This notational convention is used also in the latest edition of EPD (Jones 2003). According to Wells (1985), however, “in real life” intrusive /r/ is “just as likely to occur” as linking /r/, and the EPD notation therefore does not adequately reflect the phonetic reality (1985: 48).

Even if linking and intrusive /r/ may both be very common with native speakers, they do not enjoy the same prestige. The use of linking /r/ is generally considered correct and desirable in Mainstream RP, whereas stigmatisation has arisen for intrusive /r/, mainly on the grounds that there is nothing in the spelling to justify its use. Crystal (1992: 119) quotes from a letter addressed to his radio series *English Now*, which illustrates the widespread negative attitudes towards intrusive /r/:

On Thursday’s Today programme, I was surprised to hear the sloppy pronunciation “lawr and order” from one of your presenters. This carelessness seems to be on the increase, and I hope I am not the only one who is prepared to stand up and be counted. This kind of speech is typical of the lazy attitude found today in so many parts of society, and we should resist it at all costs...

In spite of the fact that it is commonly disapproved of and resisted by the speech conscious, intrusive /r/ “is undoubtedly widespread” (Roach 1983: 110) and “very prevalent in RP” (Wells 1994a: 202). Its long tradition and presence in RP, even among newsreaders, can be seen from the following statement, made in 1932 by Arthur Lloyd James, who was at the time a linguistic advisor to the BBC: “[intrusive /r/ is] always with us, safely entrenched, and apparently a firmly established feature of so-called Standard English” (quoted in Lewis 1975: 41). Wells (1982: 223) also mentions newsreaders on the BBC who “very generally referred to the former *Shah/r/ of Iran*”. Nevertheless, intrusive /r/ is disliked and condemned by many purists, even by speakers who may themselves use it. Thus Burchfield in his pronunciation guide to BBC presenters includes a clear recommendation about /r/ intrusion: “In the formal presentation of the news or of other scripted speech ... avoid the intrusive *r*” (Burchfield 1981: 10). He adds, however, that presenters should “be careful to make the correct liaisons, as in *a pair of*”. Linking /r/ in the appropriate circumstances is thus recommended. Lewis (1975: 42) holds that, given the appropriate prosody, “failure to make r-link is now abnormal”.

Given the difference in status between linking and intrusive /r/, one would perhaps expect the two variables to show different sociolinguistic patterning. Previous accounts have, however, generally revealed no systematic correlations with social or stylistic factors. Wells (1997a: 21) states that “speakers of all social classes in almost all parts of England” use intrusive /r/. Various studies of regional accents have found no significant sociolinguistic patterning for r-sandhi (e.g. Stoddart et al 1999, Mathisen 1999, Tollfree 1999, Trudgill 1974: 162). The only exception is Foulkes (1997), who found evidence of social and stylistic variation with speakers from Newcastle. His investigation revealed that linking /r/ was most common among older and middle-class speakers, while intrusive /r/ was mainly found in the lower socio-economic group. More surprisingly, though, the use of intrusive /r/ increased markedly among middle-class speakers in the formal reading style, indicating that intrusive /r/ is not perceived as stigmatised by these speakers. Foulkes explains this by reference to the fact that intrusive /r/ is not a salient feature of the local Newcastle dialect, and is thus seen as a non-regional feature (often heard in the media from newscasters and announcers) and perceived as prestigious and advantageous by the middle class in Newcastle (Foulkes 1997: 83-84). Descriptions of r-sandhi in RP are numerous, but empirical corpus-based studies are lacking. Bauer (1984) is one of the few systematic empirical studies of linking and intrusive /r/ in RP. Bauer investigated 37 recordings of RP speech made between 1949 and 1966, and found no evidence of the variation being linked to social factors such as age or gender. His findings are, however, based on a very limited sample where every informant was recorded reading the same short text passage.

Since r-sandhi is primarily a boundary feature of connected speech, it might be expected to increase in frequency with the speed of delivery and the degree of informality. Brown (1988) holds that r-sandhi “is a feature of fluent colloquial style, and is not so common in careful declarative style” (145). Ramsaran (1978), on the other hand, observes that linking /r/ is “frequent” in all speech styles (quoted in Wells 1982: 286), and according to Cruttenden, only intrusive /r/ is potentially affected by style, while the use of linking /r/ “is of no stylistic importance” (2001: 294). Similarly, Bauer (1984) found no correlation with speech tempo and the use of sandhi /r/ with his informants. Lewis (1975) states that r-sandhi is normal and very common in “average tempo” speech, but that absence of linking and intrusive /r/ is not uncommon in markedly rapid speech as well as in markedly slow speech. Wells (1982: 284-286) holds that extensive use of r-sandhi is primarily a feature of fluent native RP, while the omission of sandhi /r/s is characteristic of speakers of adoptive RP (cf. 2.2), who typically lack control over connected speech variants. Intrusive /r/ in particular is often avoided by

speech-conscious adoptive RP speakers, who tend to regard it as incorrect or slovenly. The stigmatisation and avoidance of intrusive /r/ in turn leads many speakers to also suppress “justifiable” linking /r/s (Wells 1982: 224, Cruttenden 2001: 289, Brown 1988: 145).

The variation in r-sandhi further seems to correlate with a number of linguistic factors. Most writers agree that the frequency of intrusive /r/ depends partly on the quality of the preceding vowel, in that r-intrusion is less likely to occur after /ɑ:, ɔ:/ than after /ə/ (cf. Jones 1960: 197-198, Cruttenden 2001: 288, Wells 1994a: 202, Lewis 1975: 40). The reason for this difference is that intrusive /r/ after /ɑ:, ɔ:/ tends to be more strongly stigmatised, possibly because the words ending in /ɑ:, ɔ:/ are much rarer than those ending in /ə/, or because intrusive /r/ in these environments represents a more recent development. Crystal (1984: 42) suggests that r-intrusion after /ɔ:/ is more noticeable because of the openness of the /ɔ:/ vowel. Brown (1988: 149) explains the difference with reference to the fact that /ə/ only occurs in unstressed syllables, while /ɑ:/ and /ɔ:/ are usually stressed. Unstressed syllables are by definition less prominent than stressed ones, and for this reason intrusive /r/s after /ə/ may be less noticed. Most stigmatised of all are the word-internal cases where a derivative suffix receives intrusive /r/, e.g. *draw/r/ing*. Foulkes (1997), however, found that the quality of the preceding vowel did not affect the presence or absence of intrusive /r/ with the speakers in his corpus.

Another phonological factor which may impede the use of sandhi /r/ is the presence of another /r/ in the preceding syllable (cf. Jones 1960: 197, Wells 1982: 224, Brown 1988: 145). Linking and intrusive /r/ would thus be avoided after words like *error, nearer, roar, straw, era*, etc., presumably because of the articulatory effort involved. According to Lewis (1975: 38), however, r-sandhi is only regularly omitted in cases where the link syllable is unstressed and begins with a consonant cluster ending in /r/, as in the words *extra, zebra, orchestra, mantra*, etc.

Prosody and syntax are also assumed to influence the use of r-sandhi. Foulkes (1997) notes that linking /r/ is more likely to be omitted before a stressed syllable (e.g. *act your 'age*). Similarly, Bauer (1984) found that linking /r/ occurs most frequently when followed by an indefinite article or an unstressed pronoun. Jones (1960: 197, 1967: xxvi) states that linking /r/ tends to be avoided before a clause boundary (which normally corresponds to a tone group

boundary³⁹) and potential pause. He holds that the insertion of /r/ is “unusual if a pause is possible between the words, even if no such pause is actually made”. He gives as example the phrase *He opened the door and walked in*. This prediction seems to be supported by Bauer’s (1984) findings. Lewis (1975: 37), on the other hand, claims that Jones’ comment “is not applicable to current General British pronunciation”, and that in such cases as the example above “the r-link is generally made in non-hesitant, not particularly deliberate styles of delivery”. Even if the presence of a syntactic boundary may not prevent the use of sandhi /r/, linking and intrusive /r/ are predictably more common between words where the syntactic linkage is close, since the effect of r-sandhi is to join together two neighbouring items.

Finally, r-sandhi is expected to have a certain correlation with lexical factors, to the extent that it is often avoided before proper names. Lewis (1975: 38) reports that linking /r/ tends to be suppressed after honorifics such as *Mr, Sir, Doctor*, etc. when a name is introduced, to avoid ambiguities and misunderstandings (*Mister Arafat*, pronounced with a linking /r/, may sound like *Mister Rarafat*). The same suppression of /r/ is expected to be found with the intrusive variant (in e.g. *Mullah Omar*).

R-sandhi is a variable feature that has been present in RP for a long time. Intrusive /r/ is possibly the more recent variant: Wells (1982: 227) estimates that “it has probably characterized RP since the early nineteenth century, though no doubt regularly disapproved of and avoided by the speech-conscious”. The variability in r-sandhi does then not represent an ongoing change in the phonology of RP. There are, however, some observations that suggest that there may be a shift in the frequency of the two r-sandhi types. Jones comments as early as 1960 on a decline in the use of linking /r/, as he notes that “there appears to be an increasing tendency, especially among younger people, not to use linking r at all” (Jones 1960: 197). In his investigation of Newcastle speech, Foulkes (1997) similarly found that the young working-class informants used linking /r/ “to a far lesser degree” than the older middle-class speakers (89), a finding which suggests a phonological change whereby linking /r/ is being erased from the accent of this group. Bauer (1984), however, found no such evidence of a diachronic shift in progress with his RP speakers. He did, on the other hand, find an indication of a change with respect to intrusive /r/, as this variable was much more common in the latest recordings, suggesting an increase over time in the use of intrusive /r/. This apparent change was also noted in 1939 by Ida Ward, who wrote that “[t]here is no doubt that the

³⁹ A tone group (or intonation unit) is a stretch of speech pronounced with one pitch pattern. A tone group normally corresponds to a syntactic unit, either a clause or a phrase, and every tone group boundary is a potential pause.

intrusive *r* is spreading” (quoted in Lewis 1975: 41). In his article ‘The Cockneyfication of RP’, Wells mentions the use of intrusive /r/ among the sound changes that have possibly been imported into RP from Cockney (Wells 1994a: 202-203). There is however no substantial empirical evidence of a change in the frequency of linking versus intrusive /r/ in current RP.

The previous accounts of r-sandhi referred to above show that this is a highly complex variable. According to Lewis (1975: 39) the use or non-use of linking and intrusive /r/ “is a notable field for idiosyncratic variation on the part of individual speakers”. One and the same speaker can be found to drop common r-links and at the same time use intrusive /r/ quite freely. The variability does however not seem to be completely random. The use of linking and intrusive /r/ potentially correlates with social, stylistic, phonetic, prosodic, lexical and syntactic factors, although no categorical patterns have been revealed in previous studies. In the present investigation of r-sandhi all the different hypotheses about the occurrence of linking and intrusive /r/ will be tested against the spoken data.

4.4.4 T-voicing

In syllable-final intervocalic position, the phoneme /t/ may undergo what is often referred to as t-voicing. This phenomenon causes /t/ in words like *better* and *British* to be realised as what sounds like a very fast [d], a “voiced t”. Phonetically it is usually a rapid alveolar tap rather than a plosive, and it is normally voiced; but as Wells (1982: 248) points out, “it is an oversimplification just to call it [d]”. T-voicing does not necessarily entail a neutralisation of the opposition between /t/ and /d/. While /d/ is [d], /t/ typically has the intervocalic realisation [ɾ], the difference between them being primarily one of rate of articulation: the duration of the alveolar contact is longer in the case of [d] than for [ɾ]. The tap is produced by a single rapid contact by the tongue, and the closure is not long enough to produce neutralisation with /d/. However, /d/ can of course also be realised as a tap in intervocalic environments, in which case neutralisation with /t/ is possible (cf. Wells 1982: 250, Zue and Laferriere 1979: 1043-1045). In such cases there is a convention of analysing [ɾ] as an allophone of /t/ if there is a *t* in the spelling (e.g. *putting*) and as belonging to /d/ if it is spelt *d* (e.g. *pudding*).

T-voicing can take place word-medially, as in *letter*, *reality*, and across word boundaries, as in *what about*, *get away*.

Scholars are in some disagreement over exactly how to classify and transcribe the phonetic outcome of t-voicing. The prevailing view seems to be that which regards voiced t as

a tap, distinct from the [d] of /d/, and as belonging to the phoneme /t/. A more appropriate term would therefore perhaps be *t-tapping* (which is used in e.g. Tollfree 1999) or just *tapping* (e.g. Shockey 2003: 29-30). Some phonologists use the term *flapping* for the same phenomenon (e.g. Holmes 1994), while others distinguish systematically between flaps and taps in terms of the articulatory movements involved (e.g. Trask 1996, Shockey 2003: 29). *T-voicing* is the term used by e.g. Wells (1982, 2000) and Bell (1982, 1984), and the one adopted in the present study. This term then embraces both tapped variants and the rarer realisations of /t/ as a voiced plosive.

With regard to notation, the voiced or tapped allophone of /t/ is often written [t̚], with [̚] indicating voicing (e.g. Wells 2000, Jones 2003, Sivertsen 1960). Other symbols that have been used are [d] (Upton et al. 2001), [ɖ] (Trudgill 1986), [D] (Bauer et al. 1980) and [r] (e.g. Wells 1982, 1984, Cruttenden 2001, Shockey 2003). The latter symbol is the one that most accurately represents the usual phonetic result of t-voicing. The tapped [r] is also a common U-RP realisation of /t/ in intervocalic position (*very sorry*) and after /θ/ (*three*) (cf. Wells 1982: 282). In the present study the symbol [t̚] will be used to represent the voiced t variant, to indicate that the sound is interpreted as a realisation of /t/.

T-voicing is one of the many phenomena that are known to affect /t/ in English accents. Within Britain, t-glottalling is the feature that has been most extensively discussed and investigated, also with reference to RP (cf. Fabricius 2000). Voicing of /t/ has not been subject to the same scholarly attention. T-voicing is familiar as a feature of North American English, where it is considered the general realisation of intervocalic /t/ in all styles of speech. Most of the descriptions and studies of t-voicing therefore pertain to American or Canadian English (e.g. Harris and Kaye 1990, Woods 1991, Patterson and Connine 2001). T-voicing is also frequently used in Australian and New Zealand English (cf. Trudgill and Hannah 2002: 18, 24, Holmes 1994). Within the British Isles, t-voicing is apparently much less common, although the phenomenon has not been systematically investigated. According to Wells (1982: 324-325), Cockney is the only English variety where the voiced tap is a typical realisation of intervocalic /t/ (second only to the glottal stop). Some recent studies have reported t-voicing in a number of accents from all parts of the British Isles, such as Newcastle (Watt and Milroy 1999), Sandwell, West Midlands (Mathisen 1999), South East London (Tollfree 1999), Cardiff (Mees and Collins 1999), Glasgow (Stuart-Smith 1999) and Londonderry, Northern Ireland (McCafferty 1999). None of these studies have t-voicing as

the main topic, and many of the observations are informal. They do show, however, that t-voicing is a supra-regional feature that is found in current accents in many areas of Britain.

RP is generally not thought of as having t-voicing, and the traditional descriptions do not acknowledge the voiced tap as a regular allophone of /t/. In the phonetic literature there are, nevertheless, a few references to t-voicing in relation to RP. Wells (1982) notes that t-voicing is sometimes to be observed “in certain casual styles in British accents ranging from RP to Cockney” (250). Similarly, in Wells (1984) he observes that a tapped realisation of intervocalic /t/ “is by no means uncommon in England”, and states that “pronunciations such as [brɪtɪʃ] *British*, [mæɾə] *matter* are characteristic of casual style in at least some accents within RP” (56). According to Cruttenden (2001: 164) voicing of /t/ “is increasingly reported for a minority of RP speakers”. Shockey (2003: 30) claims that many speakers of Standard Southern British (which is her term for RP) occasionally tap their /t/s, “especially in often-heard words such as *British*”.

Several observations seem to suggest that t-voicing is becoming more frequent in all British accents. Roach, Hartman and Setter claim that “it is increasingly common to hear the American-style tapped /t/ in England” (Jones 2003: 527), and in a study of Newcastle speech, Watt and Milroy (1999: 29) found that [ɾ] was considerably more frequent among the younger speakers than among the older. Wells (1982: 250) believes that t-voicing in British accents represents the diffusion of an American borrowing, and regards t-voicing as “the first distinctly American phonetic innovation likely to spread in time to all accents of English”, a view shared by Taylor (1998) and Ramsaran (1990b). Alternatively, the apparent spread of voiced taps may represent independent innovations in several different places, which in turn may be explained by the “naturalness” of this feature. T-voicing is one of many reduction or lenition processes that tend to occur in intervocalic environments (cf. Shockey 2003), and has been characterised as a natural and phonetically motivated type of sound change. The naturalness can be attested for example by the rapidity and ease with which British speakers adopt this feature when they move to the United States (Trudgill 1986: 19-22). The naturalness of t-voicing reflects the general tendency to aim for ease of articulation. T-voicing seems to be consistent with the “principle of least effort” (cf. Wells 1982: 94), which leads speakers to pronounce utterances with a minimum of articulatory effort. A voiceless [t] between vowels requires a switching off and on again of the vocal cords, and it is easier to maintain the voicing throughout the alveolar articulation. This articulatory benefit may be an explanatory factor in the increased use of voiced /t/s, as ease of articulation can motivate

linguistic processes and “appears to be a general factor involved in linguistic change and variability” (Chambers 2002: 257).

The exact phonetic and stylistic conditions for t-voicing are not fully known and remain largely uninvestigated, at least for British English. In rhotic American English [ɾ] can occur between a sonorant and a following vowel or a syllabic /r/ or /l/ (e.g. [sɪɾi] *city*, [pɑ:ɾti] *party*, [wɪnɾr] *winter*, [bɛɾr] *better*, [bæɾtɫ] *battle*). It is commonly thought to be restricted to post-accentual position, and never to be followed by a stressed syllable (cf. Cruttenden 2001: 164, Jones 2003: 527). This seems too restrictive, however, as t-voicing can be heard in words and phrases like *whatever* [wʌɾtʰevr], *at all* [ət ʰɔ:l]. In non-rhotic British English the distribution of voiced /t/ is somewhat more limited, in that [ɾ] can only occur between vowels or between a vowel and a syllabic /l/ (and in a small number of cases also before syllabic linking /r/, as in *a matter of*). It seems to be most frequently heard between a short stressed vowel and a short unstressed vowel, but my own observations of broadcast speech suggest that the voiced /t/ is not limited to this environment. The present study thus investigates t-voicing word-internally and across word boundaries between vowels and between a vowel and a syllabic /l/ or /r/. The first vowel is always short, while the second can be short or long. The flanking syllables can be stressed or unstressed, but the /t/ must be syllable-final⁴⁰ and the second segment must follow immediately without a pause or an intermediate glottal stop.

In British English accents t-voicing is associated with casual or informal style, and seems to be limited to relatively rapid speech (cf. Wells 1982: 325). The observations of t-voicing in current British English (cf. above) are primarily based on studies of conversational speech in informal settings. The two variants [t] and [ɾ] thus have different social status and prestige. The voiced variant has been found to be commoner with working-class speakers (cf. Mathisen 1999, Watt and Milroy 1999, Holmes 1994, Woods 1991), and is thus the “rougher” variant. The voiceless plosive is the most high-status form, and is even considered “posh” by some Cockney speakers (cf. Wells 1984: 56). According to Taylor (1998: 150) and Trudgill (1986: 20), the voiced tap can be placed in the middle between the stigmatised, rough glottal stop and the prestigious, formal alveolar plosive, and tapping is thus a “compromise solution” for many who wish to avoid either of the more socially marked variants in intervocalic position. Moreover, the voiced /t/ is found to be more often used by males than by females (cf. Watt and Milroy 1999, Holmes 1994, Woods 1991), a pattern which is not surprising in view of the fact that females are consistently found to use more prestigious speech forms than

⁴⁰ For a definition of syllable-final, see 6.4.2.

males (cf. 3.1.4). In the present investigation, then, the voiced tap [ɾ] is expected to be found in stylistic variation with [t] in certain environments. The use of t-voicing is similarly hypothesised to correlate with the rate of speech and the gender of the speaker.

4.4.5 Smoothing

Smoothing is a type of elision whereby a prevocalic diphthong loses its second element and is reduced to a monophthong, as when the /eɪ/ of *chaos* /keɪɔs/ is pronounced [e]: [keɔs]. This reduction process, sometimes referred to as *levelling* (the term *Smoothing* was introduced by Wells 1982: 241) optionally affects any closing diphthong in the environment of a following vowel. In the English closing diphthongs the main prominence and length is attached to the first element of the vowel glide, and the transition to the second quality [ɪ] or [ʊ] requires an extensive tongue movement. The final elements of these diphthongs are thus inherently liable to be weakened or even lost in certain phonetic contexts (cf. Gimson 1964: 133). This monophthonging particularly affects the vowels of PRICE and MOUTH, /aɪ/ and /aʊ/, before a following /ə/ (Wells 1982: 238). The [ə] may be an integral part of the word, as in *tyre*, *fire*, *hour*, *shower*, or a suffix appended to the root, e.g. *buyer*, *highest*.

In the present study the term *smoothing* will be used only with reference to /aɪə/ and /aʊə/, and the discussion will be confined to words containing these vowel sequences.

When the triphthongs⁴¹ /aɪə/ and /aʊə/ undergo smoothing, the second element, [ɪ] or [ʊ], is obscured or lost, reducing the triphthongal sequences to diphthongal glides, [aə] or [ɑə]. The first element is retained and sometimes lengthened, giving [a:ə, ɑ:ə]. The [ə] may or may not be syllabic, giving four possible realisations of a word like *hire*: [ˈhaɪə] (two syllables), [ˈhaɪǝ] (one syllable), [ˈha:ə] (two syllables, smoothed), or [ˈhaǝ] (one syllable, smoothed). The centring diphthong resulting from smoothing may be even further reduced to a long monophthong with a quality ranging from front [a:] to centralised-back [ɑ:], adding a fifth variant of *hire*: [ha:].

⁴¹ The term *triphthong* strictly speaking refers to “a type of vowel where there are two noticeable changes in quality during a syllable” (Crystal 2003: 476). However, many scholars give a wider interpretation to the term, using *triphthong* with reference to any vocalic glide with three distinguishable vowel qualities, whether monosyllabic or disyllabic (cf. Roach 1992: 116). The term is thus frequently employed about [aɪə] and [aʊə] (cf. e.g. Hughes and Trudgill 1996: 4). In the present study *triphthong* is used in this way, with reference to disyllabic as well as monosyllabic realisations of [aɪə, aʊə].

The centring diphthongs derived from underlying /aɪə/ and /aʊə/ may or may not be qualitatively identical; some speakers merge them, others keep them apart. The quality of the first element is the same as that of the underlying smoothed diphthong. The starting point of the PRICE and MOUTH diphthongs varies considerably, however, and current RP embraces realisations ranging from front [a] to almost fully back [ɑ]. According to Wells (1982: 293) /aɪ/ typically has a more front starting point than /aʊ/, a difference which is preserved when the vowels undergo smoothing. The first element of smoothed /aɪə/ will then usually be somewhat fronter than that of smoothed /aʊə/, [aɪ] vs. [aʊ]. This analysis is supported by Gimson (1964: 133-134), Jones (1960: 110), Hughes and Trudgill (1996: 48) and Cruttenden (2001: 139), who all keep smoothed /aɪə/ and /aʊə/ distinct as [aɪ] and [aʊ] respectively. Gimson even uses different phonemic symbols for PRICE and MOUTH: /aɪ/ and /aʊ/. In modern RP, however, the trend is rather the opposite: the starting-point in PRICE tends to be more retracted than the one in MOUTH: [ɹaɪ] vs. [ɹaʊ] (cf. Upton 2004: 225, Cruttenden 2001: 132). The potential difference in quality may then be a more back starting-point for smoothed /aɪə/ as compared to /aʊə/. For many speakers, however, the starting-points of /aɪ/ and /aʊ/ are identical, so that smoothed /aɪə/ and /aʊə/ fall together, making *tyre* and *tower* homophones: [taɪ] (cf. Wells 1982: 239). Cruttenden points out that the qualitative difference between the two centring diphthongs is tenuous, and accordingly often levelled out, with the effect that both /aɪə/ and /aʊə/ “are frequently reduced to a diphthongal glide whose first element is a central open vowel” (2001: 139). Similarly, Hughes and Trudgill (1996: 48) state that “the notional difference in the first element of the reduced forms of *tyre* and *tower* is so small that the two sounds have become homophonous for many speakers”.

When the vowel sequences /aɪə, aʊə/ are reduced to monophthongs, the same potential distinction may apply, with the former having a more front quality than the latter: [a:] and [ɑ:] respectively. Smoothed /aʊə/ may then become phonetically identical with the phoneme /ɑ:/, making *tower* homophonous with *tar*. According to Wells (1982: 292) and Brown (1990: 82) this occurs frequently in RP. Those with similar starting-points in /aɪ/ and /aʊ/ will normally merge monophthongised *tyre* and *tower* to [ta:], distinct from *tar* [tɑ:]. Alternatively, those who merge *tyre* and *tower* may make both homophonous with *tar*, pronouncing all three as [tɑ:]. Such extreme levelling is, according to Cruttenden, rare, and criticised as an affectation in RP (2001: 139). Hughes and Trudgill, on the other hand, claim that the word *tyre* “is now increasingly reduced to /tɑ:/ ... with the same pronunciation as the

word *tar*” (1996: 4). Wells (1982: 239) comments that the relationship between [a:, ɑ:] and their underlying representations /aɪə, aʊə/ is so complex that it leads to restructuring, reflected in spelling mistakes such as *sar* for *sour*, which shows that phonetic [ɑ:] is reinterpreted as the realisation of the phoneme /ɑ:/. Similarly, Gimson (1964: 133) claims that “RP listeners tend to interpret ... [tɑ:] as either *tar*, *tyre* or *tower*”.

From the discussions above we can conclude that smoothing comprises two phonological variables, which can be written (aɪə) and (aʊə). In principle, both of these can have at least five different variants: one triphthongal pronunciation, [aɪə] or [aʊə]; two diphthongal forms, [aə] and [ɑə]; and two monophthongal ones, [a:] and [ɑ:]. In the present study only two variants will be distinguished: one triphthongal and one smoothed. The variable (aɪə) thus has the variants [aɪə] and [aə]. The former incorporates all realisations where three vocalic elements can be distinguished: [aɪə ~ ɑɪə], while the latter includes all realisations where the second [ɪ] element is lost: [a:ə, ɑ:ə, aə, ɑə, a:, ɑ:]. Similarly, (aʊə) has the variants [aʊə] and [aə], representing the triphthongal realisations [aʊə ~ ɑʊə] and the reduced realisations without [ʊ]: [a:ə, ɑ:ə, aə, ɑə, a:, ɑ:].

When the sequences /aɪə/ and /aʊə/ are reduced to centring diphthongs [aə ~ ɑə], the question inevitably arises of whether the “new” centring diphthongs should be regarded as having phonemic status. In those cases where the smoothed variant is realised as a monosyllabic [aə] it would seem natural to treat it as a separate diphthongal phoneme, distinguishing a word like *hired* from *hard*. For those speakers who have different qualities in smoothed /aɪ/ and /aʊ/, [aə] vs. [ɑə], both variants represent different phonemes, distinguishing minimal pairs such as *tyre* – *tower*, *hired* – *Howard*. Similarly, when the triphthongs are reduced to a monophthong [a:], which is much fronter than the realisation of /ɑ:/, this monophthongisation will give rise to a distinctive phoneme /a:/, separating e.g. *pyre/power* from *par*. In those cases where smoothed /aɪə/ or /aʊə/ is realised as a disyllabic sequence ['a:ə ~ 'ɑ:ə], it may be analysed as a sequence of two vowels [a:] + [ə], where the first element represents a monophthongal allophone of /aɪ/ or /aʊ/, restricted to prevocalic position.

Smoothing has been characterised as one of the principal special-context and allegro variants encountered in native-speaker RP (Wells 1982: 286). Gimson (1964: 133) mentions the reduction of /aɪə, aʊə/ as one of the most notable developments which have taken place in the RP vowel system. The phenomenon is however far from new. In the first edition of EPD,

Jones (1917: xxx) mentions monosyllabic variants [aə, a:] and [ɑə, ɑ:] as common realisations of [aɪə] and [aʊə]. Wells estimates that smoothing originated in London and/or East Anglia at the end of the 19th century (1982: 241). But it is not known whether it spread upwards into RP or downwards from RP. According to Cruttenden (2001), smoothing of /aɪə, aʊə/ “has long been common in southern England” (304), and it is claimed to be “very usual” among today’s RP speakers (Wells 1982: 286). The phenomenon has however been very little investigated, and the social and geographical distribution of smoothing remains largely unknown. Trudgill (1974) reports extensive smoothing in words like *seeing, saying, brewer, going, royal* among speakers in Norwich. Wells (1982: 241) notes that the feature “seems to be shared by all classes in London”, but adds that it is apparently more common at the two extreme ends of the social scale, in broad Cockney and U-RP, than in the intermediate varieties. A similar observation is made by Cruttenden (2001), who states that the more extreme form of smoothing, which results in levelling with /ɑ:/, is particularly common in Refined RP, but also condemned as a Cockney vulgarity (139). Empirical support of this claim is provided by Tollfree (1999), who, in an investigation of South East London English, found monophthongisation of *fire* and *power* primarily in broader speech.

The association of smoothing with U-RP and socially conspicuous speech, is evident from a comment in Lloyd James (1931), where the recommendation to BBC presenters is clear:

Such pronunciations as *nevaa, faa, waaliss*, for *never, fire, wireless*, will appear as an offensive affectation to those who are unacquainted with the class variant of which these pronunciations are so characteristic a feature. (Lloyd James 1931: 9)

With regard to present-day RP it is not clear whether smoothing is becoming more common or less common. Gimson predicted that the monophthongisation of /aɪə, aʊə/ and subsequent coalescence with /ɑ:/ was to become “one of the most striking sound changes affecting southern British English in the twentieth century” (1980: 140). This position seems to have been somewhat extreme, and Cruttenden (2001: 140) modifies the claim:

The most likely situation in RP at the moment seems to be a levelling of the two triphthongs to a single diphthong, usually [a:ə], but not the further reduction to a monophthong and no loss of the contrast with /ɑ:/.

The smoothed variants of /aɪə/ and /aʊə/ are generally not included in English pronunciation dictionaries. LPD (Wells 2000) is the only current dictionary that indicates a possible omission of the second element [ɪ] or [ʊ] in all the relevant entries, symbolised by an italicised vowel: [aɪə̃] *iron*, [paʊə̃] *power*.

Smoothing is often described as a feature of connected speech, expected to vary according to speech tempo. Hughes and Trudgill (1996: 48) claim that the full triphthongal variants may be heard in careful or slow speech, while the second vowel element is “usually omitted” in faster speech. Cruttenden (2001:139) and Wells (2000: 165) also characterise smoothing as a feature of rapid or casual speech, with [aɪə, aʊə] as the slower variants. Smoothing is further predicted to be more common in those cases where the schwa is an inseparable part of the word, e.g. *fire, hire*, than in words where [ə] represents a suffix, e.g. *flyer, higher* (cf. Jones 1960: 106, Hughes and Trudgill 1996: 48). These aspects will be pursued further in the analysis of smoothing in chapter 6.5.

4.4.6 Yod coalescence

Yod coalescence is a type of assimilation where the approximant /j/ (yod) fuses, or coalesces, with preceding /t, d/, resulting in affricates /tʃ, dʒ/, e.g. *tune* /tju:n/ → /tʃu:n/. Assimilation is a type of coarticulation and can be defined as “the influence exercised by one segment upon the articulation of another, so that the sounds become more alike, or identical” (Crystal 2003: 38). Coalescence is a reciprocal influence, where two adjacent sounds influence each other and fuse into a new segment. Yod coalescence is a subcategory of place assimilation, whereby an alveolar consonant coalesces with a following palatal /j/ to produce a palato-alveolar sound. The phenomenon has also been referred to as *palatalisation* (cf. Roach 1992, Shockey 2003), as the articulation is “shifted nearer to the hard palate” (Roach 1992: 76). However, this term is somewhat misleading, because it is only the first element, the alveolar consonant, which is palatalised. *Yod Coalescence* is the term used by Wells (1982, 1994a, 2000), who restricts the term to cases involving /t, d/ + /j/. Cruttenden (2001) merely calls it *coalescence*, and includes /s/ and /z/ in his discussion of the phenomenon, as in *miss you* /mɪʃu:/, *as yet* /əʒet/ (2001: 286). The present investigation concentrates on the coalescence of /j/ with a preceding alveolar plosive /t/ or /d/. In terms of phonological variability, yod coalescence thus

embraces two phonological variables, (tj) and (dj), each with two variants representing the uncoalesced and the coalesced pronunciation respectively: /tj, tʃ/, /dj, dʒ/.

Yod coalescence can occur in three different environments:

- (a) across word boundaries, e.g. *did you* /dɪdʒu/, *won't you* /wɒntʃu/.⁴²
- (b) word-internally before an unstressed vowel, e.g. *educate* /'edʒukeɪt/, *statue* /'stætʃu:/.
- (c) word-internally before a stressed vowel, e.g. *Tuesday* /'tʃu:zdeɪ/, *reduce* /rɪ'dʒu:s/.

Yod coalescence across word boundaries is a type of contextual assimilation that typically varies according to style and speech tempo. The word-internal coalescence in (b) and (c) can be viewed as a kind of lexical-incidental variation, where individual words have alternative phonological structures (cf. 4.2.1).

Coalescence of type (a) and (b) is quite common in fluent colloquial speech, and well within the boundaries of RP (cf. Cruttenden 2001: 212, 286). In words of type (b) the coalesced pronunciation is often preferred, and Wells (1994a: 203) claims that in this environment “RP is drifting towards categorical coalescence”. The third kind of coalescence, which only occurs before the vowels /u:/ and /ʊə/, is much less frequent, and is widely perceived as non-standard or non-RP (cf. Wells 1994a: 203, Taylor 1998: 148). The present study investigates only yod coalescence of type (c), within a stressed syllable, where coalescence traditionally has been excluded from RP, and where there is evidence of an ongoing change.

Yod coalescence is related to yod dropping, which involves the elision of /j/ after consonants (*tune* /tju:n/ > /tu:n/). Both phenomena can be characterised as yod cluster reductions, and thus processes of simplification. According to Lutz' (1991) model of language change, most consonantal changes in English are phonotactically and articulatorily motivated, in that they aim for an improvement in the direction of a preferred CV syllable structure. The linguistic preference for single-consonant syllable onsets appears to be a universal feature (cf. Vennemann 1988, Shockey 2003: 33), as “a regular alternation of consonants and vowels is more natural than clusterings” (Wells 1982: 95). Yod coalescence and yod dropping can then be interpreted as devices for simplifying complex syllable onsets in order to achieve or at least approach the preferred CV structure. While yod dropping reduces the complex CCV onset by eliding the second consonant, yod coalescence offers an alternative strategy by which the alveolar and palatal consonants merge to form a palato-alveolar affricate. Both processes are

⁴² This type of coalescence usually involves clitic *you* or *your* (cf. Wells 1994a: 203).

common patterns of improvement, according to Lutz (1991: 224). Whether affricates should be analysed as one consonant or a sequence of two consonants is a matter of some controversy (cf. Wells 1982: 48-49). There is, however, plenty of evidence to support a single-unit analysis, at least for /tʃ/ and /dʒ/. All affricates are combinations of plosion and friction, but unlike other affricates, /tʃ/ and /dʒ/ have an unrestricted distribution in that they can occur initially, medially and finally in a word (cf. Crystal 2003: 16), and according to Wells, native speakers “usually think of [tʃ] and [dʒ] as single sounds” (1982: 48). Moreover, Harris (1994: 39-41) shows that, for quantitative purposes, affricates function like simplex consonants and not like clusters. In a related discussion of palatalisation of [s] before [i] and coalescence of [s] and [j], Bybee (2001: 71-72) argues that this type of assimilation is a temporal reduction, which involves at least partly simultaneous articulation of two previously sequential units.

The tendency to elide or modify /j/ in C/ju:/ sequences can also be viewed as a consequence of the low functional load of /j/ in this environment. The consonant distinguishes a few word pairs, such as *do – due*, *moo – mew*, but is in most cases redundant (cf. Rastall 1997). RP has lost historical /j/ in words like *blue*, *prude*, *rude*, and /j/ is disappearing in a number of other items, especially after /s/ and /l/, such as *sue*, *super*, *lure*, *delude*, *illusion*. After /t, d/, however, the tendency is towards yod coalescence. The London Cockney accent traditionally has yod dropping also after the alveolar plosives, but, according to Wells (1982: 330), a switch has taken place in Popular London speech towards yod coalescence instead of yod dropping in words like *tube*, *due*. The origin of this change is unknown, but Wells suggests the possibility that “the [tʃ]-, [dʒ]- forms have long existed in South London ... and that from there they have spread to other parts of London” (1982: 330).

Yod coalescence may then be seen as an intermediate variant between the traditional Cockney yod dropping and the traditional RP forms with /j/. Yod coalescence is often referred to as a typical Estuary English feature (cf. Coggle 1993: 51-52, Altendorf 2003), and Upton (2004: 229) characterises the coalesced forms as “a less formal alternative” to the more “careful” forms with /j/. Altendorf (2003) investigates yod dropping and yod coalescence with a group of school girls from London, Colchester and Canterbury. She concludes that, while yod dropping “seems to remain a London working-class variant” yod coalescence “seems to be spreading socially into higher social classes and perhaps also into other parts of the south-east” (2003: 100).

While yod coalescence may be spreading socially, it is also the target of negative attitudes and complaints in the media. Kerswill (2001: 12) quotes from a letter in the Daily Telegraph on 17 June 2000, where the reader deplores the spread of the “virus of ‘London lad’ speak”:

I am tired of hearing presenters – from weather girls to news readers – refer to “Chewsday” [Tuesday] ... and using “jew” as a word to replace many others, as in “Jew agree?” [Do you agree?] or “Jew” [due] to rain there was no play at Chrent [Trent] Bridge today” ... The insidious degradation of spoken English saddens me and someone ought to stand up and say “enough”.

There is no doubt that there has been a development taking place in RP, whereby word-internal yod coalescence has become increasingly common. The spreading of the phenomenon is evident when we consider the great number of words which now have variable coalescence. Ramsaran (1990b) mentions yod coalescence as one of the current trends within modern RP. A process of yod coalescence is the origin of the palato-alveolars used by all speakers in words such as *nature* /neɪtʃəl/, *soldier* /səʊldʒəl/, *vision* /vɪʒn/, *sugar* /ʃʊgəl/ (cf. Dobson 1957: 957-960), and it seems reasonable to assume that the coalesced pronunciation in time will become the norm also for words which today alternate between the coalesced and the uncoalesced variants. Yod coalescence in stressed syllables has long been resisted by RP, but is now beginning to be accepted as part of modern RP by a number of linguists. The changing attitude to forms such as /tʃu:b/ *tube* and /dʒu:n/ *dune* is reflected in various writings on the accent. Wells (1982: 338) states that “RP always has /j/ in words like *tune*, *duke*”, and according to Ramsaran (1990b: 187) there is “considerable resistance” to yod coalescence when /tj/ or /dj/ occurs before a stressed vowel. In 1994a, however, Wells writes that while coalescence within a stressed syllable is “still on the whole perceived as non-RP”, he considers it “likely that here, too, coalescence may penetrate RP within a few decades” (Wells 1994a: 203-204). He attributes the increasing use of coalescence to the influence of the Cockney accent. In other sources Wells seems to regard the avoidance of stressed yod coalescence as one of the hallmark features that distinguish RP from Estuary English (Wells 1997b, 1998b). In LPD (Wells 2000) pronunciations such as /tʃu:n/ *tune* and /ɪnˈdʒʊəl/ *endure* are marked as non-RP, and yod coalescence in stressed syllables is referred to as belonging to Estuary English, and thus localised to the southeast (2000: xiii). In Wells (1999a), however, the author reports from an extensive poll of British English pronunciation preferences, where

respondents from many parts of the country show a preference for coalescence in words like *tune* and *during*, indicating that yod coalescence is not confined to the southeast. Indeed, the realisation of *tune* with /tʃ/ is reported to be “most readily accepted” in Scotland (1999a: 44). Moreover, the pronunciation items selected for Wells’ survey are described as “mostly ones for which an association between response and regional origin was not anticipated” (36). The use of yod coalescence in stressed syllables can thus be characterised as a supra-regional feature and as such acceptable as part of non-regional RP.

Cruttenden (2001) has no reservations about embracing yod coalescence in stressed syllables as an RP feature. He lists this type of coalescence as one of the “changes well established” in modern RP, and among the pronunciations “now typical of a majority of speakers of General RP” (83). Likewise, Upton (2004: 229) characterises the resistance to coalescence before a stressed vowel as a feature of conservative RP, and in the introduction to ODP (Upton et al. 2001), the use of affricates in words like *tune* and *reduce* is described as “very frequent in RP” (xiii).

In light of the above observations, it is evident that there is a change in progress, whereby yod coalescence in stressed syllables is becoming increasingly common in RP. Wells (1999a: 44-45) reports a general British English increase in the preference of coalesced forms, as his opinion poll survey shows that an affricate in *tune* and *during* is much more acceptable among the younger respondents than among the oldest. The pronunciation dictionaries also reflect this development: Yod coalescence in stressed syllables is not included in the first edition of *English pronouncing dictionary* (Jones 1917). In the 1967 edition of EPD, coalescence is included as an alternative only in the word *during*. In the 16th edition (Jones 2003) the affricate is listed as a possible variant in most of the relevant words, but inconsistently: coalescence is for example allowed in words like *tune*, *due*, *produce*, *institution*, *prostitution*, but not in *endure*, *subdue*, *reduce*, *intuitive*, *studio*, *student*. LPD (Wells 2000) includes coalescence as an alternative in all the relevant entries, but the pronunciation is marked as non-RP. ODP (Upton et al. 2001) is the only dictionary that allows /tʃ, dʒ/ in stressed syllables without any reservations.⁴³

⁴³ According to the ODP editors, the inclusion of /tʃ, dʒ/ forms is a reflection of their slightly “broader” definition of RP: “Implicit in the British English model presented here ... is the view that a larger group of people can lay claim to possession of an RP accent than has often hitherto been acknowledged. ... As a result of this policy, certain regularly-occurring pronunciation features which have to date been ignored or marked prescriptively are allowed where they are now judged to be established features of RP. Notable examples of such features are [-tʃ-] in place of [-tj-] and [-dʒ-] in place of [-dj-] in such words as *destitute* and *reduce*” (Upton et al. 2001: xiii).

5 Methodology

5.0 Introduction

This chapter describes the procedures for obtaining, analysing and processing the data that form the basis for the present investigation. It discusses the principles of the methods of speaker sampling, data extraction, auditory analysis and statistical quantification, as well as the caveats associated with these approaches.

The purpose of any linguistic investigation is to be able to make generalisations about linguistic usage, at some level or other. This presupposes that the results of the investigation are *valid* and *reliable*. Validity and reliability are criteria for assessing the quality of the data collected. Validity means that the data are representative of the field of interest and that they truly provide information about what the study purports to investigate. The concept of reliability refers to the replicability of the study, or the likelihood that the results reflect actual usage patterns in the relevant population and would be consistent over time. The following sections will show how the chosen methods of data sampling and analysis aim to secure that the study is valid and reliable.

5.1 The speakers

5.1.1 Selecting a sample

In an investigation into current pronunciation, the first decision to be made concerns the selection of subjects to provide the phonetic data. In an investigation of RP the issue of speaker selection is somewhat more problematic than with most other accents (cf. 2.3). RP speakers are not concentrated in any particular geographical region or city, and they have been estimated to constitute only 3 to 5 per cent of the British population (Trudgill 1974). It is therefore not altogether a straightforward task to collect large amounts of sound recordings from a sizeable number of RP speakers in a given community. Although it is hard to find RP speakers, it is easy to find RP, as the accent is frequently heard in the broadcast media. In the newscasting genre RP still dominates as the most frequently used accent of English, despite of

the changing social status of the variety (cf. 2.6). Broadcast news is therefore well suited as a source for observing RP in use. In addition to being easily accessible, newscasts provide continuous stretches of speech produced mainly or entirely by one speaker. Moreover, the high technical quality of the sound enables the researcher to get good recordings, which in its turn facilitates the auditory analysis.

While the use of news presenters offers many practical advantages for the researcher, there are obvious weaknesses with using such a narrow selection of speakers. Within traditional variationist research, informants should be randomly selected and representative of the overall population, in order to increase the validity of any generalisations. A sample consisting of only newsreaders is clearly not representative of the RP speaking community as a whole. Modern variationists have, however, largely abandoned strict random sampling in favour of *judgement* sampling, where the investigator defines in advance the type(s) of speakers to be studied, and then finds a sample of subjects who fit the specified criteria (cf. Milroy 1987: 26). Such judgement sampling has been found to be more suitable for linguistic studies, as formal representativeness does not necessarily yield greater insight into linguistic usage patterns (cf. Romaine 1980). In the present study, the selection of speakers was made on the basis of linguistic criteria, as they should all speak RP. In addition, all channels and both genders should be equally represented. The selection was otherwise random, in that I included those speakers from whom I first managed to collect one hour of speech.⁴⁴ Even though the newsreader informants are not representative in the strict statistical sense, they represent an important group within the RP-speaking community, whose speech is heard by millions of people (cf. 3.3).

Another limitation of the present sample is that the stylistic range is fairly narrow, since all the speech is produced in the same, rather formal, setting of news broadcasting. Any generalisations can therefore not automatically be applied to RP speech in general. On the other hand, news presenters are also part of the population and the speech community, and receive linguistic and evaluative input from a variety of sources (cf. 3.3). Pronunciation features or speech patterns that occur with a large number of the speakers can thus be assumed to apply also to other RP speakers outside the field of newscasting. Moreover, because of the formality of the newsreading situation we can assume that the pronunciation features employed by the presenters are quite well-established in non-regional speech. Most linguistic changes are known to start in casual, informal styles and then spread to more formal

⁴⁴ Some presenters disappeared from the air before I had recorded enough speech data, and had to be replaced.

styles as they become more established or lose their stigma. The regular presence in newscasts of a feature “from below” can thus be viewed as an indication that it is becoming established, and we can presume that its frequency will be greater in informal styles.

The spoken data for the present study have been collected from three different television news channels: BBC World, Sky News and the ITV News Channel. Each channel is represented with ten speakers, five male and five female presenters, listed below. The speakers will throughout be referred to by their numbers.

Table 5.1. List of speakers

Speaker number	Gender	Channel
1-5	Male	BBC World
6-10	Female	BBC World
11-15	Male	Sky News
16-20	Female	Sky News
21-25	Male	ITV News
26-30	Female	ITV News

A sample of 30 speakers was deemed suitable for the purposes of the present study. The sample could then be equally divided between male and female speakers, as well as between the three channels, and it is large enough to enable statistical analysis.⁴⁵

All the newsreaders present currents affairs and international news. Business news presenters and sports newscasters have not been included. The reason for this is primarily that the number of newsreaders is larger than the number of sports and business presenters, and the majority of the newsreaders speak RP, while non-RP accents dominate in the other group. Also, the sports and business bulletins tend to be fairly short, and these presenters thus produce less continuous speech than the newsreaders. A comparison of the speech of the various types of presenters might have yielded some interesting results. However, since the present study investigates RP speech only, the most practical solution was to concentrate on newsreaders.

⁴⁵ Samples for linguistic studies are typically much smaller than for other types of surveys. It has been argued that large samples are not as necessary for linguistic research, because linguistic usage tends to be more homogeneous than other types of social behaviour, and very large samples tend to be redundant (cf. Sankoff 1974: 22-23).

Biographical information about the speakers is limited. However, some information is available in the form of presenter profiles on the web pages of the respective channels and on the TV Newsroom site (<http://tvnewsroom.co.uk/staff/staff-a.php>). These short biographies reveal that the presenters have varied geographical backgrounds. Many of them have previously worked as reporters or correspondents, and some have been presenting for several different stations.

Although there is a considerable age span between the speakers, they can all be said to belong to the middle age group, which tends to be linguistically fairly stable (cf. Hudson 1996: 16, Chambers 2002: 203). They are in those adult years that typically involve “full engagement in the work force and family responsibilities (30-59)” (Labov 2001: 101). They are past the formative periods, but there is still scope for linguistic change, not least from the pressures of the “linguistic market-place” (Chambers 2002: 194-203). By focusing on the middle generation, we avoid the most conservative and obsolete forms of the elderly and the often ephemeral trends of the young. The forms used by speakers in the middle age group can be assumed to be well-established, and in a study that aims to describe the state of current RP features, it can be argued that the speech of this group is the most important for such a purpose (cf. Gimson 1984: 48). At the same time, we lose information that would emerge in an age-graded sample with data from all age groups. A study including both adolescent and old speakers would presumably give a clearer picture of how the RP accent is changing. The speech data in the present study will be compared with existing descriptions and previous studies.

The speakers then constitute a socially homogenous group. They are all in the middle age group, and they have the same occupation and are therefore at the same level in the social hierarchy. The only social variable is gender. Gender has, however, been found to be one of the most important factors of speech variation (cf. 3.1.4), and the analysis of the six variables will therefore include a comparison of male and female scores.

As the present study is an investigation of RP as spoken in the news media, not of broadcast speech in general, I have included only those presenters who have a (Mainstream) RP accent, as defined in chapter 2. With non-localisability as the sole criterion for selecting the speakers, I have excluded speakers with regional pronunciation features, that is, features associated with a particular region or city in Britain. The most common non-RP (and thus disqualifying) features were:

- rhoticity (characteristic of Scottish, Irish and West Country accents, as well as non-British accents like American and Canadian English)

- short front vowel [a] in BATH (characteristic of Northern English)
- [ʊ] or [ə] in STRUT, (characteristic of Northern English and Welsh English)
- diphthong shift (characteristic of southeastern accents like London English)

All the speakers have accents that fall within Mainstream RP (cf. 2.2), with the basic phoneme system described in 2.5. One potential informant was excluded because his accent was closer to the conservative upper-crust RP (cf. Wells 1982: 280). There were however no limitations at the other end of the scale, since one of the main aims of the study is to examine developments in RP. The speaker selection thus “allowed” for features that are new or represent ongoing changes, such as t-glottalling, r-fronting and l-vocalisation (cf. 4.2.2), the only criterion being that the features are widespread and non-regional. The vowel features listed under B in 4.2.2 were all present to varying degrees; most of the speakers have occasional t-glottalling pre-consonantly and word-finally, but r-fronting and l-vocalisation were not observed.

The use of linguistic criteria in the selection of speakers is potentially problematic. If a study aims to find out how a group of people speak, and the informants are selected on the basis of how they speak, there is an obvious danger of circularity. The selection principles applied here would therefore have posed a serious problem if the aim of the investigation had been to find out how newsreaders speak, or what RP is. However, the main purpose of the present study is to examine the status and usage levels of a set of variable features that are known in advance to occur in RP (cf. chapter 1). The sampling method does then not invalidate the study, as long as no restrictions or qualifications are imposed with regard to the features in question. The use of linguistic criteria is not an ideal way of selecting RP speakers, but other methods (based on e.g. social criteria) are even more problematic (cf. 2.4).

In order to provide independent confirmation of my judgement of the 30 newsreaders as speakers of RP, I asked two phonetic experts, who both have written about RP, to listen to samples of the recordings and give their opinion on the accent of the speakers.⁴⁶ One of the judges listened to 22 of the speakers, the other heard 20 speakers. The sound excerpts each lasted between one and two minutes, and contained a range of features which should provide enough information for the judges to assess whether or not the accents represent RP. With the exception of one speaker, the judges agreed that the newsreaders speak RP. While there was a certain amount of inter-speaker variation, with some showing more progressive features than

⁴⁶ I am very grateful to Clive Upton and Thor Sigurd Nilsen for taking the time to listen to the excerpts and provide assessments of the speakers.

others, they were all judged to fall within the boundaries of modern RP. The only exception was speaker 6, who was judged by one of the listeners as having a near-RP accent. Such a level of agreement is deemed highly satisfactory for the purposes of the present study.

No attempt has been made here to distinguish between native and adoptive RP speakers (cf. 2.2). Such a discrimination would demand detailed information about the presenters' backgrounds. It is nevertheless likely that some of the informants are adoptive RP speakers, in view of the estimated low percentage of native RP speakers in Britain and the fact that two of the three BBC World presenters who were interviewed by the author confessed to having acquired the RP accent as adults (see 3.3.3). RP in particular is an accent that many people adopt during their years of education or when they enter the work force. The number of RP *users* is therefore much larger than the number of native RP speakers. The distinction between native and adoptive speakers is however of little relevance to the present study, which investigates RP as it is used in news broadcasting, and not necessarily RP as it is spoken by people who have had the accent since early childhood. Moreover, the definition of a native speaker is in itself problematic. As Milroy notes, "native speakers cannot be picked out easily just by listening to their accents" (1987: 23), and it is not evident how young a speaker has to be in order to acquire an accent successfully or completely (cf. Trudgill 1986: 34-37).

5.1.2 The channels

As mentioned above, the speakers are newsreaders from three British television news channels: BBC World, Sky News and the ITV News Channel. These channels were chosen because they are all available via satellite, and therefore easily accessible. BBC World is the British Broadcasting Corporation's 24-hour international current affairs TV channel, and was launched in January 1995. It was formerly known as BBC World Service Television, but unlike BBC World Service radio it is commercially funded. Sky News was Britain's first 24-hour television news channel and originally launched as part of the 4-channel Sky Television satellite package in February 1989. The ITV News Channel is also a 24 hour television news channel. It launched in August 2000 as the ITN News Channel, and became the ITV News Channel on 30 September 2002. After the relaunch, the format, schedule and logo changed, and many of the presenters were replaced. The ITV data in the present study consist of recordings from both periods.

As transmitters of news, all three channels belong to the “serious” part of broadcasting and are very similar in style and format. There is nevertheless a slight difference in formality, with BBC World as the most formal and Sky News as the most informal channel. In the former the newsreader is alone in the studio, while the Sky News presenters usually appear in pairs and the tone is often more colloquial, with occasional joking and bantering. Also, there are more commercial breaks on Sky News. On the ITV News channel they vary between the two presentation formats, and the level of formality is somewhere in the middle.

The differences in format and tone are related to differences in the channels’ self image, their expected audience and their ‘media goals’ (cf. Leitner 1980). All three channels present international news, but Sky News and ITV News are more directed towards a British audience and have an added focus on domestic stories. BBC World is part of the BBC, which has a long and proud history and is often heralded as the most widely respected broadcaster in the world. The BBC’s policy statement says about its news section that “our strength is in the depth of knowledge in our journalism, the quality of analysis and range of subjects, original first hand reporting and investigations from around the world”, (www.bbc.co.uk). On the BBC World website the channel is described as “an integral part of the BBC’s commitment to global broadcasting”. The focus is on information, education and quality: BBC World “keeps its viewers not just informed, but well-informed, with in-depth analysis and cutting edge interviews” (www.bbcworld.com).

While the BBC strives to “inform, educate and entertain”, in that order, Sky Broadcasting lists as its first commitment to “[deliver] entertainment to our customers” (www.sky.com). The Sky News channel aims to be “first with the latest”, and focuses on updating its audience: “Sky News ensures that 80 million viewers in more than 40 countries across the world are kept fully up-to-date with the latest breaking news” (www.sky.com/skynews).

ITV is the biggest commercial television network in the UK, and aims at “broadcasting the most talked about television and making a major contribution to the UK’s culture, economy and communities” (www.itv.com). The ITV News Channel presents a mixture of entertainment and information and aims at setting the agenda: “the ITV News Channel puts the drama back into news: trusted presenters breaking stories first and leading with agenda-setting interviews”.

Bell’s (1984) findings on radio speech showed that stations with similar audience share similar speech styles, and that individual differences are minimised as presenters tend to cluster around the station mean frequency for linguistic variables. The analysis of broadcast

RP will include a comparison of the three channels to look for evidence of a ‘station style’. However, because the similarities between the three channels are greater than the differences, and because only RP speakers are selected, the speech patterns of the presenters are not expected to be significantly affected by channel affiliation.

5.1.3 Style

The speech situation in news presenting is very different from the spontaneous conversation or sociolinguistic interview. It is a public and relatively formal setting that involves scripted speech and a high level of speech-consciousness (cf. 3.1.4.4). These contextual aspects are assumed to have certain linguistic implications.

As the study focuses on broadcast speech, there will be no comparison with conversational data. However, in most of the news bulletins there are occasional shifts in the speech situation, where the presenter interviews a subject, either in the studio or via satellite. In order to estimate the effect of this shift at intra-speaker level, the speech produced during the interviews will be compared with that produced during newsreading. The two speech situations will be referred to as *styles*.⁴⁷

Within the Labovian paradigm stylistic variation is typically analysed in terms of the degree of standardness of features and the level of formality of the situation. In the context of newscasting the level of formality remains fairly stable. Also, all the speech included here is within the boundaries of RP, and we can therefore hardly talk of standard versus non-standard features. Rather, the stylistic shift from newsreading to interview primarily involves a change in the amount of attention paid to speech, which is presumed to affect features of connected speech or “natural” phonological processes (cf. 3.1.4.3). When presenting the news, the newsreader is the only speaker, and he looks directly into the camera. Most of the speech is scripted and often rehearsed, and the focus of attention is on the transmission of information. This style presumably involves a high degree of speech-consciousness and articulatory care. During an interview, on the other hand, the presenter is not the only speaker, but relates to an interlocutor. The speech is no longer solely ‘transactional’, but also ‘interactional’ (cf. Brown 1990: 7). There is a stronger element of improvisation, since much of the speech is unscripted,

⁴⁷ In the present study, then, ‘reading style’ and ‘interview style’ are very different from what is normally referred to as reading style and interview style in traditional sociolinguistic studies.

and the focus shifts from the text to the interviewee. The result is presumably a decrease in the level of attention paid to speech and articulation.

The amount of data collected during interviews is rather small compared to the newsreading data. The interview speech constitutes only 10-15 per cent of the total recorded material. Moreover, the presenters do not produce equal amounts of speech in interview style, and a few have almost no interviews at all. The interview data have therefore been analysed collectively, as inter-speaker comparison was not possible. Because interview style constitutes such a small proportion of the corpus, only the most frequent variables could be analysed for stylistic comparison. R-sandhi and t-voicing are the only variables where the number of tokens is large enough for quantitative stylistic analysis. These variables are also the most interesting for the kind of stylistic comparison which is made here. While the other four variables (with the possible exception of smoothing) are realisational or lexical-incidental features that affect groups of words independently of the surrounding context, r-sandhi and t-voicing are connected speech features that mainly affect sounds across word boundaries. They are largely phonetically motivated “natural” processes that enhance the ease of articulation (cf. 3.1.4.3). A number of connected speech features have been documented to relate directly to attention to speech, and r-sandhi and t-voicing are therefore the variables most likely to display stylistic variation as defined here.

Another stylistic factor that potentially affects connected speech features is speech rate. A rapid speech rate is associated with fluency and casualness, while a slow delivery is said to be characteristic of careful or self-conscious speech and formal styles. The speech of newsreaders is generally somewhat slower than conversational speech. According to Crystal (1992: 121) 400 syllables per minute is the norm for spontaneous conversation, while BBC newsreaders tend not to exceed 300 syllables. Speech rate can also be measured in words per minute, by dividing the total word count by the total number of minutes. A rough estimate of the speech rate of the 30 presenters shows that the average rate is 168 words per minute, which equals approximately 250 syllables per minute (cf. Darley and Spriestersbach 1978: 257). The speech tempo of each individual presenter remains relatively stable. It sometimes increases slightly in interview style, but not always, and not with all speakers. The effect of speech rate can therefore not be measured at intra-speaker level, as there is not enough data. There is however considerable inter-speaker variation in the rate of utterance, as some presenters generally have a slow tempo, while others are fast speakers. The individual speech rate ranges from 151 to 187 words per minute. This variation may then be expected to correlate with variation in the usage levels for variants that reflect phonological processes.

5.2 Data collection and analysis

5.2.1 Data collection

5.2.1.1 Collecting the data

The data for the present study are drawn from recordings of television newscasts. The speakers are 30 newsreaders from three different channels, presented in the previous section. The total corpus consists of 30 hours of speech, one hour from each presenter,⁴⁸ altogether approximately 310 000 words. The recordings were made by the researcher in the period between January 2002 and April 2004.

In studies of linguistic variation and change, the most common method of data collection is the sociolinguistic interview (cf. Labov 1984). The use of broadcast speech involves a very different sampling strategy, which offers both advantages and drawbacks. Television speech is readily available and easily accessible, and offers good quality recordings (cf. 1.4 and 5.1.1). When you use the speech of newsreaders, who occur regularly on the air, it is easier to secure an equal amount of speech from each informant than when you interview more or less talkative people face to face. There are, however, considerable challenges in collecting a sizeable, balanced corpus of newsreader speech. First of all, I had to make sure that I would find five male and five female RP-speaking presenters on each television channel.⁴⁹ Second, once the speakers were identified, it was difficult to plan the recordings, since there was no way of knowing when the relevant presenters would be on.⁵⁰ Moreover, because of the repetitive nature of newscasts, there is little point in recording more than one bulletin per day per speaker, even though the same presenter may be on for several hours in successive bulletins, as you would only end up with a number of near-identical recordings.

Another major obstacle is the sheer amount of recordings necessary to obtain enough data. The material consists solely of speech produced by the newsreaders. All speech by other individuals, such as reporters and interviewees, is excluded, as are jingles and commercials. During a news bulletin, the amount of spoken data produced by the presenter is usually very

⁴⁸ The exact amount of individual speaking time varies somewhat, since the presenters have different speech rates (cf. 5.1.3), but they all produce approximately 10 000 words.

⁴⁹ I had originally planned to include more channels, with fewer presenters from each, but due to various practical difficulties I had to abandon that idea and opted instead for only three channels and a larger number of speakers from each.

⁵⁰ A few of the presenters were on air at fixed times, but not every day.

small. A half-hour broadcast may result in only a few minutes of actual speech from the newsreader. It follows, then, that a large volume of recordings is needed in order to collect 60 minutes of speech from each of the 30 presenters. The recordings then have to be edited, in that the relevant speech has to be extracted from the video tapes onto audio tapes, an elaborate and time-consuming process.

5.2.1.2 Sample size

How much data one should collect depends on a number of factors. The most important consideration is that the corpus is large enough to enable quantification and statistical analysis where relevant, which in turn secures reliability. The present study investigates variation at both inter- and intra-speaker level. In order to reveal patterns of systematic variation between speakers, one needs not only a certain number of informants, but also a certain amount of tokens from each informant: the intra-speaker patterns have to be reliable if the study of inter-speaker variation is to be meaningful. How many tokens are needed for the intra-speaker analysis depends on the number and types of constraints one wishes to investigate. There are no exact or universal principles. Labov (1972: 204) states that patterns of inter- and intra-speaker variation can emerge with “no more than five or ten instances of a given variable for each speaker”. Guy (1980) demonstrates that with fewer than 10 tokens, there is a strong chance of random fluctuation, while with more than 10, there is 90% conformity with the predicted pattern. He suggests 30 tokens per factor as a suitable goal if the results are to be reliable. While this figure is a useful point of departure, it is evident that variables differ in their sensitivity to sample size. Factors such as the frequency of the relevant phonological class in the language, and the relative frequency of the variants, should be taken into consideration.

The variables in the present study differ considerably with respect to their general frequency, and with respect to the number of tokens needed. R-sandhi and t-voicing take place primarily across word boundaries, and relevant sites occur quite frequently in connected speech. Both features are hypothesised to correlate with a number of linguistic factors, and a large number of tokens is needed in order to obtain a sufficient amount of items in the relevant environments. In the case of smoothing, we need enough tokens to be able to compare variant frequencies in tone group-final versus tone group-internal position. A satisfactory analysis of CURE lowering and yod coalescence requires a certain lexical diversity, as lexical class membership and word frequency may affect the choice of variants.

The analysis of GOAT allophony does not depend on a large number of tokens, since one can establish fairly quickly if a speaker has different realisations in *GOAT* and *GOAL*.

With the exception of intrusive /r/ and yod coalescence, all the variables are represented by a fairly large number of tokens, well above the putative statistical threshold of 30 per speaker.

In the data analysis I have followed Labov's (1972: 72) "principle of accountability", and included all tokens for every variable, which means that speakers are represented with different numbers of tokens. This has the potential of skewing the inter-speaker data, if, for example, one speaker has a high occurrence of a word or structure which favours a particular variant, while another speaker has none. In order to secure comparability between speakers, some researchers limit the number of tokens included for a particular item, so that speakers are represented with the same number of occurrences, and any linguistic bias is minimised (cf. Milroy and Gordon 2003: 162-163). However, if such limits are imposed, this may conceal important lexical patterns, and make it harder to explore the linguistic factors conditioning the variation. In the present study, the problem of unequal speaker contributions is minimal, because the number of tokens is on the whole large, and the sample is fairly homogeneous (see also 5.2.3.2).

5.2.2 Analysing the data

Nothing that we can do will ever make speech visible.
Speech is a jumble of noises and rhythms and tunes.

A. Lloyd James (*Broadcast English II*, 1930)

5.2.2.1 Identifying variants

In the investigation of phonological variation, one of the first steps is the identification of the linguistic variables to be studied. The variables in the present study were selected partly on the basis of previous descriptions of RP, and partly on the basis of my own observations of RP speech (cf. 4.3). The identification of a variable includes a specification of its variant forms. All the variables investigated here are binary, in that they involve an alternation between two forms. Variants of binary variables are implicitly assumed to be *discrete*, which means that they have relatively clear-cut boundaries, as defined in auditory, articulatory or acoustic terms (cf. Crystal 2003: 143). Consonants typically have discrete variants, while vowels are more continuous in nature and tend to have variants that range along a phonetic

continuum. Continuous variables are usually more complicated to analyse, as it may be difficult to establish the boundaries between variants (cf. Chambers and Trudgill 1998: 52). For analytical convenience, however, the researcher will have to impose some classificatory borders and a system for categorising intermediate forms. Whatever the nature of the variable, it is important to establish classificatory criteria that can be used in a consistent way. Even if the division of a continuum is inevitably arbitrary, it does not matter as long as the division is the same for all speakers.

The variables studied here can all be said to have two variants that are fairly easily distinguished. Three of the variables have variants at phoneme-level. R-sandhi involves the presence versus absence of a sound, while the other variables involve the alternation between two sounds. Even the vocalic variables have variants that are separated by a certain amount of phonetic distance, and on the whole relatively straightforward to identify. The six variables are also typically described in binary terms in the literature.⁵¹ Moreover, a binary classification can be made on the basis of a correlation with the social evaluation of the variants. Most of the variables represent a dichotomy where one variant is “prestigious” or “conservative” and the other is “vernacular” or “progressive”. The classification of variants is discussed in more detail in chapter 6.

5.2.2.2 Conditioning factors

There are two kinds of conditioning factors that have to be considered in the investigation of phonological variation. The first involves the range of a given variable. In order to identify the relevant variable sites in the material, one has to establish which words or contexts are potentially subject to the variation and which are not. Two of the variables studied here, CURE lowering and GOAT allophony, involve lexical sets (cf. Wells 1982: 122-124), and the relevant words can be identified with reference to their citation form. Smoothing and yod coalescence involve specific combinations of segments at word-internal level, while r-sandhi and t-voicing typically occur across word boundaries. R-sandhi is conditioned by the quality of the preceding vowel, and by the following context, in that it does not occur before a pause. T-voicing has a rather complex phonological conditioning which involves both vowel quantity and prosodic features (see further chapter 6).

⁵¹ The only exception is smoothing, which is often described as having three variants: one triphthongal realisation and two “smoothed” variants that represent different degrees of reduction, viz. diphthongisation to [aə] and monophthongisation to [a:]. In the analysis of smoothing I have collapsed the two reduced forms into one variant (cf. 4.4.5 and 6.5.2).

The second type of conditioning factors are those that affect the relative frequency of the variants. Linguistic units often display systematic variation between social groups, between different styles, and in different linguistic environments. A main aim of the analysis of linguistic variables is to uncover the intra- and extra-linguistic factors that favour and disfavour variants. In the present study, the speakers have been grouped according to two social characteristics, namely gender and channel affiliation, in order to test the effect of these factors on usage patterns, and illuminate the social status of the variants. In addition, the distribution of variants has been correlated with intra-linguistic aspects, such as phonetic environment, stress patterns, morphological context and word frequency. Where possible, I have compared variant frequency in reading style and interview style. The selection of relevant constraints for analysis is based on observations of usage, previous descriptions, and general principles of linguistic variation, as well as practical considerations.

5.2.2.3 Transcription and identification

Once the spoken data had been collected, the entire corpus was transcribed orthographically, using a dictaphone transcriber. An orthographic transcription facilitates the identification of relevant linguistic features, as lexical, morphological or phonological elements can be searched more efficiently. After having transcribed the corpus, all potential variable sites were identified and marked. In some cases, the final identification had to be done parallel to listening through the material, since variables may be conditioned by factors that are not reflected in the orthography (cf. 5.2.2.2). In the case of r-sandhi and t-voicing, for example, deciding what counts as an instance of the variable partly depends on whether the token is followed by a pause or a glottal onset of the next vowel.

When all the relevant token sites were identified, an auditory analysis was carried out by the researcher. Excerpts of the data were recorded to digital Wav-files and converted to mp3-files which were sent to external listeners. Parts of the material were also transferred to the Praat computer programme, which allows you to edit the recordings by extracting and moving sound segments, and thus facilitate the analysis. In order to secure a stable analysis, certain sounds or sound sequences were selected and isolated for detailed or repeated listening (see further below).

5.2.2.4 Auditory analysis

No one can hope to be a successful linguist unless he has a good ear.

Daniel Jones (*An outline of English Phonetics*, 1932)

The corpus has been analysed auditorily, which involves listening to the recorded data and categorising tokens on the basis of perceptual judgement. An auditory technique was preferred for practical as well as linguistic reasons. The variants are on the whole fairly straightforward to discriminate perceptually. In addition, the corpus is relatively large, and it would be exceedingly time-consuming to carry out an instrumental analysis of the entire material. Also, since the investigation has a sociolinguistic perspective, it is appropriate to rely on the same input that is available to the speakers and listeners in the speech community.

An alternative approach to analysing speech is acoustic measurement, which is commonly used in variationist studies. This technique involves examining the acoustic signals of speech sounds and translating these into a visual representation. The obvious advantage of acoustic analysis is the objectivity of instrumental measurement. Auditory perception always involves a degree of subjectivity. Furthermore, instrumental methods provide a greater level of detail, and may reveal distinctions that are undetectable to the ear. The value of instrumental phonetics has been demonstrated in a number of studies (cf. e.g. Labov 1994, Docherty and Foulkes 1999).

Acoustic analysis is, however, not practical or beneficial for all studies of speech variation. The technique is used mainly for measuring vowel variation, and acoustic analyses of consonantal features are rare (but see Docherty and Foulkes 1999). Holmes (1995) notes in her study of t-glottalling in New Zealand English, that instrumental techniques could not assist in distinguishing the variants (1995: 460). Moreover, virtually all instrumental studies have focused on comparisons of formant frequency values, which is not entirely unproblematic (cf. Milroy and Gordon 2003: 148-150). Because of sex-based differences in the vocal tract, formant values are relative, and comparative analyses require inter-speaker normalisation. Moreover, F_1 and F_2 values represent only part of the acoustic information that determines vowel quality, while vowel perception seems to depend on a wide range of cues in the speech signal (cf. e.g. Ohala 1992). With an auditory analysis, vowel quality is estimated on the basis of the entire speech signal, including those that presumably are relevant for speakers' perception and interpretation of sounds. As Lindblom (1980) notes, acoustic measurements are useful only to the extent that they reflect linguistically relevant factors.

Lastly, instrumental techniques are generally more time-consuming than auditory analysis, and acoustic studies typically analyse fewer tokens.

The greatest concern with auditory analysis is the matter of *reliability*, which in this context refers to the consistency of the analysis and, by extension, the stability of the perception. The reliability problem is summarised by Kerswill and Wright (1990: 258):

The main question here is whether or not a transcriber is consistent. Will he or she transcribe the same token the same way twice? And does that transcriber have a tendency to “drift” in his or her judgments over a period of time?

One way to reduce the impact of subjectivity is to analyse a large number of items, which will decrease the statistical effect of any “mislabelled” tokens. Another important measure is to test the consistency of the analysis by letting a second person analyse a sample of the data, and comparing the two transcriptions. For the present study, as a supplement to my own analysis, a trained phonetician analysed a portion of the corpus, with data from 22 of the speakers.⁵² The two analyses differed in only 3.6% of the observations (16 out of 440 tokens), which suggests that the classification of variants is reliable, and the perception stable. The same sample was also re-analysed by the author, in addition to a second sample of 513 tokens. The disagreement rate was again very low: 3.2% and 3.9%, respectively. The discrepancies in many cases involved the same items for both analysts, and some variables proved more problematic than others. Most of the disagreements involved t-voicing, smoothing and yod coalescence.⁵³ However, the disagreements went “both ways”, in that no particular variants were “favoured” by either of the transcribers or at either of the two times. This shows that there is no particular bias in the analysis, and the variability is not likely to skew the results.

A comment should also be made about the question of *validity* in regard to auditory analysis. It can be questioned whether an auditory transcription reflects articulatory facts or auditory impressions, since the two may not always coincide. Experimental evidence suggests that articulatory gestures do not necessarily have auditory (or acoustic) consequences (cf. Kerswill 1985, Kerswill and Wright 1990). This is particularly problematic in the case of gradual, or continuous, variation. Perception tends to operate categorically, and articulatory gradualness or overlapping will be disregarded. The acoustic formant model can provide more continuous data, but has the same inherent uncertainties, as there is no absolute

⁵² I thank Thor Sigurd Nilsen for kindly performing this analysis and providing detailed feedback.

⁵³ T-voicing: 21 tokens, smoothing: 11 tokens, yod coalescence: 8 tokens, GOAT allophony: 7 tokens, CURE lowering: 3 tokens, r-sandhi: 1 token.

correspondence between acoustic signals and articulatory features. It is thus hard to see how this problem can be avoided without using electropalatographic techniques.

As mentioned above, the variables investigated here are fairly easily approached auditorily. With the possible exception of t-voicing, they are all presumed to be above the perceptual threshold for “ordinary” hearers, which is evidenced by various references in the literature. Smoothing, r-sandhi and yod coalescence have been mentioned, commented on or parodied in non-linguistic sources (cf. 4.4), and according to Wells (1997a: 25), speakers are often conscious of the difference between [ɒʊ] and “ordinary” [əʊ] in *GOAL*.

Most of the recorded material has been listened to several times, and certain tokens which proved difficult to classify, many times. In some cases items were extracted and grouped, and listened to in immediate succession, which enabled a more direct comparison of sounds. In order to increase the stability of the analysis, it sometimes proved useful to compare sounds to selected items which were deemed “prototypical” of one variant. For example, if a speaker produced a very clear smoothed [aə], with no trace of the [ɪ] or [ʊ] element, this item was isolated and used as a reference point with which other potential [aə]-items from the same speaker could be compared. A few tokens have been excluded from the analysis because they were not audible enough for reliable classification, either because they were pronounced too quickly or because of background noise. This applies, however, to an insignificant amount of items, and does not affect the overall results. A more detailed account of the analysis of each variable is given in chapter 6.

Following the auditory coding of variants, data from each of the 30 speakers were extracted, sorted and quantified.

5.2.3 Quantification

5.2.3.1 Quantifying the results

Methods of quantification are fundamental for linguistic research that studies variability. The relative frequency of alternate variants can only be determined in quantitative terms.

Measuring the usage of a variable simply involves counting the number of times each of the variants is used, and calculating the percentage scores.

With the exception of *CURE* lowering and *GOAT* allophony, which have been analysed primarily at a qualitative and lexical level, all the variables have been quantified at inter-speaker level. Percentage scores are given for individual speakers as well as for groups.

Group scores are presented in terms of *mean* and *standard deviation*. The mean is simply the average score, calculated by adding the scores for each member of the group and dividing by the total number of members in that group. The standard deviation is an indication of how scores are distributed about the mean. A large standard deviation means that the scores are widely scattered or that there are a few scores which are very far from the mean.

In the analysis of linguistic constraints and stylistic variation, scores have been computed for the corpus as a whole, as inter-speaker characteristics are assumed not to affect the variation. This secures a large number of tokens for each factor, which increases the reliability of the results.

When a linguistic corpus has been analysed and quantified, we can study the distributional patterns in order to determine whether observed differences reflect real usage patterns in the overall population, or merely random variation in the sample. This can be done by using tests for statistical significance, which will be discussed in the next section.

5.2.3.2 Statistical analysis

A main goal of variationist research is to find relations between variables. Statistics provides standardised procedures to quantitatively estimate and evaluate such relations. The standard measure for statistical significance level is the p-value, which indicates the probability that the observed relationship occurred by coincidence (the null hypothesis). This probability decreases the larger the sample size. It is, however, not sufficient that a distributional difference is statistically significant. The researcher should also present well-founded hypotheses that attempt to explain the relation between variables. By using statistical procedures, we do not *prove* the research hypothesis directly, but we can strengthen it indirectly by rejecting the null hypothesis.

Most of the quantified data have been analysed statistically to test the significance of the findings. For this purpose I have used the statistics programme SPSS (Statistical Package for the Social Sciences).⁵⁴ The quantified results for r-sandhi, t-voicing, smoothing and yod coalescence, which all are represented by fairly large numbers of tokens, have been subjected to statistical analysis. The use of pronunciation variants has been correlated both with linguistic and extra-linguistic factors. The data for CURE lowering and GOAT allophony have not been quantified at inter- and intra-speaker level, and statistical analysis has therefore not been relevant here.

⁵⁴ I wish to thank Kyrre Moe at *Norsk samfunnsvitenskapelige datatjeneste* (the Norwegian Social Science Data Services) for assistance with the statistical analysis.

In the analysis of the effects of linguistic constraints and style shifts, the chi-square test has been used. This analysis is performed at the level of tokens, and quantifications are computed for the whole corpus, as if it were a single speaker. This gives the advantage of large numbers of tokens, sometimes several thousands, which increases the chances of obtaining reliable statistical findings. While this method will reveal distributional patterns in the material, it has the risk of skewing results in cases where speaker contributions are very unequal with regard to the number of tokens and type of speech. In the present study, however, the informants are extremely homogeneous in the sense that they all contribute with the same amount of text, and they are all in the same setting and talk about the same type of topics. It is therefore fairly unproblematic to compute scores for the whole corpus in the analysis of intra-linguistic variation.

The chi-square test is a non-parametric procedure widely used to test the independence or interdependence of the distribution of variables. The method is suitable if we want to compare two variables measured in the same sample, where the variables are categorical in nature (e.g. presence vs. absence of /r/ by stressed vs. unstressed vowel). The chi-square test is not recommended when the numbers in the contingency table are very small. If the expected frequency in any cell is below 5, the test is not reliable (cf. Butler 1985: 122, Woods et al. 1986: 144-145). This is not a problem in the present study, where all expected frequencies are well above 5. The chi-square test further requires that the data are reported in raw frequencies, and that the observations are probabilistically independent of each other (i.e. each entry in the contingency table is independent of the other entries).

Chi-square works by comparing the observed frequency of a variant with the frequency one would expect if the contextual factor had no effect on the distribution of the variant. This is calculated by subtracting the expected frequency from the observed frequency for each cell, squaring the result, and dividing by the expected frequency. The final chi-square value (χ^2) is obtained by summing these values for all cells. The null hypothesis is then that there is no difference in distribution between the observed and expected frequencies. A small chi-square value indicates that the null hypothesis is true, while a large value indicates that the observed frequencies differ significantly from the expected frequencies. The significance level (p-value) of the chi-square value is assessed with reference to the number of degrees of freedom (df), which is calculated by multiplying the number of rows minus one by the number of columns minus one. If the p-value is low, the chances are correspondingly low that the observed difference is only random. A significance level of 0.05 or below is generally taken as sufficient to reject the null hypothesis (cf. e.g. Hudson 1996: 153).

The material has also been analysed at the level of individuals to investigate any correlations with social factors such as gender and channel affiliation. The frequencies used in the inter-speaker analyses are the percentage scores for each informant. Percentages are preferred here because the total number of tokens is not the same for all speakers, and the raw figures would therefore not give an accurate representation of the relative frequency of a variant. Reducing the figures to percentages thus makes it easier to compare speaker scores in a meaningful way, as the contributions of all informants are given equal weight. The basis for analysis is then no longer large numbers of tokens, but the 30 speakers. This reduces any problem of unequal speaker contributions and potential skewness. On the other hand, it becomes more difficult to obtain significant results because of the relatively small number of informants.

In the analysis of inter-speaker variation I have used analysis of variance, or ANOVA, to compare speaker groups. ANOVA is used to uncover the effect of a categorical independent variable (e.g. gender) on a quantitative dependent variable (e.g. linking /r/). It is a parametric statistical procedure which tests for significant differences between groups by comparing the variation between the groups with the variation found within the groups. In ANOVA the dependent variable can be quantified on a ratio scale (such as percentages), and the test can analyse differences between more than two groups. ANOVA was therefore judged to be a suitable statistical method in the analysis of inter-speaker variation, since I am using percentage scores to represent individual results, and the analysis includes a comparison of the three news channels.

Parametric tests are based on assumptions about the distribution of the underlying population from which the sample was taken. The ANOVA test assumes that the dependent variable is normally distributed and that the amount of variance is the same across the groups.⁵⁵ Both these assumptions are met in the present data.⁵⁶

The key statistic in ANOVA is the F test of difference of group means (or averages), which tests if the means are different enough not to have occurred by chance. The null hypothesis is that group means do not differ at all. ANOVA works by calculating mean square values which represent the variance of the between-group and within-group components of the data. Every single score within a group is measured against the group average, the

⁵⁵ The reason for this assumption is that the denominator of the F-ratio is the within-group mean square, which is the average of group variances. If groups differ widely in variances, this average is a poor measure.

⁵⁶ Normality of distribution has been tested by comparing the skewness and standard error for each variable. A skewness value of more than twice its standard error is taken to indicate a departure from normality. Homogeneity of variance has been tested with the Levene test of homogeneity which is computed by SPSS.

deviation is squared, and the deviations in each group are summarised. The sum of squares is divided by the degrees of freedom to give the mean square values. The F value is then computed by dividing the mean square for between groups by the mean square for within groups. If the computed F score is significantly greater than 1, there is more variation between groups than within groups, from which we infer that the independent variable has an effect on the dependent variable. In addition to the F-value, the significance level (Sig.) is also computed in SPSS and given in the ANOVA table. Again, a level of less than 0.05 indicates significantly different group means.

The quantified data have further been analysed for correlation, to uncover any significant covariation between linguistic variables (see 7.2). This has been done by comparing the individual percentage scores for two variables (e.g. t-voicing and yod coalescence) to see if high (low) values on one variable tend to be associated with high (low) values on the other. For this purpose I have used the Pearson product-moment correlation, also known as Pearson's *r* correlation. Pearson's *r* is a bivariate statistic which measures the degree of linear association between two interval or ratio variables. The null hypothesis is then that there is no association between the two variables.

The Pearson's *r* correlation is the ratio of the cross-product of two variables divided by the square root of the sum of squares associated with those variables. It is thus the ratio of the covariance between two variables and their two standard deviations. As the ANOVA test, it is a parametric procedure which assumes underlying normal distribution. The value obtained by the Pearson test, the correlation coefficient, ranges from +1 to -1, and thus indicates both the direction and the strength of the relationship between two variables. A correlation of +1 indicates a perfect positive correlation, while -1 indicates a perfect negative (inverse) correlation. If the correlation is 0 or very close to 0, there is no association between the two variables. A strong positive relationship between variables may suggest that they mirror the same general process of change, that they are products of the same underlying phonological process, or that they have similar social or evaluative status.

6 Analyses and results

Until one listens to how the message is being spoken, rather than to what is being said, it is perfectly reasonable to have a very idealistic and starry-eyed view of how English is pronounced by public speakers.

Brown (1990: 3)

6.0 Introduction

This chapter presents the results of the analysis of the six variables described in section 4.4. The presentation includes brief discussions of the auditory analysis and discrimination of variants, and the criteria for selecting relevant target words or contexts. CURE lowering and GOAT allophony have been analysed mainly from a qualitative and lexical perspective. For all the other variables, quantitative speaker scores have been calculated and tested for statistical significance. Quantified results are given for the sample as a whole as well as for each individual speaker. Moreover, comparisons are made between male and female presenters, and between the three channels. In the case of r-sandhi and t-voicing, which are the most frequent variables, usage levels are correlated with formality and degree of attention, by comparing reading style and interview style. Since the present study focuses on a socially very homogeneous group, the variables' correlation with linguistic factors constitutes an important part of the analysis. According to Hudson (1996: 181), variations in the linguistic context are normally independent of social variations. The intra-linguistic analyses are therefore performed for the sample as a whole. Section 6.7 investigates the level of covariation between the linguistic variables.

The findings will be compared with previous descriptions of and claims about RP, as outlined in 4.4. On some occasions, reference will be made to studies of other varieties of English (t-voicing, for example, has been extensively studied for American English, but is generally not mentioned in connection with RP).

On the basis of previous writings and widely established principles of linguistic variation and change, the following general hypotheses can be set up⁵⁷:

⁵⁷ Hypotheses relating to intra-linguistic factors will have to be specified for each variable, and are therefore presented in detail in the appropriate sections.

- Usage levels for GOAT allophony, intrusive /r/, t-voicing and yod coalescence will be higher for male speakers than for females.
- Usage levels for linking /r/ and smoothing will be higher for females than for males.
- Sky News presenters will have the highest scores for GOAT allophony, intrusive /r/, t-voicing and yod coalescence, while BBC World presenters will have the lowest.
- The frequency of intrusive /r/ and voiced /t/ will increase in interview style compared to reading style.
- The frequency of r-sandhi, t-voicing and smoothing will correlate with speech tempo.
- Speakers with above average scores for t-voicing will also have above average scores for intrusive /r/ and yod coalescence.

These aspects will be explored in the following sections, and will be returned to in the concluding discussion in chapter 7.

6.1 CURE lowering

6.1.1 Introduction

The present section outlines the results of the analysis of CURE lowering, which involves the use of /ʊə/ versus /ɔ:/. This variation is indicative of an ongoing change in RP, whereby /ʊə/ is disappearing and gradually being replaced by /ɔ:/ (cf. 4.4.1). The analysis focuses on the lexical incidence of the two vowels, and explores the phonetic constraints on the lowering to /ɔ:/.

6.1.2 The auditory analysis of CURE lowering

CURE lowering is analysed as a phonological variable which involves the phonemic alternation /ʊə/ ~ /ɔ:/. The monophthongal phoneme is a back open-mid rounded [ɔ:] with little qualitative variation. The /ʊə/ phoneme incorporates all diphthongal realisations with a central end-point or off-glide and a close back or central starting point [ʊə, ʊ^ə ~ ʊə, ʊ^ə], as well as monophthongised realisations ranging from [ʊ:] to [ɜ:]. Although the exact phonetic

quality varies somewhat, especially for /ʊə/, the two phonemes remain quite distinct with most speakers. No realisations with /ɜ:/ were recorded (cf. 4.4.1).

The fact that /ʊə/ often has a centralised starting-point, especially after /j/, made it relatively straightforward to separate it from /ɔ:/. After consonants other than /j/, some speakers occasionally had a retracted starting-point in /ʊə/, approaching [o]. Such realisations were more problematic to distinguish from /ɔ:/, especially in cases where the vowel was followed by a prevocalic /r/, which obscured the [ə]-offglide. The main articulatory and auditory difference between /ʊə/ and /ɔ:/ was then the degree of openness, which could be used to distinguish the two. In addition, CURE words were compared with words of other lexical sets. In general, speakers with a back quality in CURE also had a back vowel in GOOSE, and the /u:/ vowel could therefore serve as a reference for /ʊə/. At the same time, CURE words with potential /ɔ:/ were compared with words of the lexical sets FORCE/NORTH/THOUGHT, which all have /ɔ:/ in their citation form and thus could provide a qualitative reference for the monophthongal variant.

6.1.3 Identifying CURE words

The CURE set comprises those words “whose citation form contains the stressed vowel /ʊə/ in conservative RP” (Wells 1982: 162), such as *sure*, *tour*, *pure*, *mature*, *furious*. /ʊə/ can also occur in unstressed syllables, as in *contour* /'kɒntʊə/, *tenure* /'tɛnjʊə/, *neurotic* /nɜʊə'rɒtɪk/, but such items have not been included in the analysis. Also, the word *mourn* and its derivatives *mourner* and *mourning*, which some speakers pronounce with /ʊə/, have been excluded because they belong to the lexical set FORCE and are thus primarily realised with /ɔ:/ (cf. Wells 1982: 162). Lastly, there are instances where /ʊə/ occurs as a result of the contraction of the disyllabic sequence /u:ə/ in words like *fewer*, *doer*, *jewel* (cf. Giegerich 1992: 65). Such cases have been excluded from the analysis since they do not represent CURE and because /u:ə/ can never be reduced to /ɔ:/.⁵⁸

One of the main problems involved in the investigation of the CURE vowel is that the number of CURE words is not very large, and some of the lexical items occur only once or

⁵⁸ Jones (1954: 8) claims that *skewer* can be pronounced /skjɔ:/. Such a variant is however not mentioned in any other sources.

twice in the corpus. The foundation for generalisations is therefore somewhat limited. In spite of this limitation, some striking distributional patterns can be observed in the use of /ʊə/ and /ɔ:/.

6.1.4 Results

The analysis of CURE lowering shows that both /ʊə/ and /ɔ:/ are used in CURE, but the centring diphthong is three times as frequent as the monophthong. The material contains a total of 1263 tokens, and two thirds of these are pronounced with /ʊə/, which suggests that the diphthong is still far away from being replaced by /ɔ:/. The frequency of the two vowels is shown in Table 6.1.1 below.

Table 6.1.1. CURE lowering: total scores

	N	%
/ʊə/	952	75.4
/ɔ:/	311	24.6
Total	1263	100

These figures in themselves are, however, of little interest, as they mainly reflect a lexical skewness in the data, where some words which invariably have /ʊə/ are hugely overrepresented. The further analysis will therefore focus on the lexical distribution of the two phonemes.

All the speakers use both /ʊə/ and /ɔ:/ in CURE, and there is very little inter-speaker variation. There are no notable differences between male and female presenters or between the various channels. The two vowels seem, however, to have very different lexical distributions. Table 6.1.2 shows all the CURE words that appear in the corpus and the number of times they occur with /ʊə/ and /ɔ:/.

Table 6.1.2. The lexical distribution of /ʊə/ and /ɔ:/ in CURE

word	/ʊə/	/ɔ:/	word	/ʊə/	/ɔ:/
<i>bourse</i>	-	1	<i>rural</i>	3	-
<i>bureau</i>	2	-	<i>secure</i>	17	1
<i>bureaucratic</i>	4	-	<i>securing</i>	1	-
<i>cure</i>	3	-	<i>security</i>	239	-
<i>incurable</i>	1	-	<i>sure</i>	4	28
<i>during</i>	171	-	<i>surely</i>	3	10
<i>endure</i>	-	1	<i>assure</i>	3	3
<i>enduring</i>	2	-	<i>assuring</i>	1	-
<i>Euro</i>	138	-	<i>assurance</i>	1	1
<i>Europe</i>	126	-	<i>reassure</i>	1	7
<i>European</i>	96	-	<i>ensure</i>	6	10
<i>fury</i>	2	-	<i>ensuring</i>	-	1
<i>furious</i>	6	-	<i>insure</i>	-	1
<i>infuriated</i>	2	-	<i>insurer</i>	2	-
<i>jury</i>	12	2	<i>insurance</i>	8	4
<i>lure (/lj-/) </i>	1	-	<i>tour</i>	20	31
<i>luring</i>	1	-	<i>touring</i>	1	2
<i>moor</i>	1	3	<i>tourism</i>	12	3
<i>poor</i>	-	20	<i>tourist</i>	49	17
<i>poorer/-est</i>	-	7	<i>tournament</i>	8	28
<i>poorly</i>	-	1	<i>Uruguay</i>	1	-
<i>purely</i>	3	-	<i>your</i>	-	58
<i>purify</i>	1	-	<i>you're</i>	-	71

It is evident from Table 6.1.2 that /ʊə/ is more frequently used than /ɔ:/ in CURE, and that the diphthong has the widest lexical distribution. There are in total 46 different CURE words, including all the morphological derivatives, and 22 of these occur only with /ʊə/ while as few as 9 occur exclusively with /ɔ:/. The remaining 15 words are pronounced with both vowels. The greatest amount of overlapping between /ʊə/ and /ɔ:/ is found with the words *sure* and *tour* and their respective morphological derivatives, which suggests that these items are still in the process of becoming lowered.

At the level of the individual speakers, the lexical distribution of the two phonemes is very consistent, in that every word occurs with the same vowel on all occasions. The only exceptions are speakers 21, 25 and 27, who fluctuate between the two vowels in the items *tournament*, *tourism* and *tour*, respectively.

Previous accounts of CURE lowering suggest that this change is both lexically and phonetically conditioned. Lowering is said to predominate in familiar or high-frequent words

while the diphthong is retained in rare or foreign words; and lowering is resisted in contexts where the vowel is preceded by a consonant plus /j/ (cf. 4.4.1). Nolan (1999) seems to regard CURE lowering as purely a matter of lexical variability, with “speakers preferring a different phoneme in certain words and classes of words” (1999: 79). The evidence from the present material clearly shows that the two vowels largely have separate distributions, with /ʊə/ dominating in certain words and groups of words, while /ɔ:/ is preferred in others, and only a few items show an equal use of both phonemes. In other words, this is seemingly a case of lexical diffusion (cf. 3.1.2). A closer inspection of the distribution of /ʊə/ and /ɔ:/ suggests, however, that lowering is not implemented by lexical diffusion, at least not in the strict sense of the term. The distribution is not based on lexical identity or word frequency, nor is it completely random. It seems, however, to be largely conditioned by the phonetic composition of the words. The use of /ʊə/ versus /ɔ:/ correlates fairly closely with the phonetic environment, and appears to follow two basic tendencies. The first, and most obvious, is that /ʊə/ occurs in the environment following a consonant plus /j/, as in *cure, fury, security*. 16 of the words in Table 6.1.2 have C/j/ immediately before the relevant vowel,⁵⁹ and 15 of these are realised with /ʊə/. This result is in accord with observations by Wells (1982: 287) and Tollfree (1999: 169). The tendency to retain /ʊə/ after /j/ can be explained articulatorily, with reference to the fact that both are “high” sounds. The palatal /j/ is then more likely to combine with /ʊə/, which has a close starting-point, than with the open-mid /ɔ:/.

As for words with initial /j/, which Wells claims are frequently pronounced with /ɔ:/, there is a sharp division between *your/you're*, which invariably occur with /ɔ:/, and all the others (*Euro, Europe, European, Uruguay*), which always have /ʊə/. This contrast may be explained by the fact that *your*, unlike the other items, historically had no /j/ (Wells 1982: 164). A more obvious factor, however, is the difference in phonetic context: in *your/you're* the CURE vowel is final, while in *Euro, Europe*, etc. it is followed by a prevocalic /r/. The latter context seems to disfavour lowering, and this constitutes the second principle that governs the distribution of /ʊə/ versus /ɔ:/. In words where the CURE vowel is not preceded by a consonant plus /j/, lowering is common if the vowel is final or followed by a consonant other than /r/, as in *poor, sure, bourse, tournament*. If, however, the vowel is followed by prevocalic /r/, as in *jury, rural, tourist*, it is typically realised as /ʊə/. Although this tendency is not as strong as the first, it explains why /ʊə/ dominates in words like *insurance* and *tourist*,

⁵⁹ These include *during, endure* and *enduring*, which in most cases are pronounced with yod coalescence, giving /dʒ/ rather than /dj/, as well as *lure*, where the /j/ is optional.

while *sure* and *tour* for most speakers have /ɔ:/. There are some striking examples of this pattern at intra-speaker level, for instance with speaker 14, who says *insure* with /ɔ:/ but *insurer* with /ʊə/, and speaker 20, who has /ɔ:/ in *secure* but /ʊə/ in *security*.

The reason why /ʊə/ tends to be retained before /r/ is presumably that the presence of /r/ weakens the [ə]-element of the diphthong, making the vowel more similar to the /u:/ of GOOSE. The origin of the CURE vowel is the same as that of GOOSE⁶⁰ (Wells 1982: 163). The centring diphthong only arose as a result of the r-dropping which took place in RP: when non-prevocalic /r/ was lost, /ʊə/ replaced /u:/ + /r/ (cf. Wells 1982: 218-222). In words like *Europe*, *rural* and *tourist*, where the /r/ is present, the vowel tends to be realised as a monophthong by the speakers in the present study. This monophthong may be interpreted as a variant of /u:/, and /u:/ will never be lowered to /ɔ:/. The vowel in *rural* is then not so different from the one in *ruble*, while word-finally, the [ə]-offglide in /ʊə/ is much more prominent, and distinguishes e.g. *cure* /kʃʊə/ from *cue* /kju:/.

The variability between /ʊə/ and /ɔ:/ in CURE seems, then, to be largely determined by the phonetic context, though not so systematically that it can be reduced to an allophonic split. The main tendency is that /ʊə/ occurs in words where the vowel is preceded by /j/ (except *your/you're*), while /ɔ:/ dominates in words where the vowel is preceded by other consonants. This pattern is so consistent that we can almost talk about two different lexical subsets, CURE and POOR. While lowering to /ɔ:/ is well-established, or well on its way, in most POOR-words, it is still resisted in CURE, where the use of /ɔ:/ can be regarded as innovative or progressive. The second trend is for /ʊə/ to be retained, or replaced by an [u:]-like monophthong, in the environment of a following prevocalic /r/. It remains to be seen whether words like *during*, *jury* and *rural* will eventually lower to /ɔ:/, or whether they will monophthongise to /u:/ and be reanalysed as GOOSE words.

6.1.5 Comparison with pronunciation dictionaries

Many recent statements about CURE lowering, though rarely based on corpus studies, indicate that /ʊə/ is rapidly disappearing from the vowel system of RP as it gives way to /ɔ:/ (cf. 4.4.1). Pronunciation dictionaries, by contrast, are quite conservative in their transcriptions of

⁶⁰ Many modern rhotic accents have /u(:)/ in CURE, and thus phonemic identification with GOOSE.

CURE words, but seem to be more in line with the findings recorded here. Current dictionaries give both /ʊə/ and /ɔ:/ as possible variants in all CURE words (except rarities like *dour* and *Ruhr*). The diphthong is however listed as the first choice in the majority of the cases. A brief look at the latest editions of LPD, EPD and ODP shows that they all give /ʊə/ as the main variant in CURE words where the vowel is preceded by a /j/ (*cure, during, Europe, furious, urine, etc.*). The only exception is *your*, which has /ɔ:/ first in all dictionaries. LPD and EPD prefer /ɔ:/ also in *you're*, but ODP distinguishes between the two and has /ʊə/ as first choice in the latter. All the dictionaries give the monophthong as the main pronunciation in *poor*. ODP prefers the diphthong in all other CURE words, while LPD has /ɔ:/ also in *sure* and *assure/assurance*. EPD has /ɔ:/ in *moor* and *sure*, but distinguishes between *assure* and *assurance* and for some reason has /ɔ:/ as first choice only in the latter. All other CURE words without preceding /j/ (*tour, insure, bourse, etc.*) are listed with /ʊə/ first in all the dictionaries.

6.1.6 Summary

CURE lowering is a case of phonemic alternation which potentially may affect the phoneme inventory of RP. The diphthong /ʊə/ has long been predicted to be replaced by /ɔ:/. The findings reported here, however, show that the newsreaders prefer /ʊə/ in the majority of CURE words, and that lowering is established only in a few items. This could imply a certain conservatism on the part of the newsreaders, or it could mean that CURE lowering is not as widespread as many previous descriptions have suggested. The recurrent statement that /ʊə/ is disappearing from the phoneme inventory of RP seems grossly exaggerated judging from the evidence provided by the broadcast data, which shows that all 30 speakers have /ʊə/ as an active member of their phonological system, and lowering is almost categorically resisted in certain phonetic environments. The variation in the incidence of /ʊə/ may of course in time entail a complete loss of the diphthong, but despite its reduced distribution, there is apparently a long way to go before the merger of /ʊə/ with /ɔ:/ is completed. CURE lowering is thus a change still very much in progress, and may not even affect all the members of the lexical set.

6.2 GOAT allophony

6.2.1 Introduction

The following sections present the results of the analysis of GOAT allophony. This variable, which is sometimes described as a typical Estuary English feature, involves the use of [ɒʊ] as an allophone of /əʊ/ in GOAT words before non-prevocalic /l/. The lexical set GOAT is defined as “comprising those words whose citation form has the stressed vowel /əʊ/ in [traditional] RP” (Wells 1982: 146), such as *boat, road, go, know, ocean*, etc. The subset *GOAL*, where the potential allophony occurs, comprises all words where stressed /əʊ/ is followed by final or preconsonantal /l/, e.g. *poll, hole, gold, soldier*, etc.

6.2.2 The auditory analysis of GOAT allophony

GOAT allophony involves an allophonic alternation in /əʊ/, which has the variants [əʊ] and [ɒʊ]. The former typically has a central unrounded onset, [ɜ ~ ə], while the latter is produced with a back rounded starting point with a quality ranging from [ɒ] to [ɔ]. GOAT allophony is phonetically conditioned, in that the [ɒʊ] variant only occurs before non-prevocalic /l/. Moreover, it may be articulatorily motivated: the /l/ is velarised in non-prevocalic positions and therefore likely to have a backing effect on the preceding vowel (cf. Milroy and Gordon 2003: 155). Although it is primarily the starting point of the diphthong which is retracted, from [ə] to [ɒ], there is also generally a change in the quality of the second element. The [əʊ] variant usually has an advanced or centralised, often unrounded, end point, [ʊ̟ ~ ɥ̟],⁶¹ while the offset of the [ɒʊ] allophone is typically fully back and rounded, [u ~ ʊ ~ o]. The [ɒʊ] diphthong is thus generally more back throughout the articulation.

On the whole, then, the two variants are qualitatively quite distinct. However, as with all vocalic variables, the variants are situated along a phonetic continuum, which leaves room for a number of intermediate realisations. Occasionally, therefore, individual sounds were difficult to classify, and the exact quality also varied between speakers. These factors did however not present any major problems, as the main purpose of the analysis was to listen for

⁶¹ In the southeast of England there is a growing trend towards GOAT fronting, where the offset may be extremely advanced: [ɤ ~ ɪ] (cf. e.g. Williams and Kerswill 1999, Altendorf 2003).

a systematic qualitative difference between /əʊ/ before [t] and /əʊ/ in other contexts, in the speech of individual presenters. Most of the speakers had a consistent contrast with qualitatively distinct variants. In cases where the presence/absence of GOAT allophony was difficult to establish, the relevant pre-/l/ items were isolated and compared directly to a selection of other GOAT words from the same speakers. In some cases, the presenters produced phrases where a *GOAT* word and a *GOAL* word occurred in immediate succession, such as *remote control*,⁶² *own goal*, *Golden Globe*, *no holds barred*, *its own poll*. Such phrases were particularly convenient for comparison and facilitated the identification of GOAT allophony.

6.2.3 Results

6.2.3.1 General

The results of the analysis show an overwhelming preference for GOAT allophony. 24 of the 30 speakers distinguish categorically between *GOAT* and *GOAL* and use the [ɒʊ] variant in pre-/l/ items. Four of the presenters (speakers 4, 7, 9 and 21) have variable allophony, and only two speakers (2 and 28) do not have the [ɒʊ] allophone.

Table 6.2. The use of GOAT allophony

	Categorical allophony	Variable allophony	No allophony
Number of speakers	24	4	2

These findings clearly suggest that GOAT allophony is becoming firmly established in modern RP. In comparison with the standard descriptive works on RP, where GOAT allophony is generally not recognised as an RP feature, it is evident that a change is taking place in the accent from [əʊ] in all GOAT words, to [ɒʊ] in the environment before [t]. Moreover, the allophony is not limited to accents of London or the southeast (cf. 4.4.2). In the study of broadcast speech I frequently observed GOAT allophony not only with the selected RP speakers, but also with speakers of various regional varieties. These observations establish

⁶² This phrase was used particularly often by the Sky News presenters, as they frequently referred to the channel's interactive service which the viewers can access by "pressing the red button on the remote control".

GOAT allophony as a supra-regional feature and justify its inclusion in non-localisable RP (cf. 2.4.7).

The speakers who display variable GOAT allophony fluctuate between [əʊ] and [ɒʊ] in pre-/l/ items, with no particular patterning in the distribution of the variants. The variability may indicate that these speakers are undergoing a change towards categorical GOAT allophony, which will eventually be completed. Alternatively, they may have the [ɒʊ] allophone in their native accent, but are modifying towards traditional RP by occasionally suppressing it.

The *GOAL* variants [əʊ] and [ɒʊ] do not have the same status, in that the former is the traditional RP realisation, while the latter seems to be relatively new in RP and may have been imported from the Popular London variety or Estuary English (cf. 4.4.2). With this in mind, one might expect more men than women to display GOAT allophony, as women tend to prefer “high-status” variants (c.f. 3.1.4). However, the present study revealed no gender differences in the use of [ɒʊ] versus [əʊ]. Most of the speakers have GOAT allophony, and those with no or variable allophony represent both male and female presenters. This suggests that GOAT allophony has fairly neutral social and evaluative connotations, as indicated by Tollfree’s (1999) study of London English, which reports *GOAT-GOAL* alternation along the entire accent continuum (cf. 4.4.2). It is worth noting, however, that there is a certain correlation with the formality of the news channels (cf. 5.1.2). All presenters from the “informal” Sky News channel have categorical GOAT allophony, while four of the six speakers who have no or variable allophony represent the more formal BBC World.

6.2.3.2 Phoneme split?

The investigation of GOAT allophony also included items where /əʊ/ is followed by a prevocalic /l/, in order to look for evidence of a phoneme split (cf. 4.4.2). Some *GOAL* words that end in /l/ may receive a suffix beginning with a vowel, making the /l/ prevocalic: *wholly*, *tolling*, *goalie*, *patrolling*. In London English the [ɒʊ] is typically retained in such derivative forms, even though the /l/ is no longer velarised (cf. Wells 1982: 312-313). This has given rise to minimal pairs such as *holy* – *wholly*, and phonemic status to /ɒʊ/. The present material contains very few *GOAL* items with a vowel-initial suffix, and the basis for generalisation is therefore very limited. Of the 24 speakers with systematic GOAT allophony, only 16 produce a total of 27 tokens with prevocalic /l/, including *bowler*, *controlling*, *rolling*, *wholly*. The analysis revealed that seven of the presenters use [əʊ] in these words (speakers 6, 10, 11, 24,

26, 27 and 30), while six show evidence of morphological regularisation and use [ɒʊ] (1, 12, 16, 22, 23 and 25). The remaining three vary between [əʊ] and [ɒʊ] (8, 18 and 29). It appears, then, that six of the speakers at least potentially have a separate phoneme /ɒʊ/ (it should be noted, though, that five of these only produce one token). It may be that these particular presenters are from the southeast of England, where this phoneme split is common, or that [ɒʊ] before prevocalic /l/ is becoming more widespread. The three speakers who use both vowels seem to favour [ɒʊ] after /r/: they all have [ɒʊ] in *controller/controlling* and *rolling*, but use [əʊ] in *wholly* and *bowler*. This may therefore reflect a tendency to retract the vowel after the postalveolar /r/, at least in pre-/l/ items, and not necessarily the beginning of a phoneme split.⁶³

6.2.3 Summary

The analysis of the GOAT vowel has revealed that the great majority of the presenters have a systematic qualitative variation between *GOAT* and *GOAL*, and use [ɒʊ] before non-prevocalic /l/ and [əʊ] in all other words. For most of them, this alternation is a matter of contextual allophony. Some speakers, however, seem to extend the use of [ɒʊ] to items with prevocalic /l/, which may give rise to phonologisation of [ɒʊ], although the number of tokens is too small to draw any firm conclusions. The use of [əʊ] versus [ɒʊ] showed no particular interspeaker patterns, which suggests that neither variant has strong evaluative force. The extensive use of GOAT allophony is a clear indication that this feature should be considered a part of modern non-regional RP, and that it extends beyond Estuary English or London accents.

⁶³ It should be mentioned that five of the six speakers who invariably use [ɒʊ] only have post-/r/ tokens, which means that the same phenomenon may apply to them.

6.3 R-sandhi

6.3.1 Introduction

The present chapter outlines the quantified results of the analysis of r-sandhi, and discusses the findings in light of previous claims about this feature in RP and other varieties.

R-sandhi comprises the two variables linking and intrusive (r), each of which has the variants /r/ and Ø, representing presence and absence of /r/ respectively. All cases of vowel hiatus involving an open or mid vowel in the left-hand environment have been counted as potential r-sandhi sites, except where there is a pause between the two adjacent vowels. What constitutes a ‘pause’ has been defined subjectively as any breach in the speech flow that is of such an extent that the insertion of a sandhi /r/ would be unlikely or unnatural. In “borderline” cases I have tried to be as consistent as possible, while at the same time considering the differences in individual speech tempo (cf. 5.1.3).

The auditory analysis of r-sandhi was unproblematic, as the presence or absence of /r/ is easily identified by the ear. Only a very small number of tokens have been excluded from the analysis because the realisation was too unclear for classification, either because of background noise or rapid articulation.

The /r/ variant is realised as a post-alveolar approximant [ɹ] virtually throughout. In a handful of cases it occurs as [r], but only when the *r* is part of a foreign name (*Igor Ivanov*, *Amir al Sadi*). No examples of r-fronting were recorded. The Ø variant includes zero realisation, i.e. a vocalic glide across the hiatus, and the use of a glottal stop onset to the second vowel. The most common realisation of Ø is by far the glottal stop.⁶⁴

The social, stylistic and linguistic aspects of r-sandhi have been outlined in 4.4.3, and these will form the basis of the discussion of the results. Because of the difference in status between the two variables, linking /r/ is expected to be more prevalent than intrusive /r/. The use of sandhi /r/ is further hypothesised to correlate with a number of linguistic and extra-linguistic factors.

The two variables are analysed separately in sections 6.3.2-6.3.3, which investigate the r-sandhi results in relation to social factors and a number of linguistic constraints. In the

⁶⁴ It is of course possible to analyse [ʔ] as an allophone of /r/ rather than of Ø. Wells and Colson (1971: 94), for example, talk of the “linking glottal stop”, although, as Brown (1988: 145) points out, this “seems something of a misnomer” since “its function is surely to keep the two syllables separate rather than to link them”.

section on style and speech tempo, linking and intrusive (r) are discussed together. The results are summarised in section 6.3.5.

6.3.2 Linking (r)

6.3.2.1 General

In all, 6045 tokens of linking (r) sites have been analysed, which averages just over 200 tokens per speaker. Table 6.3.1 shows the total scores for linking (r) in tokens and percentages.

Table 6.3.1. Linking (r): total scores

	N	%
/r/	3612	59.8
Ø	2433	40.2
Total	6045	100

The quantified results show that, as expected, r-link is favoured over r-suppression, as /r/ is used in almost 60% of the cases. The preference for linking /r/ is however quite small, and is far from the categorical usage often claimed for RP speakers (cf. 4.4.3). The individual frequency scores for the /r/ variant are outlined in Figure 6.3.1 below. The majority of the speakers have scores around 60%. Five presenters have scores below 50% (speakers 4, 13, 19, 29 and 30), and five have scores above 70% (7, 8, 12, 14 and 15), while the average score is 59.3%. Speaker 30 stands out as the only one with a score below 30%. At the other end is speaker 12, who has the highest frequency of linking /r/ usage, with a score of 79.5%.

Although /r/ is the dominant variant, a total percentage score of 59.8 is lower than expected, in view of the overt prestige of linking /r/ and the fact that RP speakers are thought to use linking /r/ almost categorically (cf. 4.4.3). Linking /r/ is optional, but, according to Cruttenden (2001: 289) “generally present” in RP speech. The use of linking /r/ has also been described as categorical or as the norm in several other English accents (see Foulkes and Docherty 1999a). It is therefore somewhat surprising that five speakers favour the Ø-variant in linking environments.

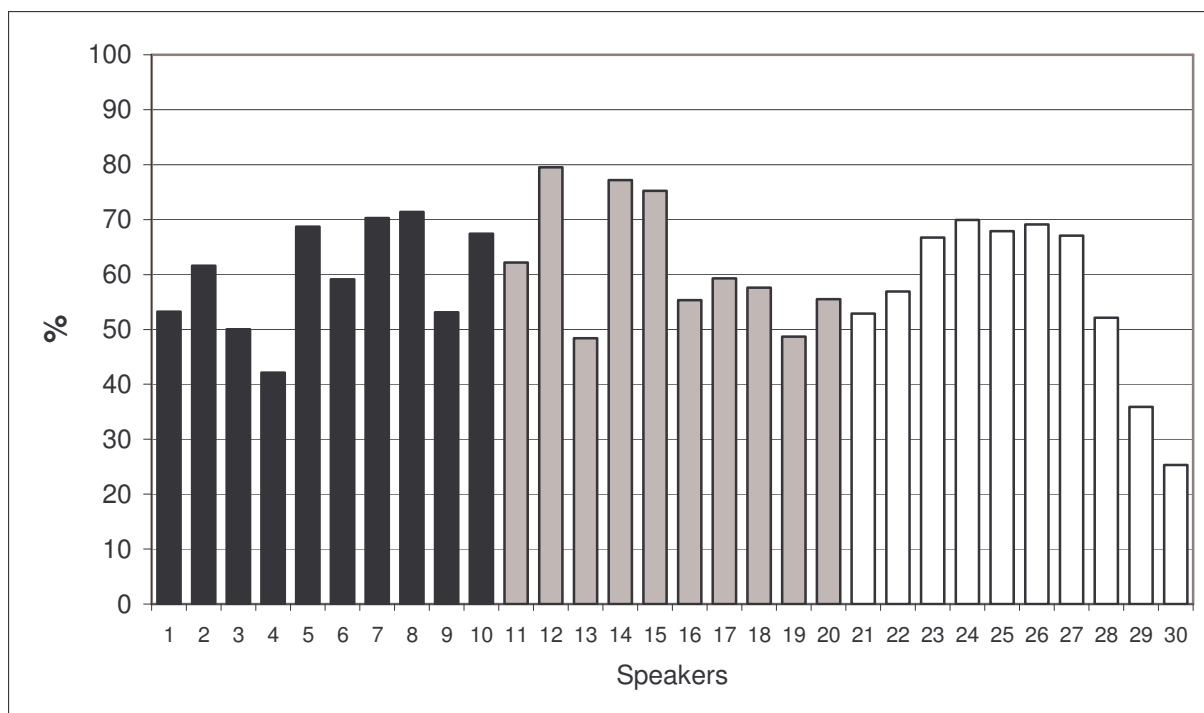


Figure 6.3.1. Percentage frequency of linking /r/, individual speakers (Black columns = BBC World, grey = Sky News, white = ITV News)

There may be several factors involved to account for the relative frequent absence of linking /r/. The relative formality of the speech situation, with its focus on clarity and comprehension, and the rate of utterance, which is somewhat slower than that of everyday conversational style, are likely to reduce the frequency of connected speech processes such as r-sandhi. Another explanation for the suppression of /r/ links may lie in the stigma connected with intrusive /r/ (cf. 4.4.3): in an effort to avoid the stigmatised intrusive /r/, many speech-conscious people also omit potential linking /r/s (cf. Wells 1982: 224; Brown 1988: 145). Moreover, omission of sandhi /r/s is said to be characteristic of adoptive RP speakers (Wells 1982: 284).

No attempt has been made here to distinguish between native and adoptive RP speakers, as that would require more biographical data. It is however very likely that at least some of the presenters speak adoptive RP (cf. 5.1.1), and this may contribute to the relative rarity of linking /r/. It is however not entirely obvious why r-sandhi in particular should be avoided by adoptive RP speakers, as this feature is present in almost all accents of England, and is not unique to RP. It is therefore likely that most adoptive RP speakers already have linking and intrusive /r/ as part of their native phonology, and do not have to “learn” them when they adopt the RP accent. A consistent use of sandhi /r/ should presumably be a

problem only for speakers who try to suppress a rhotic accent and omit linking /r/ as a form of hypercorrection. On the other hand, adoptive RP speakers may in general be more careful and speech-conscious and use a slower speech tempo than native RP speakers, and for this reason use less r-sandhi. They will then probably also use less assimilation, elision and other features of connected speech (r-sandhi and speech tempo is further investigated in section 6.3.4).

Another factor which may affect the presenters' use of linking /r/ is the focus on clarity and diction which characterises the newsreading genre. There may be a conception that sandhi /r/ potentially obscures the message, and that separating each word by omitting the linking /r/ and starting the vowel-initial word with a glottal onset enhances the clarity, as it underlines the word boundaries in connected speech.

Lastly, there may be a general diachronic shift in progress, where linking /r/ is becoming less used. This observation is made by Jones (1960: 197) with reference to RP, and by Foulkes (1997) in his study of Newcastle speech, where he reports a decrease of linking /r/ usage with younger speakers. There is however no significant empirical evidence for such a shift in RP (cf. 4.4.3).

In addition to the factors mentioned above, linking /r/ is expected to be constrained by various factors which can be quantified. A number of hypotheses about the occurrence of linking /r/ have been tested against the recorded data, and the analysis showed that the variability in the use of linking /r/ correlates with a number of factors. The first section looks at linking (r) in relation to channel and gender, while various linguistic constraints are discussed in 6.3.2.3. Stylistic aspects are discussed in section 6.3.4.

6.3.2.2 Social factors

Linking (r) is generally assumed not to vary significantly at inter-speaker level, and previous studies have revealed no correlations with social variables such as gender, age or class (cf. 4.4.3). The only exception is Foulkes' study of r-sandhi in Newcastle, where linking /r/ was found to be more frequent with older and middle-class speakers (Foulkes 1997). There were, however, no significant differences between men and women.

The overall scores for linking (r) in broadcast RP have shown that /r/ is less common than expected based on existing descriptions of RP, and that there is a certain amount of variation between speakers. The degree of speech-consciousness and articulatory carefulness, along with concepts of correctness, have been suggested as possible sources of variation. The present section examines the use of linking /r/ in relation to the inter-speaker variables, namely channel and speaker gender.

Channel

Based on the (albeit slight) differences in image, format and formality between the three channels investigated (cf. 5.1.2) and the status of linking /r/ as “correct”, one would expect the BBC World presenters to use the most r-links. A calculation of the average scores for the various channels reveals some minor differences between the three groups, but not in the direction one might expect. As shown in Table 6.3.2, the Sky News presenters have the highest /r/-scores, with an average of 61.9%, followed by BBC World with 59.7% and ITV News with 56.3%.

Table 6.3.2. Linking /r/: mean scores for channels

Channel	Mean	N	Std. Deviation
BBC World	59,700	10	9,9162
Sky News	61,890	10	11,4847
ITV News	56,380	10	15,3290
Total	59,323	30	12,2345

The ANOVA test (see Table 6.3.3) shows that the difference is not significant, as the variation between the channels is not significantly greater than the variation within the groups ($p = 0.614$). Although the highest /r/ scores are found with three Sky News speakers (12, 14 and 15), most of the other Sky presenters have scores below the average (cf. Figure 6.3.1). At the other end, two of the ITV speakers (29 and 30) stand out with markedly low scores, which skew the overall results for this channel.

Table 6.3.3. Linking /r/: ANOVA analysis for channel

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	153,929	2	76,964	,496	,614
Within Groups	4186,885	27	155,070		
Total	4340,814	29			

Gender

It is often claimed in writings on RP that linking /r/ usage is the norm in this accent, and that the feature is considered correct and desirable for native speakers as well as foreign learners (cf. 4.4.3). Such statements indicate that linking /r/ is an overtly prestigious feature. The high

status of linking /r/, combined with existing knowledge of women’s speech patterns (cf. 3.1.4) would lead us to expect higher /r/-scores for the female presenters. However, a comparison of the male and female newsreaders shows that the women use fewer r-links than the men do. The average frequency score for /r/ is 56.3% for the women and 62.2% for the men:

Table 6.3.4. Linking /r/: mean scores for gender

Gender	Mean	N	Std. Deviation
Males	62,167	15	11,3014
Females	56,480	15	12,8457
Total	59,323	30	12,2345

The gender difference is not significant, as can be seen from Table 6.3.5.

Table 6.3.5. Linking /r/: ANOVA analysis for gender

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	242,536	1	242,536	1,657	,209
Within Groups	4098,277	28	146,367		
Total	4340,814	29			

A comparison of the inter-group and intra-group variation shows that the gender contrast is most likely to be due to random factors ($p = 0.209$). This result is in line with previous observations, none of which have established any significant gender patterning for linking /r/. It is nevertheless noteworthy that all the individual scores above 75% are produced by men (speakers 12, 14 and 15), and the only scores below 40% are produced by two women (speakers 29 and 30). Of the 16 speakers who use /r/ in less than 60% of the cases, 10 are female. The tendency is therefore for women to use less r-links than men, although the pattern is not consistent. This suggests that the status of linking /r/ as “correct” is not as relevant to its usage as might be expected. Alternatively, the suppression of /r/ is motivated by a wish for phonetic explicitness, a motivation which may be somewhat stronger with the female presenters (see further 6.5.3.2).

6.3.2.3 Linguistic constraints

The use of linking /r/ is expected to be constrained by a number of linguistic factors. In the present study of linking (r), the variable has been analysed according to lexical, syntactic, prosodic and phonetic constraints.

Lexical factors

A detailed investigation of the linguistic context showed that linking /r/ occurs most frequently between short, often grammatical, words, e.g. *there are, here is, where a, or a, are also, your own*, etc. The r-link is particularly common, though far from categorical, before and after prepositions, as in *for a, before a, over a, after it, here on, more of, fear of, number of, leader of, fire in, hour of*, etc. Linking /r/ seems to occur less readily between longer lexical words, in phrases like *consumer attitude, senior administration, former economic, tighter immigration, newspaper allegations*, etc. It is found in certain common and more or less fixed expressions such as *year-old, four o'clock, far away, share index, Far East, Winter Olympics, mortar attack*, which almost invariably occur with linking /r/. Otherwise there is a tendency towards suppressing the linking /r/ across lexical words. Only 42.4% of the potential r-links are realised when they occur between two lexical items, and the majority of these involve the expression *year-old*. By comparison, the /r/ score increases to 63.2% when one or both of the surrounding words are grammatical⁶⁵. The figures are presented in Table 6.3.6.

Table 6.3.6. Linking (r) according to the grammatical category of the surrounding words

	Between lexical words		Between function words ⁶⁶	
	N	%	N	%
/r/	426	42.4	3186	63.2
∅	579	57.6	1854	38.8
Total	1005	100	5040	100

χ^2 : 151,121, p = 0.000, df = 1

⁶⁵ Grammatical words, or function words, here include determiners, pronouns, prepositions, conjunctions, auxiliaries and lexical *be*.

⁶⁶ These include cases where both the preceding and the following item is a function word, as well as cases where only one is a function word.

As linking /r/ is a connected speech feature, it is not surprising that it favours short and frequent function words and words in fixed phrases. The former are normally unstressed and carry little informational value, and are thus unlikely to be separated by a glottal stop in the speech flow. The latter are likely to be joined together by an r-link because of their close semantic relationship.

Another lexical constraint that significantly affects the occurrence of linking /r/ involves proper nouns. The data show that linking /r/ is normally dropped before names, in phrases such as *Mister Annan, Doctor Austen, Sir Alex Ferguson, prime minister Ariel Sharon, Emperor Akihito, Yasser Arafat, Lower Ormeau Road, near Al Kut, for Edward, writer Iris Murdoch*, etc. There are altogether 344 potential r-links before a name⁶⁷, and 303 of these are pronounced without /r/, and overwhelming score of 88.1% (cf. Table 6.3.7). This result confirms Lewis' claim that linking /r/ tends to be avoided before names, supposedly to secure clarity and comprehensibility (see 4.4.3), and extends the constraint to involve any word before a name, not just honorifics.

Table 6.3.7. Linking (r) before proper nouns

	Before proper nouns		Other environments	
	N	%	N	%
/r/	41	11.9	3571	62.6
∅	303	88.1	2130	37.4
Total	344	100	5701	100

χ^2 : 347,029, p = 0.000, df = 1

Stress

One of the hypotheses tested is that linking /r/ is more likely to be avoided before a stressed syllable, e.g. *for* ¹*hours*, *were* ¹*injured*, *restore* ¹*order*, *your* ¹*e-mail*. Foulkes (1997) notes that stress on the word-initial vowel may be one of the main factors preventing the use of linking /r/ in his data. The reason why linking /r/ will tend to be omitted before stress is presumably that a stressed vowel is more forcefully articulated and therefore more likely to be preceded by a glottal onset.

⁶⁷ The great majority of these are personal names. Common geographical names and names of well-known sports teams, organisations, etc. (e.g. *England, Iraq, Arsenal, IRA*) have not been included.

The analysis confirmed that /r/ is significantly less prone to occur immediately before a stressed vowel than before an unstressed one. Of the 1294 potential pre-stress r-links, only 34.9% are realised as /r/. The Ø variant, used in 65.1% of the cases, is almost invariably realised as a glottal stop [ʔ]. Table 6.3.8 shows the frequency of linking /r/ according to whether the following vowel is stressed or unstressed.

Table 6.3.8. Linking /r/ according to the weight of the following syllable

	+ stressed V		+ unstressed V	
	N	%	N	%
/r/	452	34.9	3160	66.3
Ø	842	65.1	1591	33.5
Total	1294	100	4751	100

χ^2 : 421,795, p = 0.000, df = 1

Although pre-stress linking /r/ is relatively rare, it is not universally disfavoured by the presenters: five of the speakers (12, 14, 15, 24 and 26), who all have high overall scores for linking /r/, prefer the /r/ variant also in stressed environments. Moreover, many of the pre-accentual sites are before proper nouns, where /r/ is then doubly avoided. But even if the latter examples are excluded, the general tendency is unambiguously to disfavour /r/ before a stressed syllable.

The presence of another /r/

Another linguistic constraint expected to limit the occurrence of linking /r/, is the presence of another /r/ in the immediate vicinity. Several observers have noted that linking /r/ tends to be avoided if the immediately preceding consonant is /r/ (cf. 4.4.3), e.g. *career of*, *interior of*, *lecturer is*, etc. Lewis (1975) modifies this claim and states that “if simple /r/ begins the linkable syllable, the link is generally made” (38). According to Lewis, then, it is only when the link syllable is unstressed and begins with a consonant cluster ending with /r/ that the link is avoided, as in *registrar*, *emperor* (*!emprə/*). In the present investigation all instances where /r/ precedes a potential link are included, e.g. *terror attacks*, *murderer and*, *barrier in*, *uproar across*, *there are also*, etc. In addition, I have extended the study to embrace examples where /r/ is the consonant immediately following the potential linking /r/, as in *were arrested*,

regular Iraqis, border area, Blair arrived, major air attack, etc. There are even a few rare cases where /r/ occurs on both sides of the potential r-link, e.g. *poorer areas, carrier arrived*.

The data clearly support the claim that linking /r/ is inhibited by the presence of another /r/. There are altogether 280 tokens with another /r/ nearby, and /r/ occurs in only 37 of these. This equals a percentage score of 13.2 for linking /r/ usage, compared to an overall rate of 59.8%. Table 6.3.9 gives the frequency scores for /r/ and Ø according to whether or not there is another /r/ in the immediate vicinity.

Table 6.3.9. Linking /r/ according to the presence and absence of another /r/

	Near /r/		Other environments	
	N	%	N	%
/r/	37	13.2	3575	62.0
Ø	243	86.8	2190	38.0
Total	280	100	5765	100

χ^2 : 264,402, p = 0.000, df = 1

The r-link tends to be avoided in all the relevant environments, both immediately before and after another /r/. It seems, however, to be particularly unlikely after an unstressed /r/-initial syllable in the target word (e.g. *terror alert*), where linking /r/ occurs in only 2 out of 27 tokens. No examples were recorded of potential “cluster” links (e.g. *registrar of*), and Lewis’ claim (cf. above) could therefore not be tested for linking (r). Judging on the basis of the present material, however, it seems reasonable to conclude that Lewis’ statement does not apply, as linking /r/ is on the whole avoided both before and after a “simple” /r/.

Clause boundaries

Since linking /r/ creates a phonetic bridge between two adjacent words, it is especially common between items that are syntactically close, and the majority of r-links occur between words that belong to the same phrase or clause. One potential constraint that has been investigated is the use of linking /r/ before a major syntactic boundary, or tone group boundary (cf. 4.4.3). The hypothesis is that /r/ will be avoided in such contexts, in order to separate the two syntactic units, as claimed by Jones (1960: 197; 1967: xxvi) and observed by Foulkes (1997: 85). In most cases the clause boundaries were accompanied by a pause, and such instances have not been included in the analysis. There are 203 tokens of potential

linking /r/ over a major syntactic boundary where there is no marked pause at the boundary (cf. 6.3.1). Half of these (101 tokens) are realised with /r/, as in the following examples:

- [1] *The Severn Bore, which happens twice a year/r/, is caused by the Atlantic tide.*
- [2] *Our top story this hour/r/: after a day-long assault...*
- [3] *What about there in Singapore/r/? Is there that feeling...*
- [4] *...in Tbilisi, where the celebrations were over/r/ and the streets returning to normal.*
- [5] *...since the genocide in 1994/r/, in which nearly a million people were killed.*
- [6] *Fifty-two bombers have been arriving at RAF Fairford air base in Gloucestershire/r/, a clear signal that...*

It seems, then, that the presence of a clause boundary does not prevent the insertion of /r/, although the frequency of /r/ is somewhat lower than in other environments. As can be seen from Table 6.3.10, there is a certain decrease in the frequency of linking /r/ before a clause boundary, but the effect is relatively low compared to the other linguistic constraints investigated.

Table 6.3.10. Linking (r) before clause boundary

	+ clause boundary		÷ clause boundary	
	N	%	N	%
/r/	101	49.8	3511	60.1
∅	102	50.2	2331	39.9
Total	203	100	5842	100

χ^2 : 8,731, p = 0.002, df = 1

The readiness to use linking /r/ across syntactic boundaries does however not apply to all the informants. Eight of the speakers (3, 6, 11, 18, 21, 28, 29, 30) have no instances of r-link before a clause boundary. However, these speakers typically insert a pause between two adjacent clauses, and since such cases are not included in the analysis, they contribute with a fairly small number of tokens. Only 16 of the 102 ∅-tokens are produced by these eight informants. If these instances are removed from the total, the average /r/ score for the remaining speakers is 54%.

The frequency of linking /r/ across clause boundaries corresponds fairly well with the overall patterns for linking /r/, in that those with a low total frequency of linking /r/ tend to avoid /r/ also before syntactic boundaries, whereas for those with many linking /r/s in general, the presence of a potential pause does not seem to affect the use of linking /r/. This observation contrasts with Jones' claim, and supports Lewis' (1975: 37) observation that r-links are commonly made even across clause boundaries (cf. 4.4.3).

6.3.3 Intrusive (r)

6.3.3.1 General

The number of potential intrusive (r) sites is much smaller than for linking (r): a total of 558 tokens was identified. The difference in status between the two variables (cf. 4.4.3) was expected to affect their relative frequency, and the analysis showed that intrusive /r/ is indeed significantly less frequent than linking /r/: 182 tokens are realised with /r/, which equals approximately one third of all cases. The total figures for intrusive (r) are displayed in Table 6.3.11

Table 6.3.11. Intrusive (r): total scores

	No.	%
/r/	182	32.6
∅	376	67.4
Total	558	100

As expected, then, intrusive /r/ is not as common as linking /r/. A frequency percentage of 32.6 is nevertheless relatively high for a feature which is widely stigmatised and said to be avoided by the speech-conscious in more formal styles. The individual use of intrusive /r/ is however extremely variable, as can be seen in Figure 6.3.2. Three of the speakers (10, 26 and 30) avoid intrusive /r/ completely, while three speakers (7, 12 and 25) have /r/-scores of between 70 and 85 per cent.

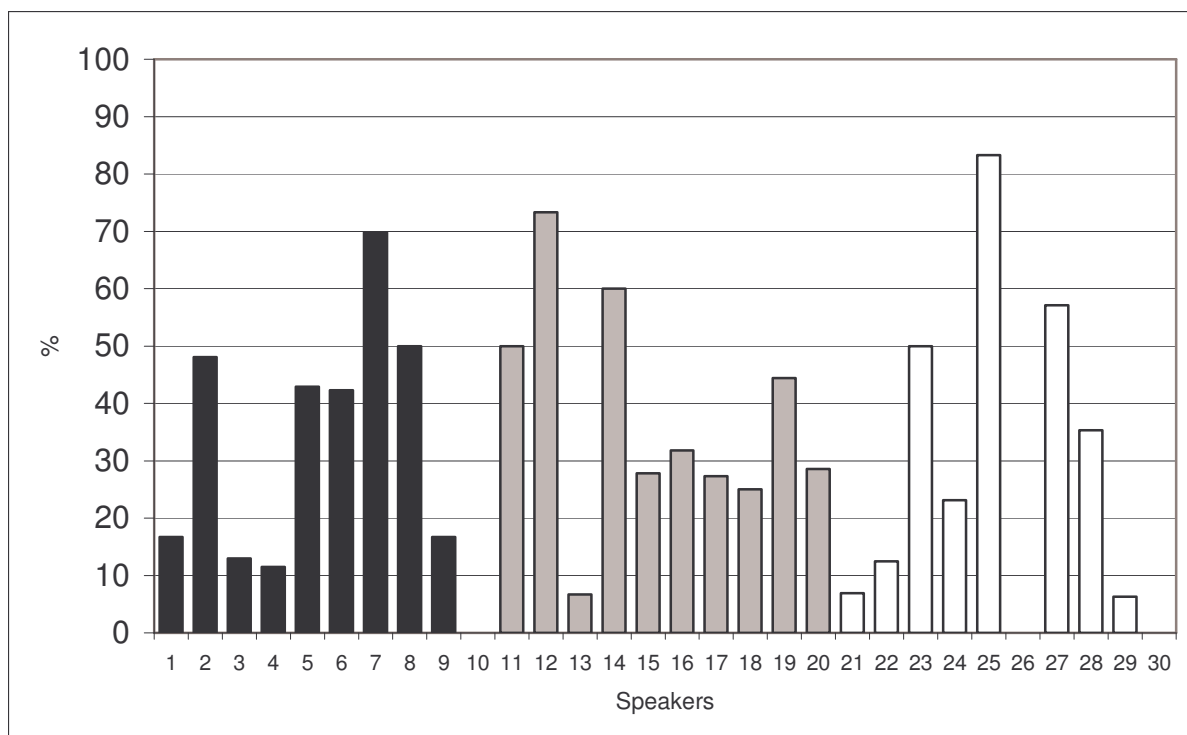


Figure 6.3.2: Percentage frequency of intrusive /r/, individual speakers (Black columns = BBC World, grey = Sky News, white = ITV News)

The extreme range in /r/ usage is partly a result of the limited number of tokens, as small numerical variations can then have a disproportionate effect on the percentage results. The variation is however not completely random: there is a certain correlation between the use of linking and intrusive /r/, as those with high scores for linking /r/ tend to have high scores also for intrusive /r/, and those who use few linking /r/s tend to have few intrusive /r/s as well (cf. Figures 6.3.1 and 6.3.2). This pattern is however not consistent. Notable exceptions are speakers 10 and 26, who both have above average scores for linking /r/, but systematically omit all intrusive /r/s. There are also instances where common r-links are dropped but intrusive /r/ is applied, within the same sentence, as in the following examples: *the border[Ø] of India[r] and Pakistan*; *Granada[r] and Carlton are[Ø] in merger talks*; *Russia[r] is a permanent member[Ø] of the Security Council*.

With the exception of speaker 25, all presenters use intrusive /r/ less than they use linking /r/, which suggests that the two variables have different status (cf. 4.4.3). The two types of r-sandhi are phonetically identical, but orthographically distinct: the intrusive /r/ has no written representation and is therefore widely regarded as incorrect or slovenly. The stigmatisation of intrusive /r/ is probably the main factor accounting for the relatively low

frequency of this variant compared to linking /r/. It does however not lead to a total avoidance of r-intrusion, as one might expect. Suppression of intrusive /r/ is believed to be particularly common with speech-conscious people or in self-conscious styles (cf. Wells 1997a: 21; Brown 1988: 145), and newsreaders, being professional speakers whose voices are broadcast to millions, are presumably among the most self-conscious users of spoken language. Moreover, as linking and intrusive /r/ are both triggered by the same preceding vowels, speakers have to use their knowledge of the spelling in order to distinguish between the two sandhi processes and avoid intrusive /r/ (cf. Wells 1982: 224; Giegerich 1992: 283). In the case of newsreaders, there is no doubt that the spelling of the target words is known to the speakers, as most of the speech is scripted. In view of the high degree of speech-consciousness involved and the immediate availability of the *r*-less orthographic forms, the fact that intrusive /r/ occurs with 27 out of 30 news presenters suggests that r-intrusion is largely an unconscious process that takes place without the speakers' being aware of it. Alternatively, the relatively widespread use of intrusive /r/ means that the feature is not as stigmatised as is often assumed and that most of the speakers therefore do not worry about it. There has undoubtedly been some change in the attitudes towards intrusive /r/ in broadcast speech: while BBC newsreaders were previously strongly advised against using r-intrusion (cf. Burchfield 1981: 10), no such explicit recommendations exist today.

The next sections look in detail at the variability in the use of intrusive (r), and explore the role of various linguistic and non-linguistic factors in the distribution of variants. Because of the comparatively low number of intrusive (r) tokens, the data-base is more limited than it was for linking (r), and any generalisations will inevitably be more tentative. Some distributional patterns can nevertheless be observed.

6.3.3.2 Social factors

The main evaluative difference between linking and intrusive /r/ is that the latter is widely condemned and disapproved of, while the former is prestigious and recommended (cf. 4.4.3). The stigmatisation of intrusive /r/ is likely to affect its usage in a public format such as newscasting. We have already seen that, although intrusive /r/ is by no means uncommon, it is relatively infrequent with most of the speakers, and is probably used less than in conversational RP. There is nevertheless a considerable amount of variation between speakers, as shown in Figure 6.3.2. The present section looks at this variation from a social point of view and examines whether the distribution of intrusive /r/ shows any correlation with the speakers' channel affiliation and gender.

There have been few sociolinguistic studies of intrusive (r), and those that do exist have generally not revealed any significant social patterning in the use of intrusive /r/ (cf. 4.4.3). The exception is again Foulkes' Newcastle study (1997), which found that r-intrusion was more common with working-class speakers than with middle-class speakers. No significant gender differences have been observed.

Channel

In view of the stigmatisation of intrusive /r/, the Sky News presenters may be expected to use more intrusive /r/s than the others, as the slightly more relaxed and informal style of this channel is likely to decrease the level of speech-consciousness and articulatory carefulness that presumably lead speakers to suppress intrusive /r/s. A comparison of the overall percentage scores for each channel, shows that the Sky presenters indeed use the most intrusive /r/s, with an average score of 37.5%, followed by BBC World with 31.1%, and ITV with an average of 27.5%. The results are summarised in Table 6.3.12.

Table 6.3.12. Intrusive /r/: mean scores for channels

Channel	Mean	N	Std. Deviation
BBC World	31,100	10	22,3663
Sky News	37,490	10	19,4550
ITV News	27,450	10	28,2598
Total	32,013	30	23,2028

Statistical analysis shows, however, that the inter-channel difference is not significant ($p = 0.635$):

Table 6.3.13. Intrusive /r/: ANOVA analysis for channel

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	516,521	2	258,260	,462	,635
Within Groups	15096,274	27	559,121		
Total	15612,795	29			

Even if the individual variation is great, it is characteristic that all the Sky presenters use intrusive /r/ to some degree, while both BBC World and ITV have speakers who avoid

intrusive /r/ completely (cf. Figure 6.3.2). Moreover, only one Sky presenter has an /r/ score below 25%, while the two other channels have altogether 11 speakers with /r/ percentages below 25, and four of the ITV presenters even have scores below 10%.

Gender

Since intrusive /r/ is widely stigmatised and considered “incorrect” by many, one would hypothesise that women would use it less than men, as women have generally been found to use less low prestige variants than men (cf. 3.1.4). A comparison of the overall scores for the male and female presenters shows that there is a certain difference between the genders: the men on average use intrusive /r/ in 35.1% of the cases, while the mean score for the women is 29% (cf. Table 6.3.14). Statistical analysis shows, however, that the difference is not significant (see table 6.3.15). This suggests that the social stigmatisation of intrusive /r/ may not be as important to its usage as one might expect, or it may be overridden by other factors. It is nevertheless noteworthy that the three speakers who categorically avoid r-intrusion are all women (speakers 10, 26 and 30), and the only scores above 70% are produced by two men (speakers 12 and 25). The tendency is thus the same as for linking (r), but, again, not consistent.

Table 6.3.14. Intrusive /r/: mean scores for gender

Gender	Mean	N	Std. Deviation
Males	35,053	15	25,0287
Females	29,040	15	21,6779
Total	32,047	30	23,2085

Table 6.3.15. Intrusive /r/: ANOVA analysis for gender.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	271,201	1	271,201	,495	,488
Within Groups	15349,133	28	548,183		
Total	15620,335	29			

6.3.3.3 Linguistic constraints

Aside from the comparatively lower status of intrusive /r/, which predictably limits its occurrence in broadcast RP, the use of intrusive /r/ is expected to be constrained by many of the same linguistic factors that were found to influence the use of linking /r/. The following sections present a detailed examination of intrusive (r) in relation to a number of such linguistic constraints. (The three speakers who consistently avoid intrusive /r/ are not included in the data discussed here.)

Lexical factors

Words with potential intrusive /r/ at the end are relatively rare and, in contrast to the linking (r) tokens, are all lexical. No generalisations can therefore be made with respect to the words preceding the intrusive /r/. The words following a realised /r/, however, tend to be monosyllabic grammatical words, e.g. *Russia is, dilemma of, China and*, etc., while r-intrusion is somewhat less likely to occur before a longer word, as in *media alliance, California investigate, Al Qaeda operative*, etc. Of the 125 tokens with potential intrusive /r/ before a polysyllabic word, only 29 (23.2%) are pronounced with /r/.

The analysis further revealed that intrusive /r/, like linking /r/, is absent immediately before names. Phrases like *Katya Adler, Patricia Amos, Muqtada Al-Sadr, Jemaah Islamiah, saw Anthony*, etc. are almost invariably pronounced without r-intrusion. 34 out of 38 (r)-sites before names are realised with Ø, which amounts to 89.5%. Intrusive /r/ after names, on the other hand, is not affected by the same restriction, even though it carries the same potential for misunderstandings and ambiguities. Intrusive /r/ is frequently heard in phrases like *Nelson Mandela/r/ is 85 today; Martina Navratilova/r/ has turned back the years; Jack Straw/r/ is meeting Russia's foreign minister; Louis Saha/r/ is not for sale; from Quetta/r/ in the southwest of Pakistan*.

Table 6.3.16. Intrusive (r) before proper nouns

	Before proper nouns		Other environments	
	N	%	N	%
/r/	4	10.5	178	34.2
Ø	34	89.5	342	65.8
Total	38	100	520	100

$\chi^2 : 9,054, p = 0.001, df = 1$

Morphology

Intrusive /r/ can occur word-internally as well as across word boundaries. Previous accounts of intrusive (r) have generally claimed that word-internal r-intrusion is the most stigmatised, and therefore least common, type of sandhi /r/ (cf. 4.4.3). Intrusive /r/ was therefore expected to be avoided immediately before a suffix, in words like *drawing*, *withdrawal*, *outlawing*. Only 12 tokens were recorded for potential word-internal r-intrusion, and four of these are realised with /r/. The number of tokens is somewhat small for quantitative analysis, but the fact that r-intrusion occurs in one third of the items at least suggests that intrusive /r/ is not particularly avoided in medial environments as compared to word-final environments.

Stress

Another linguistic constraint investigated is the effect of stress on the frequency of intrusive /r/. Cases where the potential r-intrusion is followed by a stressed vowel are expected to favour the Ø variant, realised as a glottal stop, e.g. *Asia* 'analyst, *guerrilla* 'army, *Malaga* 'airport, *Al Qaeda* 'allies. The analysis confirmed the hypothesis and supports the results for linking (r): only 7 of a total of 54 pre-accentual tokens are pronounced with /r/, which equals a percentage score of 13. All the speakers favour a glottal stop in the environment before a stressed syllable, while /r/ is virtually limited to unstressed vowels. The figures are outlined in Table 6.3.17.

Table 6.3.17. Intrusive (r) according to the weight of the following syllable

	+ stressed V		+ unstressed V	
	N	%	N	%
/r/	7	13.0	175	34.7
Ø	47	87.0	329	65.3
Total	54	100	504	100

χ^2 : 10,507, p = 0.001, df = 1

Preceding vowel

It is often claimed in previous descriptions of r-sandhi that intrusive /r/ is more likely to occur after /ə/ than after /ɔ:/ and /ɑ:/ (cf. 4.4.3). There is no evidence in my findings to support this

claim. Table 6.3.18 shows the data for intrusive (r) according to the quality of the preceding vowel.

Table 6.3.18. Intrusive (r) according to the quality of the preceding vowel

	/ə/		/ɔ:, ɑ:/	
	N	%	N	%
/r/	148	30.1	34	50.7
∅	343	69.9	33	49.3
Total	491	100	67	100

χ^2 : 11,387, $p = 0.001$, $df = 1$

Numerically, there are far more intrusive (r) sites after /ə/ than after /ɔ:, ɑ:/, but in percentage terms, intrusive /r/ actually occurs significantly more frequently after /ɔ:/ and /ɑ:/ than after /ə/ (only four tokens involve preceding /ɑ:/, but these are all pronounced with /r/). The evidence from my data thus indicates that if the quality of the preceding vowel affects the use of intrusive /r/, it is rather in the sense that /r/ is more likely to occur after /ɔ:, ɑ:/ than after /ə/. This result is in contrast to statements by both Cruttenden (2001) and Wells (1994a), and more in line with Foulkes' (1997) data for r-intrusion in Derby and Newcastle, which revealed no significant effect by the preceding vowel on the use of intrusive /r/.

The presence of another /r/

Intrusive /r/ is predicted to be inhibited by the presence of another /r/ in the immediate preceding or following environment (cf. 4.4.3). This tendency was observed with linking /r/, and the analysis showed that the same restriction applies to intrusive /r/, though not to the same extent. There are altogether 113 tokens of potential intrusive /r/ near another /r/, e.g. *extra information*, *era of*, *Korea is*, *withdraw its*, *area around*, *Pastrana arrived*, and 25 are pronounced with /r/. This amounts to 22.1%, compared to 35.3% when there is no /r/ in the immediate surroundings (cf. Table 6.3.19 below).

Table 6.3.19. Intrusive (r) according to the presence and absence of another /r/

	Near /r/		Other environments	
	N	%	N	%
/r/	25	22.1	157	35.3
∅	88	77.9	288	64.7
Total	113	100	445	100

χ^2 : 7,098, p = 0.005, df = 1

The majority of the items with realised intrusive /r/ involve either a long, stressed vowel between the two /r/-sounds, as in *Straw is* /'strɔ:r ɪz/, *draw a* /'drɔ:r ə/, or two syllables between the /r/s, as in *Korea and* /kə'ri:ər ənd/, *area of* /'eəriər əv/. Intrusive /r/ is on the whole avoided in cases where the two potential /r/s are separated only by an unstressed weak vowel, as in *camera is* /'kæmə(r) ɪz/, *Basra in* /'bæz(r) ɪn/, *drama around* /'drɑ:mə(r) əraʊnd/. The decisive factor therefore seems to be the number of syllables between the two /r/s, or whether the vowel preceding the second /r/ is full or reduced. Whether the link syllable begins with a consonant cluster or a simple /r/ does not seem to matter, and Lewis' claim that r-sandhi tends to be avoided only after clusters ending with /r/ (cf. 4.4.3) cannot be supported by the present evidence.

Clause boundary

Previous statements about r-sandhi and clause boundaries have pertained to linking (r) only, but are assumed to apply to the intrusive variant as well. The prevailing view seems to be that which is expressed by Jones (1960, 1967) and which predicts that sandhi /r/ is avoided before a syntactic boundary and potential pause (cf. 4.4.3). The investigation of linking (r) revealed that /r/ was just as often inserted as avoided before a clause boundary, and the same seems to be the case for intrusive (r) (again, those cases where there is a pause between the syntactic units are not included). There are 48 instances of potential r-intrusion before a major syntactic boundary, and 24 are realised with /r/, which amounts to 50% intrusive /r/ usage. It should be added that this type of r-intrusion is found primarily with speakers who have a relatively high overall score for intrusive /r/, and the usage is limited to 13 of the 30 speakers. With these speakers, however, the presence of a clause boundary does not seem to prevent the insertion

of /r/, as in the following examples, which illustrate how /r/ intrudes across major syntactic borders:

- [7] *Let's get the latest news from Manila/r/. Our correspondent John McClane is there.*
- [8] *...of Al Qaeda/r/. An exclusive report.*
- [9] *...if the rebels do withdraw/r/, as Barnaby Phillips now reports.*
- [10] *We shouldn't assume that this is someone from America/r/, it could be somebody living in the Philippines.*
- [11] *We're just getting some news in from the Gulf headquarters in Doha/r/. A spokesman has said that...*

It would seem, then, that the application of r-sandhi is determined more by prosody than by syntax. Moreover, the similarities in the use of linking and intrusive /r/ before clause boundaries, as well as in other linguistic environments, shows that the two variables represent the same phenomenon.

6.3.4 Style and speech rate

Linguists disagree as to whether r-sandhi is dependent on style and speech rate, and no previous studies of this phenomenon have found any clear correlation with stylistic variation (cf. 4.4.3). R-sandhi has not been established as sociolinguistically salient, and may therefore be expected to show no stylistic variation, as stylistic differentiation normally presupposes social differentiation (cf. 3.1.5). On the other hand, linking and intrusive /r/ have different connotations in terms of “correctness” which may influence their stylistic status. In general, stigmatised features tend to be particularly avoided in more formal contexts, while high-status forms show the opposite tendency. In this perspective, the use of intrusive /r/ may be expected to be greater in colloquial or informal speech and decrease in slow or careful styles, while linking /r/ may either be equally frequent in all styles or be more common in formal speech.

Another approach to r-sandhi and style is to view linking and intrusive /r/ as connected speech processes. The insertion of /r/ to avoid vowel hiatus can be said to be a phonetically motivated, “natural” process that eases the articulation of adjacent vowels and enhances the flow of speech. According to Shockey (2003: 34), the use of sandhi /r/ is a characteristic of careful speech, which favours the alternation of consonants and vowels rather than two

successive vowels. However, various other “natural” connected speech processes have been documented to be more frequent in situations where speech attention and articulatory carefulness are reduced (cf. Dressler and Wodak 1982, Nolan and Kerswill 1990). Within the framework of Natural Phonology, then, both linking and intrusive /r/ may be expected to increase in frequency in less formal styles, when the level of speech consciousness is lowered.

In order to investigate the possible stylistic correlation of r-sandhi, we need to analyse the individual speakers’ use of linking and intrusive /r/ in different situations. In this context, then, *style* refers to speech situation, and is analysed as an independent variable (cf. 3.1.5). The only situational variation that can be observed in the newscasts is the occasional shift from newsreading to interviews (cf. 5.1.3). The reading style is the most formal and involves a high degree of speech consciousness, where attention is directly focused on the text that is being transmitted. In the interview style the presenter interacts with another speaker and the amount of speech attention and articulatory care is assumed to decrease.

The number of sandhi (r) tokens produced in interview style is relatively small, and they are unequally distributed between the speakers. A total of 707 tokens occur during interviews, which corresponds to 10.7% of the total number of r-sandhi tokens, and the majority occur with speakers from Sky News and BBC World. Although the basis for comparison is limited, there is nevertheless a clear tendency towards increased use of r-sandhi in interviews, as shown in Table 6.3.20: the score for sandhi /r/ is 70.3% in interview style, as compared to 55.9% in reading style (because of the low number of tokens and the fact that the two variables show the same tendency, linking and intrusive (r) are presented together).

Table 6.3.20. R-sandhi according to style

	Interview		Reading	
	N	%	N	%
/r/	497	70.3	3297	55.9
∅	210	29.7	2599	44.1
Total	707	100	5896	100

χ^2 : 53,389, p = 0.000, df = 1

The increased frequency of /r/ usage in interviews is highly significant, and shows that r-sandhi is sensitive to situational variation, although the stylistic range investigated here is somewhat limited. The stylistic correlation is the same for linking and intrusive /r/, which

indicates that r-sandhi is in great part a “natural” process which is directly related to speech attention. The comparatively low frequency in reading style and increased usage in interviews was expected for intrusive /r/, which is a feature many speakers will suppress in self-conscious styles. The fact that linking /r/ shows the same stylistic patterning was somewhat more surprising, and suggests that speakers are less aware of, or less concerned with, the status of linking /r/ as “correct”. Alternatively, this aspect is overridden by the focus on clarity, which is more prominent in reading style, and which may lead presenters to separate adjacent words by omitting sandhi /r/s.

Another extra-linguistic factor which is closely linked to style and formality is speech tempo. A rapid speech rate is associated with fluency and informality, while a slow delivery is said to be characteristic of careful or self-conscious speech and formal styles. In the same way as style shifts were found to affect the frequency of sandhi /r/, the speed of delivery may therefore be expected to influence the use of linking and intrusive /r/, with more /r/s in the rapid deliveries.

The speech tempo of the individual news presenters is fairly stable, but there is considerable inter-speaker variation in the rate of utterance, as some speakers have markedly slow or fast deliveries. The rapid speakers were predicted to have the highest frequency of sandhi /r/s. Surprisingly, however, the rate of utterance does not seem to affect the number of linking and intrusive /r/s. The ITV presenters, who have the lowest average score for sandhi /r/, are the fastest speakers, with an average of 171 words per minute. Speaker 12, who represents Sky News, has the highest and second highest scores for linking and intrusive /r/, but is by far the slowest speaker, with 151,5 words per minute. The fastest speech rate is found with speaker 14 (187 words per minute), who does use sandhi /r/ extensively. The second fastest speaker, however, is no. 26, who omits all intrusive /r/s. A comparison of the five speakers with lowest linking /r/ scores and the five with the highest scores shows that the average speech tempo is virtually the same for both groups. The conclusion from the evidence presented here is therefore that there is no correlation between an individual’s average speech rate and his or her use of sandhi /r/. This weakens the theory that presenters who avoid r-insertions are adoptive RP speakers with particularly slow and careful deliveries (cf. 5.1.2). At the same time, there may not be a necessary correlation between speaking rate and care of articulation. The two typically interact, and their effects are difficult to isolate, but evidence suggests that care is an influence separate from rate (cf. Kerswill 1985, Dressler and Wodak 1982). It is then possible that adoptive RP speakers produce more careful and deliberate speech, but not necessarily more slow speech.

6.3.5 Summary

The findings presented here have shown that r-sandhi is a complex variable which correlates with a number of contextual factors. Intrusive /r/ is considerably less common than linking /r/, which indicates that the former is affected by a certain stigmatisation. The distributional patterns of the two variables are otherwise very similar. Neither variable shows significant gender patterning, though there is a tendency for men to use more of both sandhi /r/s. The analysis revealed a strong conditioning by certain linguistic environments on the use of linking and intrusive /r/, motivated mainly by semantic, prosodic and articulatory considerations.

Both variables further display a significant style shifting which seems to be directly related to the level of speech attention, though not to the rate of delivery. The correlation with speech-consciousness is typical of “natural” connected speech processes that are at a low level of conscious awareness. At the same time it is evident that r-sandhi is also open to direct manipulation, as it is systematically avoided in certain contexts.

The relatively low overall scores for linking and intrusive /r/ in broadcast RP indicate that the formality of the speech situation reduces the likelihood of sandhi /r/ usage. Such a finding was expected for intrusive /r/ but not for linking /r/, and suggests that the latter may not be so categorical in RP as previously believed, and that its status as “correct” RP usage is of less relevance. At the same time, it is possible that the low usage levels for sandhi /r/ may be a direct result of the nature of the speech situation. Previous studies or descriptions of linking /r/, with data from other types of context, all report /r/ levels well above the 59.8% found in the present survey (cf. Bauer 1984, Foulkes and Docherty 1999a). The stylistic comparison showed that the frequency for /r/ was significantly lower in reading style than during interviews, which suggests that the self-monitoring and focus on clarity and diction which characterise the newsreading genre, reduces the likelihood of r-sandhi usage.

6.4 T-voicing

6.4.1 Introduction

This section presents the results of the analysis of t-voicing, which involves the use of a voiced tap versus a voiceless plosive for the phoneme /t/ in intervocalic environments. 6.4.2 provides a specification of the environments investigated, and the linguistic constraints that have been distinguished. The remaining sections present the quantified results and a discussion of these in light of previous studies of t-voicing and the established hypotheses about the distribution of voiced /t/ (cf. 4.4.4).

It is not established whether t-voicing in RP represents a recent development and ongoing change, or whether it is a case of stable variation, where the voiced variant typically occurs in more casual styles or rapid deliveries. T-voicing has not previously been investigated with RP speakers, and the phenomenon is not included in the many descriptions of the accent. In various other accents, however, it has been found to display sociolinguistic patterning typical of changes in progress, and various comments and observations suggest that t-voicing is spreading in British English in general (cf. 4.4.4). If this is indeed the case, there is no reason to assume that RP should be any exception.

6.4.2 Linguistic conditions

The conditions for t-voicing are complex and, according to Shockey (2003: 29), “not yet fully understood”. The process seems, however, to be conditioned by certain linguistic constraints, and these have formed the basis for the selection of relevant environments to consider in the analysis of t-voicing.

Vowel length

Based on previous descriptions and preliminary observations of t-voicing (cf. 4.4.4), I have considered realisations of word-medial and word-final /t/ between vowels and between a vowel and a syllabic /l/ or /r/ (e.g. *pity*, *battle*, *matter of*). With regard to the preceding sound, t-voicing in RP initially seemed to be restricted to short vowels. Studies of other accents also observe that t-voicing is more common in the context of a preceding short vowel (cf. e.g. Holmes 1994: 211). As it turned out, some of the speakers exhibited occasional t-voicing also

after long vowels. I have nevertheless limited my analysis to phonemically short vowels, thus excluding items like *meeting*, *eighty*, *later*, *party*, etc. The vowel following a potentially voiced /t/ can be either short (e.g. *hitting*) or long (e.g. *at ease*).

Syllabification

T-voicing seems to be conditioned by syllable boundary, insofar as intervocalic /t/ undergoes voicing only when it is syllable-final. The syllable division is obvious when /t/ occurs word-finally (e.g. *what if*). In the case of word-medial /t/, however, the question of syllabification is a more complex issue, and a notorious analytical problem (cf. Wells 1990c, Kreidler 2004: 76-77). A widely recognised principle in contemporary phonology is the Maximal Onset Principle, which states that syllable strings should be divided in such a way that the syllable onset contains as many consonants as possible. A VCV sequence will then universally be syllabified as V.CV (cf. Crystal 2003: 325). Such an analysis would however not account for the systematic American English voicing of /t/ in words like *atom* [ætəm] and *priority* [praɪəːrɪtɪ] and the equally systematic lack of t-voicing in *atomic* [ətəːmɪk] and *prioritise* [praɪəːrɪtaɪz], as all these words contain the same V/t/V structure, but only two of the /t/s behave as syllable-final consonants.

According to Wells (1990c), t-voicing is one of several English allophonic rules which can best be described by specifying syllable boundary as part of the conditioning environment.⁶⁸ In order to account for these phonological processes, Wells proposes a different principle for English syllable division: “[s]ubject to certain conditions (...), consonants are syllabified with the more strongly stressed of the two flanking syllables” (1990c: 80). This means that the /t/ in words like *city* /'sɪt.i/, *Saturday* /'sæt.ədəɪ/ and *political* /pə'ltɪ.kəl/ belongs to the preceding, stressed syllable. /t/ is thus syllable-final and a candidate for voicing. In words like *Italian* /ɪ.'tæljən/, *determine* /dɪ.'tɜːmɪn/ and *politician* /pə'lɪtʃən/, the /t/ is syllabified with the following, stressed syllable, and voicing is not possible. The stress principle also applies to syllables with secondary or tertiary stress: in words like *retaliation* /rɪ.'tæli'eɪʃn/, *conservatism* /kən'sɜːvə.ˌtɪzəm/ and *prioritise* /praɪ'ɒrɪ.taɪz/ the /t/ is assigned to the following syllable, which in all three words is heavier than the preceding syllable. T-voicing is thus not an option in these words. It is evident from this that word-medial t-voicing is highly sensitive to prominence patterns. Among the items

⁶⁸ Other phonetic rules that are constrained by syllable division include the rule which conditions presence or absence of aspiration, that which determines the duration of vowels before fortis consonants, and those that determine possible t-glottalling, t-epenthesis and /t, d/-elision (cf. Wells 1990c: 77-79).

which do have potential intervocalic t-voicing, the only regular exception to Wells' stress-based principle is the word *whatever* /wɒt.ˈevə/, where the morpheme boundary overrides the stress principle, and the /t/ behaves like a word-final consonant and is syllabified with the preceding vowel.

With regard to adjacent syllables of equal weight, Wells posits the following codicil, which also breaks with the principle of maximal onset: “[w]here adjacent syllables are of equal grade, consonants are (...) syllabified with the leftward syllable” (Wells 1990c: 81). According to this principle, the /t/ is syllable-final, and thus potentially voiced, in words like *security* /sɪˈkjuərət.i/, *consecutive* /kənˈsekjət.ɪv/, *marketed* /ˈmɑ:kɪt.ɪd/, *orbiting* /ˈɔ:bɪt.ɪŋ/, *solicitor* /səˈlɪsɪt.ə/, *inevitable* /ɪnˈevɪt.əbl/, etc., where it is flanked by unstressed weak or short syllables. The only items which do not seem to follow this principle are words like *heretic* /ˈherə.tɪk/ and *politics* /ˈpɒlə.tɪks/, where t-voicing will normally not occur⁶⁹, even though the /t/ is flanked by unstressed syllables. This, according to Wells (1990c: 81), is because the /ɪ/ of *-ic(s)* counts as a full vowel, sufficient to outrank the weak vowel in the preceding syllable and capture the /t/.⁷⁰ The distinction between unstressed full and weak vowels is not always clear in RP, and /ɪ/ in particular is ambiguous in that it can belong to both categories (cf. Wells 2000: 844).

Wells' two main principles for syllable division provide a useful framework for identifying potential environments for t-voicing, and account for most instances of use and non-use of intervocalic voicing in English. The same syllabification principles are adopted in LPD (Wells 2000), and they have formed the basis for the identification of potentially voiced /t/s in the present study.

An alternative approach to t-voicing, used in some models of non-linear phonology, analyses a potentially voiced /t/ as ambisyllabic, i.e. belonging to both adjacent syllables simultaneously (cf. Harris 1994: 198-200). This ambisyllabicity is created through a rule of resyllabification, whereby a foot-internal /t/, in e.g. *city* /ˈsɪ.ti/, is resyllabified into the preceding rhyme, while keeping its attachment to the onset. A word-medial /t/ may then undergo voicing (or lenition) only if it is ambisyllabic. The resyllabification rule is however

⁶⁹ T-glottalling is also blocked in these words.

⁷⁰ Steriade (2000) argues that in American English, tapping of /t/ between two weak vowels may be blocked in order to secure *Paradigm Uniformity*. This explains why /t/ is realised as [t] in *militaristic* [ˌmɪlətəˈrɪstɪk], but as [ɾ] in *capitalistic* [ˌkæpərəˈlɪstɪk]. The exceptional untapped [t] in the derivative *militaristic* corresponds to untapped [t] in the base form *military* [ˈmɪləˌtəri], while the tap in *capitalistic* corresponds to the tap in *capital* [ˌkæpərəl]. The paradigm of {*military*, *militar(-istic)*} becomes less variable phonologically if the [t] is generalised to the unstressed syllable.

also stress-sensitive, as only a foot-internal /t/ is ambisyllabic, not a foot-initial one, as in *pre¹tend*. The ambisyllabicity approach is therefore a more complex, but not a more productive, model of syllabification.

Morphology

Two morphological constraints have been distinguished in the analysis of t-voicing, namely the presence and absence of a word boundary after the /t/, or word-final versus word-internal voicing (e.g. *put in* vs. *putting*). With reference to the latter, I have not distinguished between words where the /t/ is followed by a morpheme boundary (e.g. *sitting*) and words where there is no such boundary (e.g. *city*). This is mainly because the number of words with medial voicing is relatively small (see 6.4.2.3), and does not provide enough basis for comparison and generalisation. Previous studies of other English varieties have yielded contradictory results for word-medial voicing: Patterson and Connine (2001: 259) find that t-voicing is more common in simple words compared to morphologically complex items, while Bell (1977) reports the opposite result (referred in Holmes 1994).

Word-final /t/ has been counted as potentially voiced only in cases where there is no pause between the two adjacent words, and no glottal onset in the second item. Moreover, instances of word-final /t/ before an unstressed word with initial /h/, e.g. *that he*, have been included in those cases where the /h/ is elided.

Stress

It is widely assumed that t-voicing is restricted to the environment following a stressed syllable and preceding an unstressed one (cf. 4.4.4), and this environment has therefore received most attention in studies of t-voicing. Patterson and Connine (2001), for example, include only post-accentual /t/ in their study of American English tapping, as do Zue and Laferriere (1979). Woods (1991: 136) limits his study of t-voicing in Ottawa English to the same environment. Sivertsen (1960:109), on the other hand, notes that in Cockney, voiced /t/ may occur in any stress environment. The same observation is made by Holmes (1994) with reference to New Zealand English (she even records occasional t-voicing in words like *catastrophe*). Holmes' analysis shows, however, that voicing is most likely to occur between syllables of unlike stress, and is disfavoured between vowels of equal stress. In the present analysis, I have considered all instances of intervocalic /t/, regardless of stress, with the only limitation that the /t/ must be syllable-final. The data thus includes /t/ in the environments

'V__V (e.g. *'better, a 'lot of*), V__'V (e.g. *what'ever, it 'is*), V__V (e.g. *se'curity, at a*), and 'V__'V (e.g. *'that 'area*).

6.4.3 Results

The present section gives the results of the quantitative analysis of t-voicing in relation to the linguistic and non-linguistic factors which are believed to affect the use of voiced versus voiceless /t/. The overall scores are presented first, followed by the results for social factors (channel and gender) and a number of linguistic constraints, and lastly the stylistic variation (presentation vs. interview). The findings are summarised in 6.4.2.5

6.4.3.1 General

In the analysis of t-voicing two variants have been identified, which are distinguished primarily by voicing: [t̚], a voiceless alveolar plosive which in intervocalic environments is normally aspirated or affricated, and [t̚], a voiced variant, normally realised as a tap. No attempt has been made to distinguish between taps and other voiced realisations. The [t̚] variant thus potentially includes the realisations [ɾ], [t̚] and [d]. The most frequent realisation is by far the tapped [ɾ].

A total of 5615 tokens was analysed, which corresponds to an average of 187 tokens per speaker. The analysis showed that, as expected, the voiceless [t̚] is the most frequent variant of /t/ in intervocalic environments. The voiced [t̚] is however far from uncommon: in total, more than one third of all tokens are realised with t-voicing. The exact figures and percentages are given in Table 6.4.1.

Table 6.4.1. T-voicing: total scores

	N	%
[t̚]	1976	35.2
[t̚]	3639	64.8
Total	5615	100

A score of 35.2 % for [t̚] is strikingly high, considering the relative formality of the newscast setting, and the fact that t-voicing is generally not considered a common feature of

RP (cf. 4.4.4). The result clearly suggests that [t̚] is a well-established variant of intervocalic /t/ in RP, and thus not limited to non-standard accents or casual style.

The individual [t̚] scores are shown in Figure 6.4. Although there is a great deal of inter-speaker variation, with percentage scores ranging between 10 and 60, all the speakers have t-voicing as part of their phonetic inventory. Four of the speakers (12, 15, 17 and 23) even favour the voiced variant over [t̚] in the intervocalic contexts considered here. The individual variation is primarily linked to speaker gender, which will be discussed in section 6.4.3.2. below.

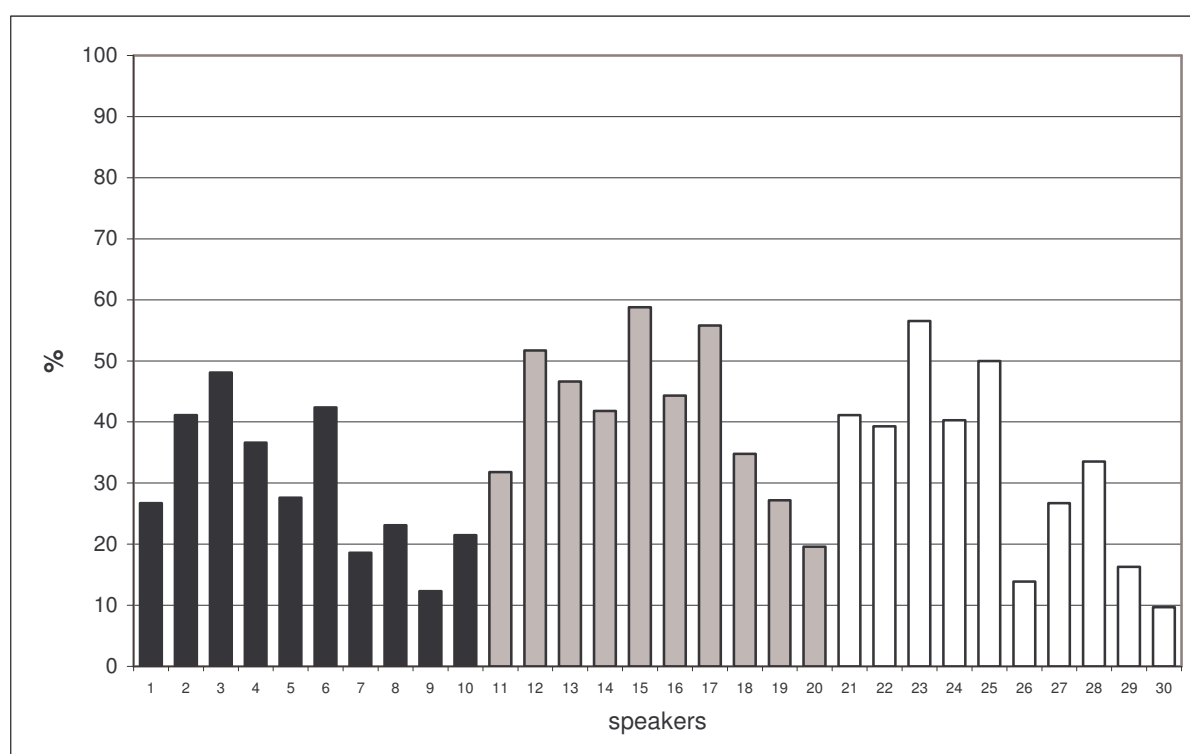


Figure 6.4. Percentage frequency of [t̚], individual speakers (Black columns = BBC World, grey = Sky News, white = ITV News)

6.4.3.2 Social factors

T-voicing appears to have many of the characteristics of a vernacular variant. It has been described as typical of casual or informal style, and is reported to be most common with working-class speakers, with younger people, and with males (cf. 4.4.4). The social salience of voiced /t/ is expected to have certain implications for the distributional patterns of [t̚] at inter-speaker level. The present section discusses t-voicing according to channel and speaker gender.

Channel

According to Holmes (1994: 216) newsreading “is a relatively distinctive activity which is likely to favour voiceless over voiced variants”. She refers to New Zealand English, where t-voicing is well-established, and the claim seems likely to be even more true of RP, where t-voicing is not a recognised feature. The most interesting observation is therefore the considerable extent to which t-voicing is used in the British news channels investigated here, which all represent a relatively high level of formality, with regard to both style, topic and setting.

Bell’s (1982) analysis of New Zealand broadcast speech demonstrates a covariation of language style with the character and “image” of various radio stations, and t-voicing is one of the variables that consistently reflect this style shift. Bell’s study revealed that newscasters increased their levels of t-voicing on the commercial local stations compared to the more prestigious national network station. In view of the differences in style and image between the three channels investigated here (cf. 5.1.2), and the social and stylistic connotations of t-voicing (cf. 4.4.4), it is expected that voiced /t/ will be used most frequently by the newsreaders on Sky News. A comparison of the mean scores for the three channels seems to confirm this hypothesis. As can be seen in Table 6.4.2, the Sky News presenters have a mean [t̥] score of 41.2%, while the scores for ITV and BBC World are 32.7% and 29.8% respectively.

Table 6.4.2. T-voicing: mean scores for channels

Channel	Mean	N	Std. Deviation
BBC World	29,800	10	11,6786
Sky News	41,240	10	12,7548
ITV News	32,730	10	15,7256
Total	34,590	30	13,9254

The inter-channel differences are however not consistent. Although the scores for the Sky presenters tend to be overall relatively high, the scores for the other channels are extremely variable. The widest range of variation is found with the ITV speakers, who exhibit scores ranging from 9.7% (speaker 30) to 56.4% (speaker 23). The ANOVA test shows that the variation within each channel is so extensive that the inter-channel differences are not statistically significant ($p = 0.163$).

Table 6.4.3. T-voicing: ANOVA analysis for channel

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	706,262	2	353,131	1,939	,163
Within Groups	4917,305	27	182,122		
Total	5623,567	29			

In Bell's (1982) study of t-voicing in New Zealand radio speech, the stations' corporate style (audience design) proved to be the most important factor influencing the presenters' use of voiced /t/, massively overriding any individual differences between speakers. The same contrast cannot be observed in the present study, mainly because the stylistic differences between the channels in the present study are fairly small. Moreover, Bell's informants were all male, and his study thus excludes the potentially most important social factor, namely gender.

Gender

Various studies of t-voicing have shown that men use more voiced /t/s than women (cf. 4.4.4), a common pattern for vernacular variants. The same distribution was hypothesised to appear in the present study. A comparison of the two speaker groups confirms the expectation that the male speakers use considerably more voiced /t/s than the females. As outlined in Table 6.4.4, the average percentage frequency for [t̥] is 42.5 for the men and 26.4 for the women.

Table 6.4.4. T-voicing: mean scores for gender

Gender	Mean	N	Std. Deviation
Male	42,533	15	9,6035
Female	26,647	15	13,2004
Total	34,590	30	13,9254

The gender contrast is fairly consistent across the sample. A detailed look at the individual [t̥] scores shows that 12 of the 15 female speakers have scores below the average (cf. Figure 6.4). The three exceptions are speakers 6, 16 and 17. Only one woman (speaker 17) uses [t̥] more than [t], while four men have [t̥] scores of 50% or more. Ten of the women presenters have results below 30%, with 9.7% as the lowest score, while the lowest male

score is 26.7%, and only two men use [t] in less than 30% of the cases. The gender difference is thus quite systematic, and statistical analysis shows that it is highly significant ($p = 0.001$). The data from the ANOVA test are displayed in Table 6.4.5 below.

Table 6.4.5. T-voicing: ANOVA analysis for gender

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1892,896	1	1892,896	14,207	,001
Within Groups	3730,671	28	133,238		
Total	5623,567	29			

The most significant individual variation in t-voicing is thus found between the two genders. This result for broadcast RP is consistent with previous findings for t-voicing in other varieties of English. The overwhelming male dominance in the use of voiced /t/ thus supports the established theory that men prefer the more “rough”, vernacular variants, while women aim for the more overtly prestigious pronunciations.

T-voicing can also be seen as a phonological process involving phonetic reduction and increased ease of articulation. In this perspective, the reported results suggest that men are more prone to using reduced variants, while women aim for phonetic explicitness and prefer to articulate the “full” form (cf. 3.1.4).

6.4.3.3 Linguistic factors

As a process of phonetic reduction, t-voicing is predicted to correlate with certain linguistic factors, notably word-position and stress. Moreover, since t-voicing may represent a new, or at least an increasing, trend in RP, it is expected to have a limited lexical distribution. The following sections look at t-voicing in relation to morphological, lexical and prosodic constraints.

Word position

Based on preliminary observations of RP and previous studies of other varieties (e.g. Holmes 1994), t-voicing was expected to be more common across word boundaries than within words. An investigation of the linguistic environment of the voiced /t/s confirmed this prediction. In fact, the analysis showed that the overall most important factor affecting the likelihood of t-voicing is whether the /t/ occurs word-finally (e.g. *what about*) or word-internally (e.g.

matter). The quantified results revealed that t-voicing is almost four times as likely to occur across word boundaries than within words: 66.9% of the word-final /t/s are realised as [t̚], while only 17.5% of the word-medial /t/s are voiced (cf. Table 6.4.6).

Table 6.4.6. T-voicing according to word position

	Word-final		Word-medial	
	N	%	N	%
[t̚]	1344	66.9	632	17.5
[t]	664	33.1	2975	82.5
Total	2008	100	3607	100

χ^2 : 1380,811, p = 0.000, df = 1

A score of 66.9% for final [t̚] shows that voicing is significantly favoured over the voiceless [t] in word-final intervocalic environments. Eight of the 30 speakers have scores of more than 80% for [t̚] across word boundaries, and only six speakers have scores below 50%.

The marked difference in frequency between final and medial [t̚] suggests that in RP t-voicing is primarily a feature of connected speech, and is not so well-established that it will appear regularly in citation forms. The preference for word-final voiced /t/ has however been observed also with speakers of New Zealand English (Holmes 1994), where t-voicing is a much more prominent feature. Holmes suggests that the reason for this may be the difference in salience associated with the two environments: the word-medial (post-accentual) intervocalic position (e.g. *letter*, *nutty*) is a perceptually very conspicuous environment, and the one most often associated with the popular stereotype of American speakers. Speakers tend to be less aware of the word-final environment, which is perceptually less salient and generally more prone to reduction (cf. Shockey 2003: 15). The alleged difference in salience has, to my knowledge, not been empirically measured. But similar findings for t-glottalling seem to support such an analysis. Various studies have reported that word-medial intervocalic position is the most salient and stigmatised environment, and the last environment for diffusion, for t-glottalling (cf. Mathisen 1999, Fabricius 2000: 20, Trudgill 1986: 20, Milroy et al. 1994: 21). The higher level of t-voicing in the less salient word-final environment may thus be an indication that t-voicing is a change from “below the level of consciousness” (Labov 1994: 78, Holmes 1994: 218-19).

Lexical factors

Lexical information has been identified as relevant to the extent that t-voicing is hypothesised to be more likely in high frequency function words such as *it, that, what, get, but*, etc. and common lexical words such as *British* and *getting*, and less likely in longer and less frequent words or phrases such as *solicitor, Nottinghamshire, private engineering*, etc. This assumption is based on established facts about reduction processes and the spread of linguistic changes, as well as previous observations of t-voicing. Features of phonological reduction are more likely to affect short, frequent words (cf. e.g. Shockey 2003). Moreover, if t-voicing represents an ongoing change in RP, a frequency-based pattern is to be expected, as reductive changes typically affect the most common words first, through the process of lexical diffusion (cf. 3.1.2). A similar distributional pattern for voiced /t/ has also been observed for accents where t-voicing is a firmly established feature: in their study of American English tapping, Patterson and Connine (2001) found that low-frequency words and morphologically complex items showed a lower incidence of tapping than high-frequency words and morphologically simple words.

A detailed look at the words affected by t-voicing showed that, as predicted, the lexical distribution of [t̚] is not random. T-voicing across word boundaries, which is the most common type, primarily affects word-final /t/ in monosyllabic function words such as *it, that, but, at, what*, etc., in phrases such as *it is, that it's, but if, at a, what about*, etc. T-voicing is also commonly found in other short frequent words, such as *not* and *get*, and in expressions like *a lot of, a bit of, a set of*. Although t-voicing usually involves a monosyllabic word in the left hand environment, the same restriction does not apply to the second segment, which is often a longer lexical item, e.g. *not exactly, get elected, what appeared, that Israel, it affected, that explosion, at Eton, but Osama*, etc. In total, 68.3% of the monosyllabic items are realised with word-final voiced /t/ (cf. Table 6.4.7), which makes this the most favoured environment for t-voicing. One of the main reasons why t-voicing primarily affects monosyllabic function words is presumably that since these words are short and frequent, they are usually unstressed and quickly articulated, and are thus more prone to undergo phonetic reduction in running speech.

Table 6.4.7. Word-final t-voicing according to the length of the target word

	Monosyllabic words		Polysyllabic words	
	N	%	N	%
[t̚]	1301	68.3	43	41.7
[t]	604	31.7	60	58.3
Total	1905	100	103	100

$\chi^2 : 31,113, p = 0.000, df = 1$

Longer content words are more often stressed and carefully articulated, and are therefore more likely to favour the voiceless plosive. Word-final t-voicing was thus predicted to be much less common in polysyllabic words. Lexical combinations such as *rocket attack*, *consulate of*, *market activity*, are numerically much rarer (only 103 tokens), which makes quantification less reliable. The analysis suggests, however, that longer words disfavour t-voicing, as only 43 of the polysyllabic items are realised with a final [t̚] (e.g. *target of*, *forget it*, *budget airlines*, *separate incident*, *private operations*). Although a score of 41.7% is well above the overall frequency for voiced /t/, the result demonstrates that final t-voicing is significantly favoured in monosyllabic words, and that a long or heavy target word reduces the likelihood of voiced /t/ across word boundaries.

The analysis of word-medial, or word-internal, t-voicing suggests that voicing mainly affects a limited number of frequent lexical items. Word-medial t-voicing is relatively rare, and the majority of the target words are always realised with [t]. Only 30 different words are pronounced with voiced [t̚] on two or more occasions, and only 13 are voiced ten or more times. The words most frequently realised with [t̚] are listed in Table 6.4.8 below. The table shows the total number of tokens, the number of times they are realised with t-voicing, and the percentage frequency of the voiced variant. The BNC FrQ indicates the frequency per million words in the British National Corpus of spoken and written English (from Leech et al. 2001).

Table 6.4.8. Lexical distribution of word-medial t-voicing in words with N ≥ 10

Word	% [t̥]	N [t̥]	Total N	BNC FrQ
<i>pretty</i>	83.3	30	36	52
<i>whatever</i>	76.9	10	13	132
<i>getting</i>	66.7	66	99	203
<i>putting</i>	59.1	13	22	76
<i>British</i>	45.5	207	455	357
<i>hospital</i>	38.7	36	93	180
<i>Scottish</i>	37.7	20	53	98
<i>capital</i>	35.9	55	153	138
<i>criticism</i>	32.3	10	31	59
<i>better</i>	18.6	11	59	143
<i>little</i>	16.3	20	123	306
<i>security</i>	12.7	27	212	158
<i>city</i>	8.4	23	273	275

The relative rarity of word-medial t-voicing and the limited number of tokens make generalisations difficult and conclusions tentative. We can however observe certain tendencies in the lexical distribution of voiced /t/. As can be seen from Table 6.4.8, medial voicing primarily affects words of relatively high lexical frequency. The majority of voiced tokens occur in very common words and are limited to a relatively small number of items. The word *British* alone, which is realised with voicing in almost half the cases, constitutes almost one third of the instances of [t̥]. The only words that are more often realised with [t̥] than with [t] are *getting*, *putting*, *whatever*, and the adverb *pretty*. Another frequently voiced item is the colloquial form *gotta* (contraction of *got to*), which occurs only five times in the corpus, but each time with a voiced [t̥]. It is also noteworthy that words like *capital*, *hospital* and *little* are among the words most frequently pronounced with [t̥], which suggests that t-voicing is equally likely to occur before a syllabic /l/ as before a vowel.⁷¹

By comparison, the words that never receive internal t-voicing are overall much less frequent. Of the 234 different items that occur only with [t], only 11 have a BNC frequency above 100 (including *authority*, *community*, *letter*, *majority*, *university*), and 110 are not even included in the BNC list (among others *brutality*, *chaotic*, *obesity*, *pragmatic*, *senator*).

The relatively close correlation between t-voicing and lexical frequency observed in the distribution of medial voiced /t/, may indicate a process of lexical diffusion (cf. 3.1.2), and

⁷¹ The number of tokens with potential t-voicing before syllabic /t/ (e.g. *matter of*) were too few for quantification.

thus support the theory that t-voicing is spreading in RP. There is further a connection between t-voicing and informality, as the words that favour voicing, such as *getting*, *pretty* and *gotta*, are all relatively informal in a newscasting context. This can be seen to strengthen the association of t-voicing with casual style.

Stress

One of the potentially relevant constraints on t-voicing is whether the preceding or following syllable is stressed. Previous studies of t-voicing generally hold that [t̚] is most common in the environment between a stressed and an unstressed syllable (cf. 4.4.4). With regard to voicing across word boundaries, the analysis of broadcast RP revealed that the weight of the second vowel appears to have little effect on the likelihood of t-voicing. The majority of the tokens involve an unstressed syllable immediately after the word-final /t/ (e.g. *what a'bout*, *it ap'pears*, *at a*, *a 'lot of*, etc.), and t-voicing is very common in these types of phrase. Voicing is however equally likely before a word-initial stressed syllable (e.g. *what 'else*, *but 'even*, *but 'obviously*, *at 'Aintree*, *separate 'incident*, etc.), although these are numerically in a minority.⁷² However, the presence or absence of t-voicing does seem to be affected by whether or not there is stress on the syllable immediately preceding /t/. As can be seen from Table 6.4.9, word-final voicing is somewhat less likely in the environment after a stressed syllable than after an unstressed syllable. A comparison of the two environments showed a significant difference in the frequency of [t̚] ($p = 0.000$). The environment which least favours t-voicing is /t/ between two stressed syllables, as in *'not 'all* (though it should be noted that the number of tokens is rather low).

The preference for [t̚] after an unstressed vowel contrasts, as mentioned above, with findings for other varieties. This difference may be related to the fact that t-voicing in RP primarily is a reduction phenomenon of connected speech, while in American English, for example, voicing of intervocalic /t/ is the “normal” pronunciation. If t-voicing is viewed as a phonological process of lenition, it is to be expected that final [t̚] is more common after unstressed vowels, as weak syllables are generally more prone to undergo reduction in connected speech.

⁷² The comparatively low number of t-voicing sites before a stressed vowel partly results from the fact that a word-initial stressed vowel is often accompanied by a glottal onset, which excludes word-final voicing. An unstressed initial syllable is more likely to follow immediately without any intervening glottal stop.

Table 6.4.9. Word-final t-voicing by stress environment

	'V__V		'V__'V		V__V		V__'V	
	N	%	N	%	N	%	N	%
[t̚]	343	63.9	67	51.1	702	69.4	232	70.7
[t]	194	36.1	64	48.9	310	30.6	96	29.3
Total	537	100	131	100	1012	100	328	100

χ^2 : 13,956, $p = 0.000$, $df = 1$ (comparing 'V_V/'V_'V with V_V/V_'V).

With regard to word-medial t-voicing, the number of possible stress patterns is more limited. There is only one word which potentially contains a voiced [t̚] before a stressed syllable, namely *whatever* (cf. 6.4.2 above). Word-internal t-voicing is thus virtually restricted to the environment before an unstressed syllable. The stress of the preceding syllable, however, again seems to influence the likelihood of /t/ being voiced, but with the opposite result from the one observed for word-final voicing. Among the words that most often exhibit voicing, the great majority have a stressed vowel before /t/ (cf. Table 6.4.8 above). T-voicing between two unstressed syllables occasionally occurs in frequent words like *capital*, *hospital*, and *security*, but is otherwise rare. This finding is in accordance with several previous observations of t-voicing in other accents, which all agree that the post-accentual position is the preferred phonetic environment for medial voicing (e.g. Patterson and Connine 2001, Zue and Laferriere 1979). It does, however, not fit with the idea that unstressed syllables are the ones most likely to undergo lenition. It seems, then, that word frequency is the decisive factor when it comes to internal t-voicing.

6.4.3.4 Style and speech rate

T-voicing is generally viewed as characteristic of informal style in British English (cf. 4.4.4), although empirical evidence is scarce. This stylistic correlation has been established in studies of t-voicing in non-British accents of English (e.g. Holmes 1994, Woods 1991), where the voiced variant is generally found to be more common in conversations and informal contexts than in formal interviews and reading style. Moreover, the social differentiation of t-voicing increases the probability that the feature will also show stylistic differentiation. T-voicing has been demonstrated to occur more frequently with working-class speakers than with middle-class informants, and is used more by males than by females, while the voiceless [t] is by many associated with middle-class or “posh” speakers (cf. 4.4.4). Such variation on the social

dimension is presumed to be paralleled on a stylistic continuum, with increased use of [t̥] in informal styles (cf. 3.1.5).

The relatively high overall score for voiced /t/ in the present study shows that t-voicing is not limited to casual or conversational style (cf. 6.4.3.1). There is nevertheless no reason not to believe that the use of [t̥] would be even higher in a less formal context. The comparison of the channels has already indicated that degree of formality affects the use of t-voicing (cf. 6.4.3.2). In order to establish whether there is also intra-speaker stylistic variation, the individual speakers' use of [t̥] in presentation style has been compared with the use of [t̥] in interviews. The latter style is considered less formal (cf. 5.1.3), and is predicted to be accompanied by an increased frequency of voiced /t/.

As noted in the discussion of /r/-sandhi (6.3.4), the number of tokens obtained during interviews is limited, and the amount of time spent on interviews varies greatly from speaker to speaker. In all, 892 (t)-tokens are uttered in the course of interviews, which constitutes 15.9% of the total number of tokens, and most of these are produced by presenters on Sky News or BBC World. The basis for comparison of the two styles is therefore not ideal, but the amount of data is large enough to perform a quantitative analysis. And the findings revealed that the frequency of t-voicing increases significantly in interview style: the total score for [t̥] is 51.5% in interviews, compared to 32.1% in reading style, as shown in Table 6.4.10.

Table 6.4.10. T-voicing according to style

	Interview		Reading	
	N	%	N	%
[t̥]	459	51.1	1517	32.1
[t]	433	48.9	3206	67.9
Total	892	100	4723	100

χ^2 : 123,023, p = 0.000, df = 1

The increase in the use of voiced /t/ in interview style is consistent with previous stylistic results for t-voicing in other accents of English, and confirms the informal status of voiced /t/.

The stylistic variation of t-voicing, and the fact that it is a reduction process that enhances the ease of articulation (cf. 4.4.4), suggests that the use of [t̥] is likely to increase with the rate of speech. According to Wells (1982: 325), t-voicing does not occur in slow

speech in an accent such as Cockney, where t-voicing is otherwise very prevalent. This indicates that t-voicing is closely related to the speed of delivery. The present survey has shown, however, that t-voicing is not limited to fast speech, since newsreading is performed at a slightly slower rate than ordinary conversational speech (cf. 5.1.3). As already noted, the speaking rate of the individual presenters remains fairly stable, while there are considerable inter-speaker differences in speech tempo. The effect of speech rate on the use of t-voicing could then not be tested on intra-speaker level, but the individual use of t-voicing has been compared with the presenters' general rate of utterance. The analysis showed that the fastest speakers do not necessarily use more voiced [t̚]s than those with slow deliveries. The speaker with the highest [t̚] score (speaker 15) has a speech rate below the average, and a comparison of the presenters with the highest and lowest percentage of [t̚] revealed that the average tempo is almost exactly the same for the two groups. It seems, therefore, that, as was observed with r-sandhi, the individual speech tempo does not influence the use of t-voicing.

6.4.4 Summary

The present section has presented the findings for intervocalic t-voicing after short vowels. In view of previous descriptions of RP, the formality of the newscasting situation, and the social and stylistic connotations of t-voicing, the overall frequency of the voiced [t̚] in relation to [t] is surprisingly high. Traditional descriptions of RP do not include t-voicing as a feature of this accent. Wells (1982: 125) states that the voiced tap [ɾ] “is rare in RP (RP /t/ remaining voiceless in all environments)”, and according to Chambers (2002: 256), post-tonic prevocalic /t/ “occurs as alveolar, fortis [t]” in “standard British accents”. Based on the evidence presented here, such statements need to be modified. The prevalent use of voiced /t/ with the speakers in the present study suggests that a change may be taking place in RP, whereby t-voicing is becoming increasingly common. The lexical distribution of [t̚] is a further indication of change, as it shows signs of lexical diffusion, with voicing affecting the most frequent words first.

Previous accounts of t-voicing in other accents of English, suggest that the feature originates in non-standard varieties and casual styles. The social and stylistic correlations demonstrated here, seem to confirm this interpretation: [t̚] is significantly more common with male speakers and in the less formal register, a pattern which has become familiar for

vernacular variants. Consideration of the linguistic environments showed that the most important linguistic constraint is word-position: t-voicing is much more common across word boundaries than word-medially. [t̚] is actually by far the preferred variant of intervocalic /t/ word-finally, which establishes t-voicing as a prominent connected speech feature of RP.

T-voicing is a purely phonetic type of variation, taking place at the realisational allophonic level, and is not subject to the same attention and social stigmatisation as t-glottalling. The motivation for using t-voicing is possibly related to ease of articulation (cf. Chambers 2002: 256-257, Trudgill 1986: 22). The use of [t̚] instead of [t] between vowels is articulatorily economical, and thus a natural and phonetically motivated sound change. It is therefore likely that t-voicing is largely a subconscious process with RP speakers. The data analysed in this study seem to allow for such an interpretation, as t-voicing is found to be most frequent in both linguistic and stylistic contexts where the speakers are (presumed to be) less conscious of their usage, namely across word boundaries and in interview style.

6.5 Smoothing

6.5.1 Introduction

Smoothing involves the reduction of the triphthongal sequences /aɪə, aʊə/ (FIRE, POWER) to a centring diphthong [aə] or a monophthong [a:]. Smoothing can affect all closing diphthongs in the environment of a following /ə/, but the present study focuses on /aɪ/ and /aʊ/ before /ə/, which are the vowels most readily affected (cf. 4.4.5). The analysis includes all the words that are pronounced with /aɪə/ or /aʊə/ in their citation form, such as *fire*, *science*, *society*, *power*, *hour*. I have disregarded words that have an alternative pronunciation with only /aɪ/ or /aʊ/⁷³, such as *violence*, *diamond*, *inquiry*, as well as the word *our*, which has the monophthongal realisational variant /ɑ:/ alongside and equal to /aʊə/. I have also excluded words where the triphthongal sequences occur in an unstressed syllable preceding the primary stress, as in *scien^tific*, *vio^lation*, *bio^logical*, because these environments are especially prone to phonetic reduction and thus difficult to analyse auditorily. The same problem does however not apply to cases where /aɪə, aʊə/ occur after the primary stress, such as ¹*sapphire*, ¹*gunfire*, ¹*qualifier*,

⁷³ *Longman pronunciation dictionary* has been used as the main reference in identifying these words.

'sunflower, and these items are therefore included in the analysis. The great majority of the relevant triphthongs, however, appear in stressed syllables.

6.5.2 The auditory analysis of smoothing

Smoothing was perhaps the most difficult variable to analyse impressionistically, as it is the one which most clearly involves variants situated on a continuum of vowel quality. In the auditory analysis I have only distinguished two variants: [aɪə, aʊə] and [aə], which represent absence and presence of smoothing, respectively. The former includes all realisations where a second vocalic element, representing [ɪ] or [ʊ], can be identified. This mid element is often very weak, particularly in those cases where the triphthong is pronounced as one syllable, and is far from as close as the phonetic symbols indicate. The tongue rarely reaches the [ɪ] and [ʊ] position. The actual triphthongal realisations are something closer to [aɛə] and [aɔə] (cf. Jones 1960: 106, 109). The smoothed variant, symbolised [aə], incorporates all realisations where no mid element can be identified. The phonetic distinction between the most “weak” triphthongs and a smoothed centring diphthong can be very slight. I have nevertheless tried to be as consistent as possible in discriminating between the two categories. All pronunciations where even a slight raising of the tongue can be distinguished have been counted as “full” variants, giving a relatively conservative estimate of the amount of smoothed variants. No attempt has been made to distinguish between monosyllabic and disyllabic variants or between diphthongal and monophthongal smoothing (cf. 4.4.5).

6.5.3 Results

6.5.3.1 General

The quantitative analysis shows that close to half the tokens are reduced by smoothing. From a total of 1339 words, 621 are realised with [aə], which amounts to 46.4% (see Table 6.5.1). The result indicates that smoothing is a common phenomenon in RP, and a well-established realisational variable. A comparison of /aɪə/ and /aʊə/ shows that FIRE words are somewhat more frequent (739 tokens) and slightly more prone to undergo smoothing, with an [aə] score of 48.7%. The 600 words representing POWER are smoothed in 43.5% of the cases.

Table 6.5.1. Smoothing: total scores

	N	%
[aə]	621	46.4
[aɪə, aʊə]	718	53.6
Total	1339	100

Although the analysis only distinguishes between presence and absence of smoothing, a brief comment should be made about the phonetic quality of the two variants. Most descriptions of RP smoothing hold that /aɪə/ usually has a more front starting point than /aʊə/, [a] versus [ɑ] respectively (cf. 4.4.5). This qualitative difference will then be preserved when the sequences undergo smoothing. No such systematic difference is observed in the present study. The quality of the first vocalic element is generally front for both FIRE and POWER, whether full or reduced. If there is a qualitative difference it is rather the other way around, that /aɪə/ has the more retracted and /aʊə/ the more front, starting point. This is in line with Upton's (2004: 225) description of the current RP PRICE vowel, which he represents as [ʌɪ], and possibly reflects a development towards London PRICE and MOUTH qualities (cf. Wells 1982: 308-309). The smoothed variant is in most cases realised as a centring diphthong [aə] for both FIRE and POWER. It is occasionally reduced to a monophthong, but usually with a front quality, avoiding a merging with the phoneme /ɑ:/. This finding seems to confirm Cruttenden's (2001: 140) prediction about RP smoothing (cf. 4.4.5).

The individual scores for smoothing, displayed in Figure 6.5, show that all the speakers display this features to some extent. There is however a great deal of inter-speaker variation, which will be discussed in the next sections.

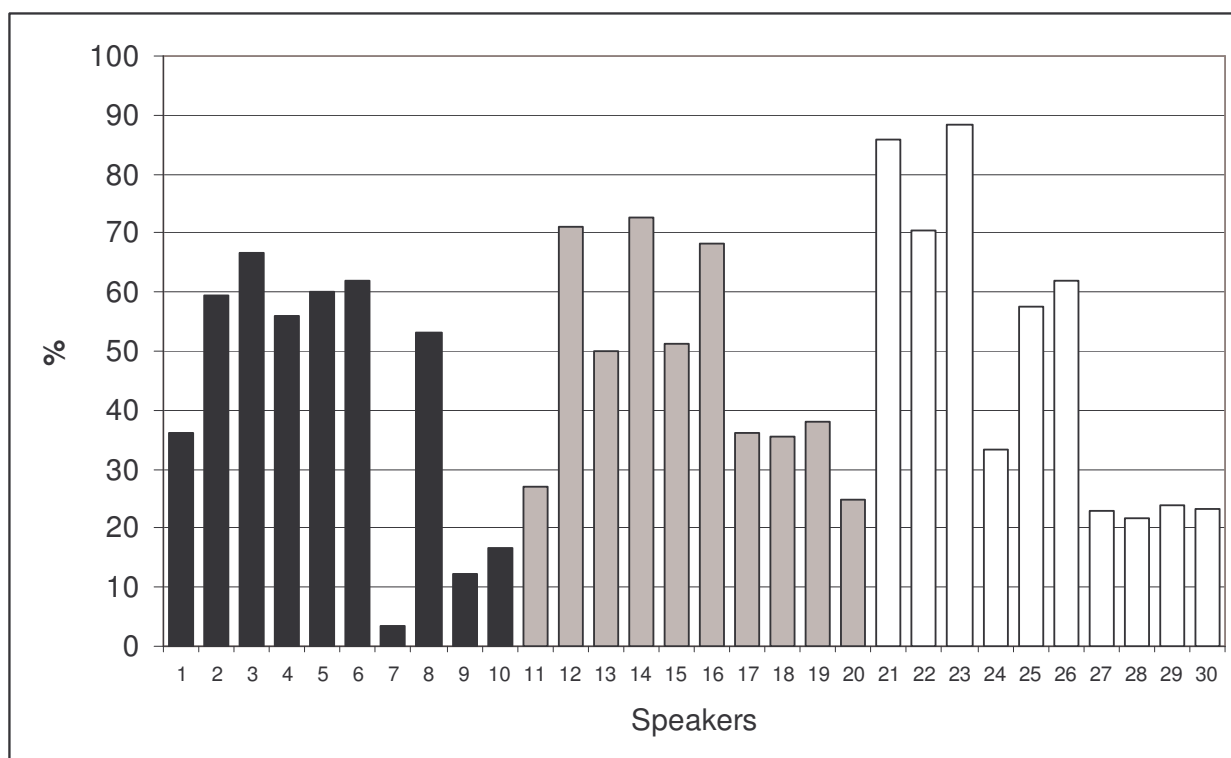


Figure 6.5. Percentage frequency of smoothed [aə], individual speakers (Black columns = BBC World, grey = Sky News, white = ITV News.)

6.5.3.2 Social factors

RP smoothing has been very little investigated, at least from a sociolinguistic point of view. The exact social connotations and evaluations of smoothing have therefore not been empirically established. It seems, however, that the phenomenon is associated with the socially conspicuous U-RP or affected, “posh” speech. This is evident from various comments in the literature on RP (cf. 4.4.5). RP smoothing has been caricatured in spellings such as “empah” for *empire* (cf. Wells 1982: 293). According to Rosewarne (1984), a posh RP speaker might produce a sentence like the following: “So tarred darling; ar harred car’s been in the mar for an ar” (‘So tired darling; our hired car’s been in the mire for an hour’). It should be pointed out, though, that this stigmatisation of smoothing pertains mainly to the more extreme monophthongised variant which results in convergence with /a:/. Cruttenden (2001: 139) notes that monophthongisation of FIRE and POWER is primarily a feature of so-called Refined RP – an upper-class variety with forms that are considered affected even by other RP speakers. The more moderate reduction to a centring diphthong is presumably widespread also in unmarked RP. This seems a reasonable conclusion based on the findings presented here, which are based on data from Mainstream RP.

Smoothing is not a new phenomenon in RP, and seems to represent a case of stable variation rather than change in progress (cf. 4.4.5). Smoothing can be viewed as an alternative pronunciation of FIRE and POWER words, which some RP speakers have as part of their phonetic inventory. Smoothing can however also be approached as a phonological process which reduces the degree of articulatory effort: the production of a full triphthong with a raised mid element involves more tongue movement than a centring diphthong or a monophthong. In this perspective, smoothing would be expected to vary according to the level of speech attention and the nature of the linguistic surroundings. These aspects will be further discussed in sections 6.5.3.3 – 6.5.3.4.

Channel

A comparison of the three channels shows very little difference in the overall mean scores (cf. Table 6.5.2). BBC World has an average slightly below the total mean, and ITV slightly above, but these are due mainly to a few extreme scores that skew the overall results (cf. Figure 6.5).

Table 6.5.2. Smoothing: mean scores for channels

Channel	Mean	N	Std. Deviation
BBC World	42,580	10	23,5885
Sky News	47,460	10	18,0599
ITV News	48,910	10	27,0129
Total	46,317	30	22,5377

The ANOVA test firmly refutes the hypothesis that the use of smoothing is in any way related to the speakers' channel affiliation. It shows that the variation within each group hugely outweighs any inter-channel differences:

Table 6.5.3. Smoothing: ANOVA analysis for channel

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	219,953	2	109,976	,205	,816
Within Groups	14510,509	27	537,426		
Total	14730,462	29			

Gender

In view of the reported connotations of smoothing as refined or affected, one would expect the female speakers to have the highest scores for [aə]. On the other hand, if smoothing is regarded as a phonological reduction process, the expectations might be reversed, as men have been found to produce more reduced forms than women, who tend to prefer “full” articulations (cf. 3.1.4). A comparison of the average scores for the two gender groups supports the latter analysis: the men have a mean score of 59%, compared to the much lower 33.6% for women (cf. Table 6.5.4).

Table 6.5.4. Smoothing: mean scores for gender

Gender	Mean	N	Std. Deviation
Males	59,040	15	17,8404
Females	33,593	15	19,6724
Total	46,317	30	22,5377

The gender contrast is greater than for any of the other variables investigated. It is also remarkably consistent: 12 of the 15 males have scores of more than 50%, while only four females have scores above the overall mean (speakers 6, 8, 16 and 26). The lowest male score is 27%, while eight women have scores below this. Not surprisingly, then, the ANOVA test shows that the gender difference is highly significant ($p = 0.001$).

Table 6.5.5. Smoothing: ANOVA analysis for gender.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4856,496	1	4856,496	13,772	,001
Within Groups	9873,965	28	352,642		
Total	14730,462	29			

The gender results for smoothing are very similar to those for t-voicing (cf. 6.4.3.2), which also reflects a phonological process of reduction or simplification. There is however an important difference between the two variables, in that t-voicing is also socially stratified, with [t̥] being more frequent with working class speakers and in casual styles. T-voicing is thus recognised as a “rough” variant, while smoothing in RP seems to have the opposite connotations. The overwhelming male dominance in the use of [aə] is therefore somewhat

surprising. One possible explanation may be found in the fact that smoothing is also a prominent feature in broad Cockney (cf. 4.4.5). Women may then be more sensitive to the “rough” connotations which stem from Cockney than men are to the “posh” associations with U-RP. Based on the references available, however, this seems a somewhat dubious interpretation, as smoothing of the present kind is never described as having vernacular associations. A more satisfactory explanation seems to be that which sees the gender patterns as reflecting a difference in the attitude towards phonetic explicitness, with women using pronunciations closer to the citation form and men being more prone to reduce. In this perspective, the results for smoothing tie in with those for t-voicing and, to a certain extent, r-sandhi. The female preference for phonetic explicitness may in turn be seen as a manifestation of the tendency for women to be more listener-oriented and concerned with the needs of the interlocutor (cf. Woods 2001), which in newscasting presumably would result in an increased focus on diction and articulatory clarity.

6.5.3.3 Linguistic factors

Since smoothing affects a relatively small number of very similar words, the amount of variation in the linguistic environment is limited. Factors believed to influence the use of smoothing include the morphological complexity of the target words and the phonetic nature of the environment immediately following the /aɪə, aʊə/ sequence (cf. 4.4.5).

Morphology

Smoothing is claimed to be particularly common in simple words, where the [ə] is an integral part of the word, and less common in morphologically complex items, where [ə] belongs to a suffix (cf. Jones 1960: 106, Hughes and Trudgill 1996: 48). The sample investigated here contains relatively few complex words, but the analysis shows a clear tendency towards keeping the triphthongs in these items: only 21 of 83 complex words are smoothed (cf. Table 6.5.6), which equals 20.2%. By comparison, the simple words undergo smoothing in 48.6% of the cases.

Table 6.5.6. Smoothing according to the morphological complexity of the target words

	Complex words		Simple words	
	N	%	N	%
[aə]	21	20.2	600	48.6
[aɪə, aʊə]	83	79.8	635	51.4
Total	104	100	1235	100

χ^2 : 31,090, $p = 0.000$, $df = 1$

The result clearly supports the hypothesis that morphologically complex words are less prone to smoothing. The items investigated here are all /aɪə/ words, such as *drier*, *highest*, *buyer*, *qualifier*, *defiant*, *alliance*. The reluctance to reduce these words is presumably morphologically motivated. The stem words constitute complete lexical units also without the [ə] (*dry*, *high*, *buy*, *defy*, etc.), whose meaning would be distorted by removing the [ɪ] element. Keeping the triphthongal realisation thus helps the receiver to decode the word structure. Morphologically simple words like *power* and *tyre*, on the other hand, have [ə] as an integrated part (there is no word *pow* or *ty*), and the triphthongal sequence cannot be divided by removing the [ə]. The mid [ɪ, ʊ] element, however, can more easily be left out in these items, since it does not mark a morpheme boundary.

Phonetic environment

According to Jones (1960: 106, 110), smoothing is particularly common before a consonant, as in *quiet*, *society*, *alliance*, *fiery*, *powerful*. There are however many /aɪə, aʊə/ items that potentially end in a vowel sound, namely those that have a final *-r(e)*, such as *hire*, *hour*, *shower*. These words will however end in a consonant when they are realised with linking /r/, as in *fire/r/ away*, *hour/r/ of*. It is thus to be expected that reduction of these words is less likely when they occur before a pause, or before a word-initial vowel without an intermediate linking /r/. Jones' claim has therefore been tested by investigating to what extent words that end in *-r(e)* are smoothed in continuous speech when the linking /r/ is present, compared to when it is omitted. The number of relevant tokens is quite small (133), but there seems to be a close correlation between smoothing and a following consonant. The figures are outlined in Table 6.5.7. The frequency of [aə] increases significantly to 73.8% in those cases where it is immediately followed by a linking [r], while the corresponding score is 37.7% when the [r] is omitted.

Table 6.5.7. Smoothing of *-r(e)* items according to the presence and absence of a following linking /r/

	+ /r/		+ Ø	
	N	%	N	%
[aə]	59	73.8	20	37.7
[aɪə, aʊə]	21	26.2	33	62.3
Total	80	100	53	100

χ^2 : 17,145, p = 0.000, df = 1

The result for smoothing and linking /r/ strengthens the notion of smoothing as a feature of continuous speech, which increases in frequency with the fluency of delivery. This aspect has been further investigated by comparing the use of [aɪə, aʊə] and [aə] before a pause. The hypothesis is that smoothing will be less common when the target word occurs at the end of a clause or before a tone group boundary, as this environment is typically lengthening and reduces the likelihood of phonological reduction (cf. Shockey 2003: 16). The analysis revealed that smoothing is significantly less likely to take place in the environment of a following pause (cf. Table 6.5.8). The [aə] score is 32.9 in pre-pausal environments, while it increases to 50.2% when the target word occurs in the middle of a phrase or sentence.

Table 6.5.8. Smoothing according to the presence and absence of a following pause

	+ pause		÷ pause	
	N	%	N	%
[aə]	98	32.9	523	50.2
[aɪə, aʊə]	200	67.1	518	49.8
Total	298	100	1041	100

χ^2 : 28,057, p = 0.000, df = 1

6.5.3.4 Speech rate

The analysis of the environment immediately following /aɪə, aʊə/ has shown that smoothing is sensitive to breaches in the flow of speech, and has strengthened the interpretation of smoothing as a feature of reduction and simplification. Smoothing might therefore be expected to correlate with extra-linguistic factors such as style and speech rate. Previous

descriptions of smoothing in Mainstream RP generally agree that the reduced forms are more likely to occur in rapid or casual speech, while the triphthongs are “slow” variants (cf. 4.4.5). Newsreading is primarily scripted speech which is performed at a rate somewhat lower than that of a spontaneous conversational (cf. 5.1.3). The overall high proportion of [aə] realisations in broadcast RP has therefore established that smoothing is common also in slow, formal speech. The number of /aɪə, aʊə/ tokens is too low for a quantitative comparison of reading style and interview style, which was performed for r-sandhi and t-voicing. Smoothing can, however, be tested for correlation with individual speech tempo, as there is considerable inter-speaker variation in rate of delivery. The hypothesis is then that the fastest speakers will have the highest percentage of smoothing. The analysis revealed that the two speakers with the highest and lowest [aə] scores (speakers 23 and 7) are both averagely fast speakers with the exact same rate of speech. Among the presenters who use smoothing most extensively we find both the fastest and the slowest speaker (14 and 12). Inter-speaker variation in speech rate thus seems at first glance to have little influence on smoothing. However, a comparison of the ten highest and ten lowest [aə] scores shows that the speakers in the former group have an average speech rate of 170,6 words per minute, while the latter group speak slightly slower, with an average of 165,5 words per minute. There may therefore be a certain correlation between smoothing and speaking rate, which was not observed for r-sandhi and t-voicing. The amount of data is however too small to draw any definite conclusions.

6.5.4 Summary

The analysis of smoothing has shown that reduction of /aɪə, aʊə/ to [aə] is very frequent in broadcast RP. In view of the nature of the newscasting situation it is evident that smoothing is not limited to rapid or casual speech. Smoothing may have contrasting social connotations, as the feature is said to be characteristic of U-RP as well as of Cockney. The significant gender difference observed here indicates that the association with U-RP either has little effect, or is limited to the extreme smoothing of /aɪə, aʊə/ to /ɑ:/. The male dominance in the use of reduced variants is perhaps best explained by regarding smoothing as a reduction or simplification process. Such an interpretation is supported by the investigation of linguistic

constraints, which shows that the use of smoothing decreases when the speech flow is interrupted.

6.6 Yod coalescence

6.6.1 Introduction

Yod coalescence involves the coalescent assimilation of /t, d/ with a following /j/, producing the affricates /tʃ, dʒ/. This section presents the results of the auditory and quantitative analysis of yod coalescence, and discusses the findings in light of the theoretical description of the variable in 4.4.6.

The present study focuses on word-internal yod coalescence in stressed syllables (e.g. *duty* /'dʒu:ti/), where the replacement of /tj, dj/ by /tʃ, dʒ/ traditionally has been excluded from RP (cf. 4.4.6). Yod coalescence in unstressed syllables (e.g. *educate* /'edʒukeɪt/, *gradual* /'grædʒuəl/, *situate* /'sɪtʃueɪt/) is firmly established in RP and is well on its way to becoming categorical (cf. Wells 1994a: 203). Coalescence before a stressed vowel represents a new development in the variety, and is not universally accepted as a feature of RP.

The analysis includes only words with primary stressed /u:, ʊə/ (e.g. *tune*, *'during*, *insti'tution*, *re'duce*). It should be noticed, however, that yod coalescence is traditionally resisted (and marked as non-RP in LPD) in all strong syllables where vowel reduction is blocked, and where some will identify a secondary or tertiary stress (e.g. *'attitude*, *'residue*, *'prostitute*) (cf. Wells 1982: 247).

Yod coalescence is a type of assimilation and as such a phonetically motivated process (cf. 4.4.6). However, it differs from the typical contextual assimilation or the coalescence that takes place across word boundaries, in that it may have a permanent effect on the phonological structure of the target words. Moreover, it represents a change “from below” in the sense that it originates in non-standard varieties. Even though yod coalescence is becoming increasingly frequent in modern RP, it still has negative connotations for many. The scholarly comments on the status of yod coalescence in relation to RP vary considerably (cf. 4.4.6), and apart from Altendorf's (2003) small-scale study of coalescence among a group of Estuary English-/near-RP-speaking girls, the subject has not been empirically investigated.

6.6.2 The auditory analysis of yod coalescence

Yod coalescence incorporates two phonological variables, (tj) and (dj), which have the variants /tʃ/, /tj/ and /dʒ/, /dj/ respectively, representing coalesced and uncoalesced realisations. /tj/ and /dj/ are sequences of two consonants, involving the alveolar plosives /t, d/ followed by the palatal approximant /j/. In the coalesced realisations the two consonants have fused into a palato-alveolar affricate.⁷⁴ The variants are auditorily quite distinct, and for the most part relatively easy to identify. The most problematic cases involved tokens where /j/ was fricated, approaching [ç, ʝ]. This occurred particularly frequently in the case of /tj/, where /j/ is devoiced. However, the /tʃ, dʒ/ variants remain distinguishable from /tj, dj/ by the fact that they involve a sibilant sound [ʃ, ʒ], and therefore more “hissing” noise. Also, the alveolar consonant is more prominent in /tj, dj/ than in the coalesced variants, where it merges with /j/. In doubtful cases, the relevant items were isolated and compared with each other as well as with distinct “prototypical” /tj, dj/ and /tʃ, dʒ/ tokens. In order to secure a consistent analysis, the coalesced variants were compared with words that contain palato-alveolar affricates in their citation form. Words such as *choose*, *Jewish* and *jury* were thus used to provide references for /tʃ/ and /dʒ/. No instances of yod dropping were recorded.

6.6.3 Results

6.6.3.1 General

The number of tokens with potential yod coalescence is relatively small. A total of 617 words were identified, which constitutes an average of just over 20 tokens per speaker. 13 speakers have less than 20 tokens, while 4 have less than 15, though none fewer than 10. The basis for quantitative analysis is in other words somewhat limited, but still large enough to reveal some important tendencies in the use of yod coalescence in RP. The main observation is that /tʃ, dʒ/ are quite common in stressed syllables. The analysis shows that, overall, the coalesced variants are used in 46.4% of the cases, as outlined in Table 6.6.1. This is a strikingly high figure, in view of the status of stressed yod coalescence in RP and the formality of the speech

⁷⁴ It is debatable whether affricates should be analysed as one or two consonants (cf. 4.4.6). A biphonemic interpretation regards an affricate as a sequence of plosive plus fricative. The realisation of (tj), (dj) as /tʃ, dʒ/ would then be analysed as a progressive assimilation, where only /j/ is modified under the influence of preceding /t, d/. In the present investigation, /tʃ, dʒ/ will be regarded as single phonemes.

situation (cf. 4.4.6), and clearly suggests that yod coalescence has penetrated the boundaries of RP.⁷⁵

Table 6.6.1. Yod coalescence: total scores

	N	%
/tʃ, dʒ/	286	46.4
/tj, dj/	331	53.6
Total	617	100

The individual percentage scores for the coalesced variants are shown graphically in Figure 6.6 below. It reveals that all the speakers use /tʃ, dʒ/ to some extent in stressed syllables, though there is a great amount of variation. It should be noted that the great range in percentage scores is partly a consequence of the limited number of tokens, such that a percentage difference of 50 may reflect a usage difference of less than 10 tokens. Speaker 4 has the lowest score of 5.9%, while speaker 25 uses the coalesced variants on all occasions. With the exception of speaker 25, then, all speakers vary between /tj, dj/ and /tʃ, dʒ/.

A comparison of Figure 6.6 with the individual scores for t-voicing (Figure 6.4) reveals that the use of yod coalescence does not covary with the use of [t̚], in that high usage levels for one does not systematically entail high usage levels for the other. One might expect some covariation between yod coalescence and t-voicing, since they reportedly have similar social and stylistic connotations. There is, however, no apparent correlation between the two, as speakers who avoid [t̚] may have high scores for /tʃ, dʒ/, and vice versa (see further 6.7).

⁷⁵ It is noteworthy that no instances of coalescence of /s, z/ and /j/ were observed. Words like *assume*, *presumably* and *consumer* are generally pronounced with /sju:, zju:/, though a few cases of yod dropping were recorded. It is evident from this that yod coalescence involving the alveolar fricatives does not have the same status as coalescence of /tj, dj/. It remains to be seen whether yod coalescence will eventually affect these items, too, or whether yod dropping will prevail, as it has in words like *super*, *suitable*, etc.

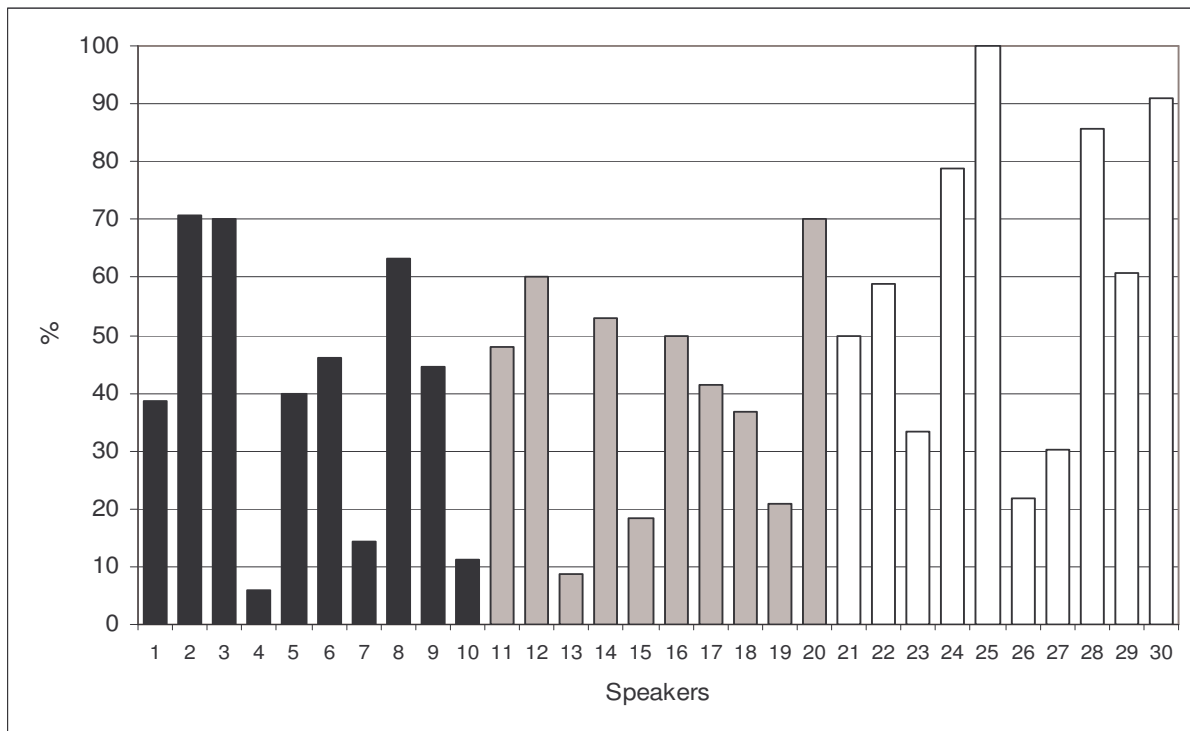


Figure 6.6. Percentage frequency of /tʃ, dʒ/, individual speakers (Black columns = BBC World, grey = Sky News, white = ITV News)

6.6.3.2 The lexical distribution of yod coalescence

There are altogether 25 different words with potential yod coalescence before a stressed /u:/, uə/. The words are listed in Table 6.6.2 below, along with the number of times they are realised with /tʃ, dʒ/ versus /tj, dj/.

Table 6.6.2. Yod coalescence: the lexical distribution of /tʃ, dʒ/ and /tj, dj/

word	/tʃ, dʒ/	/tj, dj/	word	/tʃ, dʒ/	/tj, dj/
<i>constitution</i>	3	26	<i>reduce</i>	14	6
<i>deduce</i>	-	1	<i>Stuart</i>	6	6
<i>dual</i>	1	-	<i>student</i>	8	20
<i>due</i>	33	63	<i>studio</i>	8	46
<i>duke</i>	2	1	<i>subdue</i>	2	1
<i>during</i>	142	29	<i>tube</i>	6	2
<i>duty</i>	9	19	<i>Tudor</i>	2	-
<i>endure</i>	2	1	<i>Tuesday</i>	14	26
<i>institution</i>	1	7	<i>tumour</i>	1	2
<i>introduce</i>	6	23	<i>tune</i>	1	8
<i>opportunity</i>	9	14	<i>Tusa</i>	6	6
<i>produce</i>	10	18	<i>undue</i>	-	4
<i>prostitution</i>	-	2			

It is evident from the lexical distribution of yod coalescence that the overall /tʃ, dʒ/ score of 46.4% is somewhat inflated due to the high frequency of the word *during*. Almost half of the coalesced items are realisations of this word, which is extremely frequent and is pronounced with an affricate in 83% of the cases.⁷⁶ If the *during* items are removed from the total, the score for the coalesced variants is reduced to 32.3%. This is nevertheless still a fairly high score, and leaves no doubt that yod coalescence in stressed syllables is becoming established in RP speech.

Although yod coalescence is traditionally judged as falling outside RP, *during* constitutes a special case. In earlier editions of EPD (from 1967 and up until the fifteenth edition in 1997) *during* is the only word where yod coalescence is accepted as an alternative before a stressed vowel. In LPD (Wells 2000), /tʃ, dʒ/ in stressed syllables is marked as non-RP, except in the word *during*. It seems, therefore, that *during* has occurred with an affricate for a long time, and may be leading a process of lexical diffusion (cf. 3.1.2). It is by far the most frequent of the items in Table 6.6.2,⁷⁷ and it is the only grammatical word with potential yod coalescence. It is therefore a perfect candidate for spearheading a frequency-based lexical diffusion of a reductive change. The results from the present study confirm the exceptional status of *during*, as this word is pronounced with yod coalescence more often than any other item. 21 of the speakers categorically use an affricate in *during*, while the remaining nine vary between /dʒ/ and /dj/. Six of the presenters (1, 4, 7, 10, 13 and 21) have yod coalescence only in *during* and use /tʃ, dj/ in all the other target words.

Apart from the frequent use of /dʒ/ in *during*, there are no particular patterns in the lexical incidence of yod coalescence. Only three items are never realised with an affricate (*deduce*, *prostitution*, *undue*⁷⁸), but these words only appear between one and four times in the corpus. It is noteworthy that yod coalescence also occurs after /s/, in words like *student*, *studio*, *Stuart*.⁷⁹

⁷⁶ In view of the high frequency of /dʒ/ in *during* observed here, it is interesting to note the results from Wells' opinion poll survey of pronunciation preferences (Wells 1999a, 2000), where 65% of the respondents report that they prefer /dj/ in *during*.

⁷⁷ *During* is listed with a frequency of 440 per million words in Leech et al. (2001).

⁷⁸ It is notable that all these words are relatively rare, and are not even included in the BNC frequency list (Leech et al. 2001), which means that they have less than one occurrence per million words.

⁷⁹ In /stj/ clusters there is the alternative possibility of so-called ST palatalisation, which produces /ʃtj/ or /ʃtʃ/ (cf. Altendorf 2003: 69-70). This palatalisation seems, however, to be restricted to non-RP, especially Estuary English, speech (Altendorf 2003: 101, Coggle 1993: 51).

At the level of the individual speakers, the lexical distribution of yod coalescence is somewhat inconsistent. Most of the speakers vary between /j/ and coalescence in at least one lexical item. This vacillation is seemingly random, and suggests that, with the possible exception of *during*, yod coalescence is not associated with particular words, or with particular phonetic surroundings (in contrast to, for example, CURE lowering (cf. 6.1)). The intra-speaker variation can be seen in relation to the fact that yod coalescence is a mechanism for improving articulatory ease (cf. 4.4.6), and as such will typically sometimes be applied and sometimes not, under the same circumstances. In this sense, yod coalescence reflects much of the same usage patterns as t-voicing (cf. 6.4) and r-sandhi (cf. 6.3).

6.6.3.3 Social factors

In terms of social evaluation, the coalesced and uncoalesced variants do not have the same status. The “traditional” /tj, dj/ variants are the prestige forms, while /tʃ, dʒ/, being new in RP, do not have the same status, and are by many associated with non-standard accents, notably Estuary English (cf. 4.4.6). Wells (1982: 331) describes /tj, dj/ as the “elegant” alternatives in words like *tune* and *due*, and yod coalescence is characterised as belonging to casual speech and avoided in “careful RP” (1982: 207, 2000: 50).

The results presented here show, first of all, that yod coalescence is not confined to non-RP accents, and further that, within RP, it is not exclusive of casual style. It is, however, reasonable to assume that the use of /tʃ, dʒ/ would be even greater in informal contexts, in view of its social and stylistic connotations (cf. 4.4.6). Because of the relatively small number of tokens, it has not been possible to compare reading style and interview style, and the style hypothesis could therefore not be tested for (tj) and (dj).

Channel

The social and stylistic connotations of yod coalescence would lead us to expect the highest /tʃ, dʒ/ scores for the presenters on Sky News, which is the most “informal” channel (cf. 5.1.2). The analysis shows, however, that, somewhat surprisingly, it is the ITV presenters who have the highest usage scores for yod coalescence, with an average of 61.1%, while BBC World and Sky News have virtually identical scores of 40.4% and 40.7% respectively (cf. Table 6.6.3).

Table 6.6.3. Yod coalescence: mean scores for channels

Channel	Mean	N	Std. Deviation
BBC World	40,430	10	23,7953
Sky News	40,710	10	19,5736
ITV News	61,050	10	27,3366
Total	47,397	30	24,9595

The high ITV scores are mainly due to contributions by a few speakers who then skew the overall results (cf. Figure 6.6). The ANOVA test results, outlined in Table 6.6.4, reveal that the inter-channel difference is not statistically significant ($p = 0.103$), as the individual variation is greater than the contrast between the channels.

Table 6.6.4. Yod coalescence: ANOVA analysis for channel

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2796,595	2	1398,297	2,472	,103
Within Groups	15269,715	27	565,545		
Total	18066,310	29			

It is noteworthy, nevertheless, that four of the ITV newsreaders have /tʃ, dʒ/ scores above 75%, and that none have scores below 20%, while among the other presenters there are five speakers with scores below 20%, and only one above 70% (cf. Figure 6.6). The results contrast with those for t-voicing (cf. 6.4.2.2), where the Sky presenters had the highest scores. This indicates, again, that there is apparently no correlation between the use of the two variables, and that yod coalescence has a different status.

Gender

Yod coalescence is seemingly a change towards a more vernacular norm. The /tʃ, dʒ/ variants represent traditional, high-status pronunciation, while the coalesced forms have entered RP from below, in the sense that they have their roots in non-standard speech. Our hypothesis is therefore that the male speakers will use more yod coalescence than the females, based on the typical linguistic behaviour of men and women (cf. 3.1.4). A comparison of the two groups revealed, however, that they are very similar in their use of /tʃ, dʒ/. The mean scores are

48.9% for the men and 45.9% for the women (cf. Table 6.6.5), a difference which is far from statistically significant (cf. Table 6.6.6).

Table 6.6.5. Yod coalescence: mean scores for gender

Gender	Mean	N	Std. Deviation
Males	48,940	15	26,0308
Females	45,853	15	24,6524
Total	47,397	30	24,9595

Table 6.6.6. Yod coalescence: ANOVA analysis for gender

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	71,456	1	71,456	,111	,741
Within Groups	17994,853	28	642,673		
Total	18066,310	29			

The similarity between the two genders in their use of /tʃ, dʒ/ is again in contrast to the findings for t-voicing and smoothing, which both showed a significant male preference for the marked variants (cf. 6.4.3.2 and 6.5.3.2). Women do not avoid coalesced forms more than men, and some of the female presenters who have extremely low scores for t-voicing, intrusive /r/ and smoothing (notably speakers 29 and 30), have an extensive use of yod coalescence.

The fact that men and women have the same usage level for yod coalescence may indicate a shift in the evaluative status of this feature. Yod coalescence may be moving towards a prestige variant, much in the same way as t-glottalling. The status of the glottal stop has changed from a stigmatised vernacular form to a “fashionable”, supra-local middle-class variant in the course of the last decades. Females are reported to be leading the implementation and diffusion of t-glottalling, and possibly instrumental in the social re-evaluation of the feature (cf. 3.1.4). Of the six variables studied here, yod coalescence and GOAT allophony are the only ones that are commonly referred to as typical Estuary English features (cf. 4.2.3). With the rise of Estuary English as a new, popular “standard”, yod coalescence (and GOAT allophony) may, on some level, be associated with prestige and “modernity” in a way that does not apply to t-voicing and smoothing.

6.6.4 Summary

The analysis of yod coalescence has shown that forms such /tʃu:n/ *tune* and /dʒu:/ *due*, are quite common in broadcast RP, although the pronunciations with /j/ are still in a majority. All the speakers use yod coalescence to some extent, and there is no difference between the genders. The lexical distribution of /tʃ, dʒ/ versus /tj, dj/ revealed that yod coalescence is particularly common in the high-frequency word *during*, which may be spearheading a frequency-based lexical diffusion.

The relatively frequent use of coalesced forms among the news presenters in the present study is a clear indication that yod coalescence in stressed syllables is becoming an established feature in RP, and is not restricted to casual style. The high /tʃ, dʒ/ scores are particularly striking in comparison to Altendorf's (2003) findings for teenage speakers from London, Colchester and Canterbury. Her informants, who are all female pupils with Estuary English or near-RP accents, use yod coalescence in less than 10% of the cases (2003: 98). The results from the present investigation support previous claims that yod coalescence is entering Mainstream RP speech and should be included in updated descriptions of the accent (cf. Upton et al. 2001: xiii, Cruttenden 2001: 83), and prove that Wells' (1994a: 203) prediction that coalescence in stressed syllables "may penetrate RP within a few decades" was well-founded.

6.7 Covariation

The present section expands the results discussed above by exploring the correlations between the linguistic variables. Features that are thought to represent similar phenomena or have similar evaluative status, were expected to show patterns of covariation, where high usage levels for one variant corresponds to high levels for the other. Yod coalescence, for instance, is likely to co-occur with GOAT allophony, since they both represent typical Estuary English features. High levels for these may also correlate with frequent use of "progressive" t-voicing, but probably not with the "conservative" smoothing of /aɪə, aʊə/. Linking and intrusive /r/ both represent consonant epenthesis, and as such are predicted to covary. Intrusive /r/, however, is widely stigmatised, and may therefore be more frequent with speakers who also

have high scores for t-voicing and yod coalescence, which both have connotations of vernacular or casual speech.

A comparison of the individual scores for these variables shows that the usage patterns do not always display the expected covariation. Table 6.7.1 outlines the various usage levels for each of the 30 speakers. It gives the percentage frequencies for r-sandhi, t-voicing, smoothing and yod coalescence, and indicates use/non-use/variable use of GOAT allophony.

Table 6.7.1. Overview of individual frequencies for variables 2-6

Speaker	GOAT allophony	Linking /r/	Intrusive /r/	T-voicing	Smoothing	Yod coalescence
1	+	53.2	16.7	26.7	36.2	38.5
2	-	61.6	48.1	41.1	59.3	70.6
3	+	50.0	13.0	48.1	66.7	70.0
4	variable	42.1	11.5	36.6	56.1	5.9
5	+	68.7	42.9	27.6	60.0	40.0
6	+	59.1	42.3	42.4	62.1	46.2
7	variable	70.3	69.8	18.6	3.4	14.3
8	+	71.4	50.0	23.1	53.1	63.3
9	variable	53.1	16.7	12.3	12.2	44.4
10	+	67.4	-	21.5	16.7	11.1
11	+	62.2	50.0	31.8	27.0	48.0
12	+	79.5	73.3	51.7	71.1	60.0
13	+	48.4	6.7	46.6	50.0	8.7
14	+	77.2	60.0	41.8	72.7	52.9
15	+	75.2	27.8	58.8	51.2	18.5
16	+	55.3	31.8	44.3	68.2	50.0
17	+	59.3	27.3	55.8	36.1	41.4
18	+	57.6	25.0	34.8	35.4	36.8
19	+	48.7	44.4	27.2	37.9	20.8
20	+	55.5	28.6	19.6	25.0	70.0
21	variable	52.9	6.9	41.1	85.7	50.0
22	+	56.9	12.5	39.3	70.3	58.8
23	+	66.7	50.0	56.5	88.4	33.3
24	+	69.9	23.1	40.3	33.3	78.9
25	+	67.9	83.3	50.0	57.6	100
26	+	69.1	-	13.9	62.1	21.7
27	+	67.1	57.1	26.7	22.9	30.3
28	-	52.1	35.3	33.5	21.6	85.7
29	+	35.9	6.3	16.3	23.8	60.9
30	+	25.3	-	9.7	23.4	90.9
Average		59.3	32.0	34.6	46.3	47.4

GOAT allophony is used almost categorically and is therefore not so interesting from a variational perspective. The other Estuary-feature, yod coalescence, is used much less frequently and more variably. This suggests that the two variables do not have the same status, in that GOAT allophony is less marked and more generally “accepted”. (It is noteworthy, however, that the two speakers who do not have [ɒʊ] (2 and 28) have very high scores for yod coalescence.) More importantly, though, both features are used more frequently than expected for RP, which suggests an increased approximation to Estuary English (see further 7.6).

A few speakers (1, 4, 9, 10 and 26) stand out as typically “conservative”, with overall low scores, except, in some cases, for linking /r/ and smoothing. At the other extreme, speakers 2, 12, 14, and 25 stand out with high scores for all variants. On the whole, however, usage levels do not follow any systematic patterns. A bivariate correlational analysis reveals that very few of the variables show a significant degree of covariation. Table 6.7.2 outlines the results of a Pearson correlation test, comparing usage levels for all the quantified variables.

Table 6.7.2. Bivariate correlational analysis of variables 3-6
 N = number of speakers, R = Pearson’s correlation coefficient. Significant p-values are set in bold.

Variables		N	R	p-value
Linking /r/	Intrusive /r/	30	0.602	0.000
	T-voicing	30	0.342	0.064
	Smoothing	30	0.239	0.203
	Yod coalescence	30	-0.136	0.474
Intrusive /r/	T-voicing	30	0.298	0.109
	Smoothing	30	0.104	0.584
	Yod coalescence	30	0.192	0.310
T-voicing	Smoothing	30	0.627	0.000
	Yod coalescence	30	0.026	0.891
Smoothing	Yod coalescence	30	0.054	0.776

Not unexpectedly, there is a significant positive correlation between linking and intrusive /r/, in that speakers with above average scores for one usually have above average scores for the other. Notable exceptions are speakers 10, 15, 24 and 26, who use considerably less intrusive /r/. Such deviations can be explained by reference to the contrasting social evaluation of the two /r/s. Intrusive /r/ does, however, not show any regular co-occurrence

with t-voicing or yod coalescence. The Pearson correlation test shows that these variables are not significantly related. Yod coalescence seems in general to have a somewhat different status than the other non-traditional RP features. A few speakers (20, 28, 29 and 30) have overall low scores for the other variants, but strikingly high scores for yod coalescence. It may be, of course, that these presenters are adoptive RP speakers and have yod coalescence from their native accents, but lack control of the typical informal and connected speech features of native RP (cf. Wells 1982: 284). Lastly, and somewhat surprisingly, there is a high degree of covariation between t-voicing and smoothing ($r = 0.627$), which may be related to the fact that both are processes of phonological reduction.

Based on the present sample, then, it seems that shared social evaluation is a fairly poor predictor for co-occurrence, and that pronunciation variables rather tend to covary if they reflect related phonological processes.

7 Conclusions

7.0 Introduction

This chapter presents a summary of the results of the analyses in chapter 6, and considers the empirical and theoretical implications of the data. It attempts to unite the results for the six variables into a coherent account of the current status of these features in broadcast RP. It discusses how the present findings relate to the more large-scale tendencies of accent levelling, and how they can contribute to an updated description of RP. Lastly, it points to some possible implications of the present study for the teaching of English as a foreign language, and proposes suggestions for further research.

7.1 Summary of results

The empirical analysis of the six variables has shown that RP displays considerable variability and that the phonetic properties of the accent have not remained static.

The results for CURE lowering showed that the replacement of /ʊə/ by /ɔ:/ is far from completion in RP. All the presenters used both vowel phonemes in CURE, and the majority of the words were realised with /ʊə/. The two vowels seem, however, to be largely conditioned by the phonetic environment, as /ʊə/ is retained after consonant + /j/, as in *pure*, *secure*, and before prevocalic /r/, as in *Europe*, *tourist*, while /ɔ:/ dominates in other environments. There was also some evidence to suggest that certain words may transfer to /u:/ rather than to /ɔ:/.

While the analysis of CURE lowering revealed rather conservative tendencies, the analysis of GOAT allophony showed the opposite result. The use of [ɒʊ] for /əʊ/ before non-prevocalic /l/ seems to be firmly established in broadcast RP. 24 of the 30 newsreaders had categorical allophony, and only 2 speakers used [əʊ] in all GOAT words. There was also evidence of a potential phoneme split, as some speakers used [ɒʊ] also in derivative forms such as *wholly* and *controlling*, where the /l/ is prevocalic.

The analysis of r-sandhi showed a surprisingly low overall score of 59.8% for linking /r/, while intrusive /r/, which one might expect to be suppressed in newsreading, was used in as much as 32.6% of the cases. The use of sandhi /r/ seems to be influenced by several, partly competing, factors. R-sandhi is a phonetically motivated connected speech process which enhances articulatory ease. Both linking and intrusive /r/ displayed strong conditioning by the linguistic context, motivated largely by prosodic and articulatory considerations. At the same time, the use of r-sandhi appears to be directly linked to the level of speech attention. This was evident from the increased frequency of linking and intrusive /r/ in interview style compared to reading style. The greater avoidance of intrusive /r/ in reading style can be explained by reference to the stigmatisation of the r-intrusion. The suppression of “correct” linking /r/s may be the result of newscasters’ focus on clarity: there may be a general conception that separating words by leaving out the r-link leads to greater articulatory distinctness. The systematic avoidance of /r/ before proper names further suggests that the “naturalness” of r-sandhi may be overridden by the wish to secure comprehensibility.

The findings for t-voicing showed an overall usage level of 35.2%, which clearly demonstrates that [t̥] is a common realisational variant for /t/ in RP, and that this feature is not limited to vernacular speech or casual style. The morphological analysis further showed that voicing of intervocalic /t/ is primarily a feature of continuous speech: the percentage score for [t̥] across word boundaries was a staggering 66.9%, which establishes t-voicing as one of the most common connected speech processes in modern RP. T-voicing was largely resisted word-internally, except in certain frequent and informal items, such as *British*, *getting*, *gotta*. All the presenters used t-voicing to some extent, but, as expected, there was a significant increase in the less formal interview style, as well as a significant male preference for [t̥].

Smoothing is a familiar feature of RP that is mentioned already in the earliest descriptions of the accent. The present analysis showed that the reduction of /aɪə, aʊə/ remains a common process in RP, as almost half the FIRE/POWER tokens were realised with [aə]. All the speakers used smoothing to various degrees, but there was again a significant male dominance, which was somewhat surprising in view of the association of smoothing with conservative speech. The investigation of linguistic constraints revealed that smoothing is primarily a feature of continuous speech, as its frequency significantly increased with the fluency of delivery and decreased before a pause or other interruptions of the speech flow.

The analysis of yod coalescence showed that the shift from /tj, dj/ to /tʃ, dʒ/ in *tube*, *due*, etc. is becoming established in broadcast RP. The overall score for /tʃ, dʒ/ was 46.4%,

which is strikingly high in view of the “dubious” status of stressed yod coalescence, and the formality of the speech situation. The lexical analysis of yod coalescence showed evidence of frequency-based diffusion, with *during* in the clear lead with a /dʒ/ score of 83%. All the speakers displayed yod coalescence to some extent (though some only in *during*), but there was no significant gender difference.

The correlational analysis of the quantified variables showed significant association only between linking and intrusive /r/ on the one hand and between t-voicing and smoothing on the other, which suggests that features which reflect similar phonological processes (epenthesis and reduction, respectively) tend to covary regardless of any contrasting evaluative status.

Among the most important theoretical contributions have been the discussion of how RP should be defined, and the investigation of the norm concept. I have argued that RP is best defined in linguistic terms as an accent which excludes localisable features, and shown that such a definition can be operationalised. The small-scale attitudinal survey among a group of BBC World presenters demonstrated that RP still has a great amount of prestige, attested by the fact that two of the presenters had abandoned their native accent in favour of RP. The BBC interviews further suggested that the widespread use of RP in newscasts is the manifestation of an internalised norm, partly motivated by ‘audience design’.

7.2 Channels

The analysis revealed that the stylistic differences between the three channels did not significantly affect the usage levels for the variables investigated here. The Sky News presenters typically had higher usage levels for the “progressive” variants, but this was not a consistent pattern. There was thus little evidence of ‘station styles’ such as Bell (1982) observed for New Zealand radio stations, where presenters “clustered” around channel mean frequencies for certain variables, minimising individual differences.

The lack of significant differences between the channels strengthens the notion that the speakers represent a homogeneous speech community, sharing “abstract patterns of variation” (Labov 1972: 121), and suggests that they have essentially the same norms.

7.3 Gender

The study revealed some interesting differences between the usage patterns of male and female presenters, which did not always match the established hypotheses. On the basis of the typical linguistic behaviour of men and women, established in numerous sociolinguistic studies, the females speakers were expected to have higher usage levels for the more conservative or prestigious variants, and lower levels for those at the progressive or vernacular end. This hypothesis was, however, only partly confirmed. There was no difference in the use of /ʊə/ versus /ɔ:/ in CURE, which suggests that the two variants have an equal evaluative status. The distribution of the vowels seemed to be determined primarily by the phonetic context. Somewhat more surprising was the lack of gender difference in the use of the “new” features GOAT allophony and yod coalescence. [ɒʊ] and /tʃ, dʒ/ had a high overall frequency and were not particularly avoided by the female speakers. This may indicate that these features have undergone an evaluative shift, and are becoming established as part of a modern “standard”. The only marked gender contrast was found with t-voicing and smoothing and, to a certain extent, r-sandhi. I have argued that these gender differences are more satisfactorily handled in terms of phonological explicitness than by explanations in terms of overt and covert prestige. The male scores for [t̚] and [aə] were significantly higher than the female scores, even though the features are thought to have different social connotations. The difference in the use of linking and intrusive /r/ was not significant, but there was a clear tendency for the males to use more of both variants. T-voicing, smoothing and r-sandhi are all phonetically motivated processes that are characteristic of continuous speech and enhance the ease of articulation. This suggests that men are more open to “natural” phonological processes and tend to be articulatorily more economical than women, who are more “careful” speakers. This gender contrast may be interpreted as a manifestation of the tendency for women to be more addressee-oriented (cf. 3.1.4). In face-to-face interaction, women have been found to be more cooperative and concerned with their interlocutor’s needs. Transferred to the newscasting situation, a high focus on the listener will presumably lead to increased focus on diction and phonetic explicitness.

For the variables and speakers studied here, then, the gender patterns do not necessarily reflect the variants’ degree of conservatism or prestige. There is, however, a certain correlation between articulatorily motivated processes and social prestige. Phonetic explicitness is often linked with “correctness” and high-status varieties, while phonetic

reduction or simplification may have associations of “sloppiness” or vernacular speech (cf. 3.1.5.3). In this perspective, the present results can be seen to reflect the familiar Sex/Prestige pattern.

7.4 Style

This descriptive account of RP is based solely on speech produced in the relatively formal setting of newscasting. Generalisations can therefore not automatically be made to RP as spoken in other situations. The comparison of reading style and interview style showed that there is intra-speaker variation, and that the speakers are not static in their speech patterns. The usage levels for r-sandhi and t-voicing both increased significantly when the focus was no longer uniquely on the presentation of news. It is therefore reasonable to assume that the results outlined here (at least for these two variables) are a conservative estimate of the usage levels for RP in general. Moreover, this stylistic contrast clearly demonstrates that the level of speech attention and self-monitoring directly affects pronunciation.

At the same time, the findings illustrate how features such as t-voicing and yod coalescence seem to have expanded their stylistic territory, and how notions of “appropriateness” and formality in broadcasting have changed over time.

An investigation of formal, public style is perhaps particularly relevant for RP. As a codified variety, RP is inherently linked to issues of ‘correctness’ and ‘standardness’. It is therefore particularly interesting to gain insight into broadcast RP produced in a context where the focus on such aspects is great.

7.5 Variation and change in RP

Because of the codification of RP, the image given of the accent is often monolithic and static. The codified version of the accent necessarily lags somewhat behind the newest trends and does not reflect all the variability of native speech. As a prestige variety, however, native RP may also be less prone to change than other varieties. Kroch (1978) claims that prestige accents are more resistant to phonetically motivated variation and change because their speakers wish to distance themselves from speakers of low-status accents. Also, prestige

speakers, as the most educated group in society, have traditionally been more influenced by the written language and by the literary tradition to resist change. Kroch further argues that prestige varieties are spoken by social elites whose members tend to be conservative, and that “[t]he use of conservative linguistic forms is for them a symbol of their whole value system” (1978: 30). Conversely, there may be reasons to expect *more* variation and change in modern RP than in many other accents. RP is a supra-regional variety with native speakers in all parts of Britain. RP speakers do then not form a close-knit local community of the type that is reported to inhibit linguistic change (Milroy 1987). The network model predicts that change enters a variety to the extent that its speakers develop weak-tie contacts with speakers of other varieties. According to Milroy (2001b), “[t]he very noticeable social changes that have more generally come about since 1960 or so ensure that numerous weak-tie conditions are present for changes to enter RP ... more rapidly than before” (29-30). Moreover, the reduction in status that RP has undergone in recent decades may leave the accent more open to change. Milroy (2001b: 32) argues that, as a minority accent, RP is particularly vulnerable to change when the favourable social conditions that helped sustain its dominant position have declined.

Establishing the reality of either of these claims would require more empirical data. However, my study of broadcast RP has revealed a substantial amount of variability even within an extremely homogeneous group of speakers. There is therefore little reason to believe that modern RP is less susceptible to variation and change than other varieties.

Even if RP in general is undergoing considerable change, it is nevertheless reasonable to assume that the RP of newscasts lags somewhat behind conversational RP, because of the formality of the speech situation and the fact that most of the speech is scripted. Many of the tendencies observed in the present study do, however, represent a deviation from the standard descriptions of RP, and suggest that the accent has undergone several changes, primarily at the realisational level.

The current changes in RP seem to be both internally and externally motivated. CURE lowering can be seen as part of a general tendency to monophthongise the RP centring diphthongs, combined with the low functional load of /ʊə/. T-voicing is a lenition feature that can be explained by the principle of least effort. The spread of t-voicing may also be attributed to the increasing influence of American English (cf. Wells 1982: 250, Ramsaran 1990b: 179, 183). GOAT allophony and yod coalescence are probably imported into RP from external sources. They are both said to be typical of Estuary English, and represent recent trends in RP. These “new” features are, however, also found in other varieties from different parts of Britain, which is crucial to their inclusion in RP as defined here.

Much of what has been referred to as “Estuary English” merely represents a continuation of the old tendency for features of popular London or southeastern English to spread socially and geographically. This phenomenon is today typically related to a more general, large-scale process of accent levelling that is currently taking place in Britain. Some of the contemporary changes in RP (including features not studied here, such as t-glottalling and l-vocalisation) can be seen as part of this levelling process, in that they reduce the linguistic distance between RP and other varieties.

The levelling process typically involves the reduction or disappearance of features that are “marked” or in a minority (cf. Trudgill 1986). Many traditional RP features are undoubtedly in a minority, and may in some contexts be stigmatised as “posh”. We may then see a development where, for instance, [əʊ] in *goal*, /dj/ in *reduce* and [t] in *British* will be socially “marked” as distinctly conservative or upper-class, and therefore avoided.

If RP is defined as a non-localisable accent, it is to be expected that RP will widen its range “downwards” as pronunciation features with lower-class origins achieve nationwide distribution. It is not a new phenomenon that features of non-standard origins make their way into the prestige variety (cf. Mugglestone 1995). Labov has shown that linguistic innovations usually do not originate in the highest classes, but tend to arise in the middle section of the social hierarchy, in the upper working class or lower middle class (Labov 2001). And Milroy (2001b: 18) states:

the replacement of prestige forms from external (lower prestige?) sources rather than through internal change has analogues in the earlier history of English. It seems to have happened before, and so it may well be happening again.

It is of course debatable how much an accent can be modified before it is no longer the same accent. Milroy (2001b) and Kerswill (2001), for example, argue that, in the current climate of accent levelling, with the rise of new “standards” and decline of RP’s social status, the RP accent is disappearing. I will argue that, if defined as a non-localisable variety, RP continues to exist as a linguistic entity (if not as a social institution), but that its phonological ingredients are changing.

7.6 Updating RP

The present findings may be seen to have several implications for the phonological description of current RP. The six variables studied here have never previously been investigated in RP on such a large scale. The results confirm many previous observations and claims about RP, but also add new insights. And, more importantly, while previous statements are tentative or impressionistically based, the present study provides empirical, quantified evidence.

The most important finding in this perspective is the prevalence of GOAT allophony, yod coalescence and t-voicing, which conventionally are excluded from RP descriptions. The empirical analysis has demonstrated that these non-traditional features should be part of the modern descriptions, and I have argued that we can still justify calling the accent *RP*. The extensive use of GOAT allophony and yod coalescence brings RP one step closer to Estuary English, and supports Wells' (1997b) statement that RP should be upgraded by gradually incorporating features typical of Estuary English, as this is the direction in which RP seems to be heading. CURE lowering, on the other hand, is a change in slow progress, and /ʊə/ seems to retain its phoneme status in RP for some time yet, despite recurrent statements to the contrary. Similarly, the frequency of smoothing seems to have remained stable in RP, refuting Gimson's (1980: 140) prediction that /aɪə, aʊə/ would fall together with /ɑ:/ in the course of the twentieth century. With regard to r-sandhi, the study has demonstrated a significant style shifting for linking and intrusive /r/, which has not previously been established for RP. Also, the low usage level for linking /r/ suggests that this variant is not as categorical in RP as implied in many previous accounts, at least not in formal contexts.

These empirical results should be incorporated into an updated phonological account of RP. Furthermore, the theoretical discussions of RP have illustrated the importance of keeping the definition of the accent updated and in line with social changes, and underlined the value of specifying which definition is the basis for descriptions of RP or for discussions of its status.

7.7 Pedagogical aspects

RP has always had special relevance for foreign language pedagogy, in the teaching of English in general and phonetics in particular. It can of course be questioned whether RP is the most suitable model for imitation and pronunciation teaching, and many writers have argued for the introduction of new educational standards (cf. e.g. Modiano 1999, Macaulay 1988). The main argument in favour of RP is its supra-regional character and its wide prestige (cf. e.g. Ramsaran 1990b, Nolan 1999). RP is, moreover, still the variety represented in pronunciation dictionaries (and other dictionaries), and used as a reference in the description of other varieties. The accent will therefore inevitably be encountered by foreign learners of English.

To the extent that RP has continued relevance for the teaching of English as a foreign language, it is important to keep the descriptions updated. Wells (1997a: 27) states that “EFL teachers working within a British English-oriented environment should continue to use RP (though not necessarily under that name) as their pronunciation model. But this model must be revised and updated from time to time”. At the same time, the teaching model should not be *too* progressive. It should reflect those innovations which have survived, but not the ephemeral trends (cf. Nolan 1999: 92, Gimson 1979). In this context, broadcast RP can be considered a suitable reference for a relatively conservative estimate of established RP usage.

7.8 Further research

The present work has advanced the case that RP has a place in variationist research on a par with other varieties, and that its phonological properties and social status can be explored within a descriptive sociolinguistic framework.

The analysis of broadcast RP has revealed a considerable amount of variability for most of the features investigated. A parallel investigation of conversational style would provide a valuable supplement to the present results, and might add support to the claim that usage levels for t-voicing, r-sandhi and yod coalescence would be greater in casual spontaneous speech.

In the current study I have concentrated on a homogeneous group of speakers in the same social setting. I have compared my findings with previous descriptions of RP and on

that basis suggested the direction of some ongoing changes in the accent. A sociolinguistic investigation of a wider range of speakers, with data from all age groups, would provide a stronger empirical basis for observations of change and might give a clearer picture of the development of RP.

With regard to gender, my study has shown that male speakers tend to have higher usage levels for variants that are phonologically motivated and enhance the ease of articulation. This observation should be an incentive to investigate whether this is a general tendency which applies to other varieties and social contexts.

Prevalent stereotypes about RP can contribute to obscuring the linguistic reality, and empirical data are necessary in order to keep the descriptions updated and in line with current usage. The present results, combined with Fabricius' (2000) data on t-glottalling, provide empirical facts on a wide range of RP features. There remain, however, several areas where RP exhibits variation, and where corpus-based research is needed in order to gain a full picture of the phonological status of the accent. Features such as GOOSE-fronting, monophthongisation of SQUARE and l-vocalisation, for example (cf. 4.2.2), have yet to be empirically investigated in RP.

It is also to be hoped that future research will continue to reconsider not only the phonetic ingredients of RP, but also the criteria for defining the accent.

Garrett et al. (2003: 62) argue that varieties should be viewed as ideological as well as linguistic entities. This is particularly true of RP, with its unique social history. While the accent is intimately bound up with issues of correctness and standardisation, there is no doubt that the social meaning of RP has undergone drastic changes in the dynamics of modern society. There are thus a great number of evaluative issues to explore in order to add empirical support to anecdotal evidence, and determine the current status of RP and how it relates to changing notions of standard and prestige. Large-scale attitudinal surveys among different social groups would therefore be a welcome addition to the linguistic descriptions. The present study has demonstrated that RP still has an important role in certain types of newscasting. A more general investigation of broadcast speech as such would yield valuable insight into accent use and the relative position of RP compared to other varieties.

An exploration of all the aspects outlined here would require ambitious variationist studies of a kind which is typically reserved for non-standard varieties. But, as Fabricius (2002b: 357) argues, "quantitative sociolinguistics can (and should) accommodate RP as one of the varieties of English which can be observed sociolinguistically". Such studies would entail certain theoretical problems related to the prescriptive status of RP (cf. 3.2-3-3), as well

as practical difficulties in identifying a relevant population (cf. 5.1). They would, nevertheless, provide valuable insights into the role of modern RP in the overall speech community.

References

- Abercrombie, D. 1992. RP Today: its position and prospects. In C. Blank (ed.), *Language and civilisation: a concerted profusion of essays and studies in honour of Otto Hietsch*. Frankfurt: Peter Lang. 6-10.
- Altendorf, U. 2003. *Estuary English: Levelling at the interface of RP and south-eastern British English*. Tübingen: Narr.
- Andersen, H. 1989. Understanding linguistic innovations. In L. E. Breivik and E. H. Jahr (eds.), *Language change: contributions to the study of its causes*. Berlin: Mouton de Gruyter. 5-28.
- Ascherson, N. 1994. Britain's crumbling ruling class is losing the accent of authority. *The Independent* 7 August 1994.
- Ball, M. 1984. The centring diphthongs in Southern English: a sound change in progress. *Journal of the International Phonetic Association* 14, 38-44.
- Bartsch, R. 1987. *Norms of language: theoretical and practical aspects*. London: Longman.
- Bauer, L. 1984. Linking /r/ in RP. Some facts. *Journal of the International Phonetic Association* 14, 74-79.
- Bauer, L. 1985. Tracing phonetic change in the received pronunciation of British English. *Journal of Phonetics* 13, 61-68.
- Bauer, L. 1994. *Watching English change*. London: Longman.
- Bauer, L., Dienhart, J., Hartvigson, H. and Jakobsen, L. K. 1980. *American English pronunciation*. Copenhagen: Gyldendal.
- Bayard, D. et al. 2002. Evaluating English accents worldwide.
<http://www.otago.ac.nz/anthropology/Linguistic/Accents.html>
- Bell, A. 1977. The language of radio news in Auckland: a sociolinguistic study of style, audience and subediting variation. PhD thesis, University of Auckland.
- Bell, A. 1982. Radio: the style of news language. *Journal of Communication* 32: 1, 150-164.
- Bell, A. 1983. Broadcast news as a language standard. *International Journal of the Sociology of Language* 40, 29-42.
- Bell, A. 1984. Language style as audience design. *Language in Society* 13: 2, 145-204.
- Bell, A. 1991. *The language of news media*. Oxford: Blackwell.
- Bell, A. 1997. Language style as audience design. In Coupland and Jaworski 1997, 240-250.

- Bex, T. and Watts, R. J. (eds.) 1999a. *Standard English: the widening debate*. London: Routledge.
- Bex, T. and Watts, R. J. 1999b. Introduction. In Bex and Watts 1999a, 1-9.
- Bourdieu, P. 1991. *Language and symbolic power*. Cambridge: Polity Press.
- Bourhis, R. and Giles, H. 1976. The language of co-operation in Wales: a field study. *Language Sciences* 42, 13-16.
- Briggs, A. 1985. *The BBC: the first fifty years*. Oxford: Oxford University Press.
- Broe, M. B. and Pierrehumbert, J. B. (eds.) 2000. *Acquisition and the lexicon. Papers in laboratory phonology V*. Cambridge: Cambridge University Press.
- Brown, A. 1988. Linking, intrusive and rhotic /r/ in pronunciation models. *Journal of the International Phonetic Association* 18, 144-151.
- Brown, G. 1990. *Listening to spoken English*. 2nd ed. London: Longman.
- Burchfield, R. 1981. *The spoken word. A BBC guide*. London: British Broadcasting Corporation.
- Butler, C. 1985. *Statistics in linguistics*. Oxford: Blackwell.
- Bybee, J. 2000. Lexicalization of sound change and alternating environments. In Broe and Pierrehumbert 2000, 250-268.
- Bybee, J. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
- Bybee, J. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14, 261-290.
- Cameron, D. 1990. Demythologising sociolinguistics: why language does not reflect society. In J. E. Joseph and T. J. Taylor (eds.), *Ideologies of language*. London: Routledge. 79-93.
- Chambers, J. K. 2002. *Sociolinguistic theory*. 2nd ed. Oxford: Blackwell.
- Chambers, J. K. and Trudgill, P. 1998. *Dialectology*. 2nd ed. Cambridge: Cambridge University Press.
- Chambers, J. K., Trudgill, P. and Schilling-Estes, N. (eds.) 2002. *The handbook of language variation and change*. Oxford: Blackwell.
- Cheshire, J. 2002. Sex and gender in variationist research. In Chambers et al. 2002, 423-443.
- Coates, J. 1998. Gossip revisited: language in all-female groups. In J. Cheshire and P. Trudgill (eds.), *The sociolinguistics reader. Vol. 2: Gender and discourse*. London: Arnold. 127-152. (Originally published in J. Coates and D. Cameron (eds.), *Women in their speech communities*. Longman 1989.)
- Coates, J. 2004. *Women, men and language*. 3rd ed. Harlow: Pearson Education.

- Cogle, P. 1993. *Do you speak Estuary? The new Standard English*. London: Bloomsbury.
- Coupland, N. 1980. Style-shifting in a Cardiff work-setting. *Language in Society* 9, 1-12.
- Coupland, N. 1990. 'Standard Welsh English': a variable semiotic. In N. Coupland (ed.), *English in Wales: diversity, conflict and change*. Clevedon: Multilingual Matters. 232-257.
- Coupland, N. and Jaworski, A. (eds.) 1997. *Sociolinguistics. A reader and coursebook*. Basingstoke: Palgrave.
- Crowley, T. 2003. *Standard English and the politics of language*. 2nd ed. Basingstoke: Palgrave Macmillan.
- Cruttenden, A. 1994. *Gimson's Pronunciation of English*. 5th ed. London: Edward Arnold.
- Cruttenden, A. 2001. *Gimson's Pronunciation of English*. 6th ed. London: Arnold.
- Crystal, D. 1984. *Who cares about English usage?* Harmondsworth: Penguin.
- Crystal, D. 1992. The changing English language – fiction and fact. In M. Pütz (ed.), *Thirty years of linguistic evolution. Studies in honour of René Dirven on occasion of his 60th birthday*. Philadelphia: Benjamins. 119-130.
- Crystal, D. 1995. *The Cambridge encyclopaedia of the English language*. Cambridge: Cambridge University Press.
- Crystal, D. 2003. *A dictionary of linguistics and phonetics*. 5th ed. London: Blackwell
- Culf, A. 1994. BBC radio seeks to reach parts cut-glass accents cannot reach. *The Guardian* 27 January 1994.
- Darley, F. L. and Spriestersbach, D. C. 1978. *Diagnostic methods in speech pathology*. 2nd ed. New York: Harper & Row.
- Davenport, M. and Hannahs, S. J. 2005. *Introducing phonetics and phonology*. 2nd ed. London: Hodder Arnold.
- Deumert, A. and Vandenbussche, W. (eds.) 2003a. *Germanic standardizations: past to present*. Amsterdam: John Benjamins.
- Deumert, A. and Vandenbussche, W. 2003b. Standard languages: taxonomies and histories. In Deumert and Vandenbussche 2003a, 1-14.
- Dobson, E. J. 1957. *English pronunciation 1500-1700. Volume II: phonology*. Oxford: Clarendon.
- Docherty, G. and Foulkes, P. 1999. Derby and Newcastle: instrumental phonetics and variationist studies. In Foulkes and Docherty 1999a, 47-71.
- Dorian, N. 1993. Internally and externally motivated change in language contact settings: doubts about dichotomy. In Jones 1993, 131-155.

- Dressler, W. and Wodak, R. 1982. Sociophonological methods in the study of sociolinguistic variation in Viennese German. *Language in Society* 11: 3, 339-370.
- Eckert, P. 1997a The whole woman: sex and gender differences in variation. In Coupland and Jaworski 1997, 212-231.
- Eckert, P. 1997b. Age as a sociolinguistic variable. In F. Coulmas (ed.), *The handbook of sociolinguistics*. Oxford: Blackwell. 151-167.
- Eckert, P. 2000. *Linguistic variation as social practice*. Oxford: Blackwell.
- Ellis, A. J. 1869-89. *On Early English Pronunciation*. London: Asher & Co.
- Elyan, O., Smith, P., Giles, H. and Bourhis, R. 1978. RP-accented female speech: the voice of perceived androgyny. In P. Trudgill (ed.), *Sociolinguistic patterns in British English*. London: Arnold. 122-131.
- Fabricius, A. 2000. T-glottalling between stigma and prestige: a sociolinguistic study of modern RP. Unpublished PhD thesis, Copenhagen Business School.
- Fabricius, A. 2002a. Ongoing change in modern RP. Evidence for the disappearing stigma of t-glottalling. *English World-Wide* 23: 1, 115-136.
- Fabricius, A. 2002b. RP as sociolinguistic object. *Nordic Journal of English Studies* 1: 2, 355-372.
- Fosler-Lussier, E. and Morgan, N. 1999. Effects of speaking rate and word frequency on pronunciations in conversational speech. *Speech Communication* 29, 137-158.
- Foster, B. 1968. *The changing English language*. London: Macmillan.
- Foulkes, P. 1997. English [r]-sandhi – a sociolinguistic perspective. *Histoire, Epistémologie, Langage* 19: 1, 73-96.
- Foulkes, P. and Docherty, G. (eds.) 1999a. *Urban voices. Accent studies in the British Isles*. London: Arnold.
- Foulkes, P. and Docherty, G. 1999b. Urban voices – overview. In Foulkes and Docherty 1999a, 1-24.
- Garrett, P., Coupland, N. and Williams, A. 2003. *Investigating language attitudes: social meanings of dialect, ethnicity and performance*. Cardiff: University of Wales Press.
- Giegerich, H. J. 1992. *English phonology: an introduction*. Cambridge: Cambridge University Press.
- Giles, H., Coupland, N., Henwood, K., Harriman, J. and Coupland, J. 1990. The social meaning of RP: an intergenerational perspective. In Ramsaran 1990a, 191-211.

- Giles, H. and Powesland, P. 1997. Accommodation theory. In Coupland and Jaworski 1997, 232-39. (Reprinted from H. Giles and P. Powesland (1975), *Speech style and social evaluation*, 154-170. London: Academic Press.)
- Gimson, A.C. 1962. *An introduction to the pronunciation of English*. London: Edward Arnold.
- Gimson, A.C. 1964. Phonetic change and the RP vowel system. In D. Abercrombie et al. (eds.), *In honour of Daniel Jones: papers contributed on the occasion of his eightieth birthday, 12 September 1961*. London: Longman. 131-136.
- Gimson, A.C. 1970. British English pronunciation: standards and evolution. *Praxis des Neusprachlichen Unterrichts* 17, 17-20.
- Gimson, A.C. 1977. Daniel Jones and standards of English pronunciation. *English Studies* 58, 151-158.
- Gimson, A.C. 1979. English RP: Ancient or modern? *Praxis des Neusprachlichen Unterrichts* 2, 149-56.
- Gimson, A.C. 1980. *An introduction to the pronunciation of English*. 3rd ed. London: Edward Arnold.
- Gimson, A.C. 1981. The twentyman lecture, 1981: the pronunciation of English: its intelligibility and acceptability in the world. *Modern Languages* 62:2, 61-68.
- Gimson, A.C. 1984. The RP accent. In Trudgill 1984, 45-54.
- Greenberg, J. H. 1979. Rethinking linguistics diachronically. *Language* 55, 275-290.
- Greenberg, S. and Fosler-Lussier, E. 2000. The uninvited guest: information's role in guiding the production of spontaneous speech. Proceedings of the CREST workshop on Models of Speech Production: Motor Planning and Articulatory Modelling, Kloster Seeon, Germany, May 1-4.
- Guy, G. 1980. Variation in the group and the individual: the case of final stop deletion. In W. Labov (ed.), *Locating language in time and space*. New York: Academic Press. 1-36.
- Harris, J. 1994. *English sound structure*. Oxford: Blackwell.
- Harris, J. and Kaye, J. 1990. A tale of two cities: London glottalling and New York City tapping. *The Linguistic Review* 7, 251-274.
- Haugen, E. 1997. Language standardization. In Coupland and Jaworski 1997, 341-352. (Reprinted from E. Haugen (1966), Dialect, language, nation. *American Anthropologist*, 68: 6, 922-935.)
- Hickey, R. 1999. Dublin English: current changes and their motivation. In Foulkes and Docherty 1999a, 265-281.

- Holmes, J. 1986. Functions of *you know* in women's and men's speech. *Language in Society* 15: 1, 1-22.
- Holmes, J. 1992. *An introduction to sociolinguistics*. London: Longman.
- Holmes, J. 1994. New Zealand flappers: an analysis of T Voicing in New Zealand English. *English World-Wide* 15: 2, 195-224.
- Holmes, J. 1995. Glottal stops in New Zealand English: an analysis of variants of word-final /t/. *Linguistics* 33: 3, 433-463.
- Honey, J. 1985. Acrolect and hyperlect: the redefinition of English RP. *English Studies* 66: 3, 241-257.
- Honey, J. 1988. "Talking proper": schooling and the establishment of English "Received Pronunciation". In G. Nixon and J. Honey (eds.), *An historic tongue: studies in English linguistics in memory of Barbara Strang*. London: Routledge. 209-227.
- Honey, J. 1991. *Does accent matter? The Pygmalion factor*. London: Faber and Faber.
- Honey, J. 2000. The establishment of the English RP accent: a flawed interpretation? *Bulletin of the International Association of University Professors of English*, issue for autumn 2000. <http://www.phon.ucl.ac.uk/home/estuary/honey-muggles.htm>
- Hooper, J. B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In W. M. Christie (ed.), *Current progress in historical linguistics*. Amsterdam: North Holland. 95-105.
- Hudson, R. A. 1996. *Sociolinguistics*. 2nd ed. Cambridge: Cambridge University Press.
- Hughes, A. and Trudgill, P. 1996. *English accents and dialects*. 3rd ed. London: Arnold.
- Johansson, S. 1973. Linking and intrusive /r/ in English: a case for a more concrete phonology. *Studia Linguistica* 27, 53-68.
- Jones, C. (ed.) 1993. *Historical linguistics: problems and perspectives*. London: Longman.
- Jones, D. 1917. *An English pronouncing dictionary*. London: J. M. Dent & Sons.
- Jones, D. 1954. Falling and rising diphthongs in Southern English. *Miscellanea Phonetica* II, 1-12.
- Jones, D. 1960. *An outline of English phonetics*. 9th ed. Cambridge: W. Heffer & Sons.
- Jones, D. 1967. *English pronouncing dictionary*. 13th ed. London: J. M. Dent & Sons.
- Jones, D. 1997. *English pronouncing dictionary*. 15th ed. Edited by P. Roach and J. Hartman. Cambridge: Cambridge University press.
- Jones, D. 2003. *English pronouncing dictionary*. 16th ed. Edited by P. Roach, J. Hartman and J. Setter. Cambridge: Cambridge University Press.

- Jurafsky, D., Bell, A., Gregory, M. and Raymond W. D. 2001. Probabilistic relations between words: evidence from reduction in lexical production. In J. Bybee and P. Hopper (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins. 229-254.
- Kerswill, P. 1985. A sociophonetic study of connected speech processes in Cambridge English: an outline and some results. *Cambridge Papers in Phonetics and Experimental Linguistics* 4.
- Kerswill, P. 1987. Levels of linguistic variation in Durham. *Journal of Linguistics* 23: 1, 25-49.
- Kerswill, P. 1994. Estuary English. <http://www.phon.ucl.ac.uk/home/estuary/kerswill.htm>
- Kerswill, P. 1996. Children, adolescents and language change. *Language Variation and Change* 8: 2, 177-202.
- Kerswill, P. 2001. Mobility, meritocracy and dialect levelling: the fading (and phasing) out of Received Pronunciation. In P. Rajamäe and K. Vogelberg (eds.), *British studies in the new millennium: challenge of the grassroots*. Tartu: University of Tartu. 45-58.
- Kerswill, P. 2003. Dialect levelling and geographical diffusion in British English. In D. Britain and J. Cheshire (eds.), *Social dialectology. In honour of Peter Trudgill*. Amsterdam: John Benjamins. 223-243.
- Kerswill, P. and Wright, S. 1990. The validity of phonetic transcription: limitations of a sociolinguistic research tool. *Language Variation and Change* 2, 255-275.
- Killen, M. 2001. When posh came to shove. *The Guardian* 6 November 2001.
- Knight, I. 2001. Speak proper? Not likely. *Sunday Times* 11 November 2001.
- Kreidler, C. W. 2004. *The Pronunciation of English: a course book*. 2nd ed. Oxford: Blackwell.
- Kristiansen, T. 2001. Two standards: one for the media and one for the school. *Language Awareness* 10: 1, 9-24.
- Kroch, A. S. 1978. Toward a theory of social dialect variation. *Language in Society* 7, 17-36.
- Labov, W. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, W. 1981. Resolving the Neogrammarian controversy. *Language* 57: 2, 267-308.
- Labov, W. 1984. Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (eds.), *Language in use: readings in sociolinguistics*. Englewood Cliffs: Prentice-Hall. 28-53.
- Labov, W. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2: 2, 205-254.

- Labov, W. 1994. *Principles of linguistic change. Vol. 1: internal factors*. Oxford: Blackwell.
- Labov, W. 2001. *Principles of linguistic change. Vol. 2: social factors*. Oxford: Blackwell.
- Lass, R. 1980. *On explaining language change*. Cambridge: Cambridge University Press.
- Leech, G., Rayson, P. and Wilson, A. 2001. *Word frequencies in written and spoken English: based on the British National Corpus*. London: Longman.
- Leitner, G. 1980. BBC English and Deutsche Rundfunksprache: a comparative and historical analysis of the language on the radio. *International Journal of the Sociology of Language* 26, 75-100.
- Leitner, G. 1983. Introduction. *International Journal of the Sociology of Language* 40, 5-7.
- Lewis, J. W. 1972. *A concise pronouncing dictionary of British and American English*. London: Oxford University Press.
- Lewis, J. W. 1975. Linking /r/ in the General British pronunciation of English. *Journal of the International Phonetic Association* 5, 37-42.
- Lewis, J. W. 1985. British non-dialect accents. *Zeitschrift für Anglistik und Amerikanistik* 33:3, 244-257.
- Lewis, J. W. 1990. *HappY* land reconnoitred: the unstressed word-final -y vowel in General British pronunciation. In Ramsaran 1990a, 159-167.
- Lindblom, B. 1980. The goal of phonetics, its unification and application. *Phonetica* 37, 7-26.
- Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (eds.), *Speech production and speech modelling*. Dordrecht: Kluwer. 403-439.
- Lloyd James, A. 1931. *Broadcast English I*. 2nd ed. London: British Broadcasting Corporation.
- Lutz, A. 1991. *Phonotaktisch gesteuerte Konsonantenveränderungen in der Geschichte des Englischen*. Tübingen: Niemeyer.
- Macaulay, R. 1988. RP R.I.P. *Applied Linguistics* 9, 115-124.
- MacMahon, M. K. C. 1998. Phonology. In S. Romaine (ed.), *The Cambridge History of the English Language. Vol. 4: 1776-1997*. Cambridge: Cambridge University Press. 373-535.
- Maidment, J. A. 1994. Estuary English: hybrid or hype? Paper presented at the 4th New Zealand Conference on Language & Society, Lincoln University, Christchurch, New Zealand, August 1994. <http://www.phon.ucl.ac.uk/home/estuary/maidment-sil.htm>
- Mathisen, A. G. 1999. Sandwell, West Midlands: ambiguous perspectives on gender patterns and models of change. In Foulkes and Docherty 1999a, 107-123.

- McCafferty, K. 1999. (London)Derry: between Ulster and local speech – class, ethnicity and language change. In Foulkes and Docherty 1999a, 246-264.
- McCrum, R., Cran, W. and Mac Neil, R. 1986. *The story of English*. New York: Elisabeth Sifton Books/Viking.
- McKenzie, M. 2004. An investigation into the accents used in UK television news programmes. Unpublished MA thesis, University of Glasgow.
- McMahon, A. 1994. *Understanding language change*. Cambridge: Cambridge University Press.
- McMahon, A., Foulkes, P. and Tollfree, L. 1994. Gestural representation and Lexical Phonology. *Phonology* 11, 177-316.
- Mees, I. M. and Collins, B. 1999. Cardiff: a real-time study of glottalisation. In Foulkes and Docherty 1999a, 185-202.
- Miller, G. M. 1971. *BBC pronouncing dictionary of British names*. London: Oxford University Press.
- Milroy, J. 1992. *Linguistic variation and change*. Oxford: Blackwell.
- Milroy, J. 2001a. Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*. 5: 4, 530-555.
- Milroy, J. 2001b. Received Pronunciation: who “receives” it and how long will it be “received”? *Studia Anglica Posnaniensia* 36, 15-33.
- Milroy, J., Milroy, L. and Hartley, S. 1994. Local and supra-local change in British English: the case of glottalisation. *English World-Wide* 15: 1, 1-33.
- Milroy, J. and Milroy, L. 1997. Network structure and linguistic change. In Coupland and Jaworski 1997, 199-211.
- Milroy, J. and Milroy, L. 1999. *Authority in language: investigating standard English*. 3rd ed. London: Routledge.
- Milroy, L. 1987. *Observing and analysing natural language*. Oxford: Blackwell.
- Milroy, L. 1999. Standard English and language ideology in Britain and the United States. In Bex and Watts 1999a, 173-205.
- Milroy, L. and Gordon, M. 2003. *Sociolinguistics: method and interpretation*. Oxford: Blackwell.
- Modiano, M. 1999. Standard English(es) and educational practices for the world’s lingua franca. *English Today* 15: 4, 3-13.
- Morrish, J. 1999. The accent that dare not speak its name. *The Independent on Sunday* 21 March 1999.

- Mugglestone, L. 1995. *'Talking proper': the rise of accent as a social symbol*. Oxford: Clarendon Press.
- Mæhlum, B. 2003. Normer. In Mæhlum et al. 2003, 86-102.
- Mæhlum, B., Akselberg, G., Røyneland, U. and Sandøy, H. 2003. *Språkmøte: innføring i sosiolingvistikk*. Oslo: Cappelen.
- Nevalainen, T. 2003. English. In Deumert and Vandebussche 2003a, 127-156.
- Newbrook, M. 1999. West Wirral: norms, self reports and usage. In Foulkes and Docherty 1999a, 90-106.
- Nolan, F. 1999. The shifting sands of English pronunciation. In P. Rajamäe (ed.), *New Britain: The heritage of the past and the challenge of the future*. Proceedings of the 2nd International Tartu Conference on British Studies, University of Tartu, August 24-25 1998. 76-93.
- Nolan, F. and Kerswill, P. 1990. The description of connected speech processes. In Ramsaran 1990a, 295-316.
- O'Donnell, W. R. And Todd, L. 1991. *Variety in contemporary English*. London: Routledge.
- Ohala, J. 1992. The segment: primitive or derived? In G. Docherty and D. R. Ladd (eds.), *Papers in laboratory phonology II: gesture, segment, prosody*. Cambridge: Cambridge University Press. 166-189.
- Ohala, J. 1993. The phonetics of sound change. In Jones 1993, 237-278.
- Patrick, P. L. 2002. The speech community. In Chambers et al. 2002, 573-597.
- Patterson, D. and Connine, C. 2001. A corpus analysis of variant frequency in American English flap production. *Phonetica* 58: 4, 254-275.
- Pointon, G. 1988. The BBC and English pronunciation. *English Today* 4, 8-12.
- Przedlacka, J. 2002. *Estuary English? A sociophonetic study of teenage speech in the Home Counties*. Frankfurt am Main: Peter Lang.
- Quirk, R. 1982. *Style and communication in the English language*. London: Edward Arnold.
- Ramsaran, S. 1978. Phonetic and phonological correlates of style in English: a preliminary investigation. PhD thesis, University of London.
- Ramsaran, S. (ed.) 1990a. *Studies in the pronunciation of English. A commemorative volume in honour of A.C. Gimson*. London: Routledge.
- Ramsaran, S. 1990b. RP: fact and fiction. In Ramsaran 1990a, 178-190.
- Rastall, P. 1997. How d'you say 'consume'? *English Today* 13: 4, 19-21.
- Reith, J. C. W. 1924. *Broadcast over Britain*. London: Hodder and Stoughton.

- Roach, P. 1983. *English phonetics and phonology: a practical course*. Cambridge: Cambridge University Press.
- Roach, P. 1992. *Introducing phonetics*. London: Penguin English.
- Romaine, S. 1980. A critical overview of the methodology of urban British sociolinguistics. *English World-Wide* 1: 2, 163-198.
- Romaine, S. and Reid, E. 1976. Glottal sloppiness? A sociolinguistic view of urban speech in Scotland. *Teaching English*, 9: 3. Edinburgh, C.I.T.E.
- Rosewarne, D. 1984. Estuary English. *Times educational Supplement*, 19 October 1984.
- Rosewarne, D. 1994. Estuary English: tomorrow's RP? *English Today* 37, 3-8.
- Rosewarne, D. 1998. Tony Blair and William Hague: two northerners heading south. *English Today* 56: 14: 4, 53-54.
- Sandøy, H. 2003. Skriftvariasjon. In Mæhlum et al. 2003, 166-195.
- Sankoff, G. 1974. A quantitative paradigm for the study of communicative competence. In R. Bauman and J. Sherzer (eds.), *Explorations in the ethnography of speaking*. Cambridge: Cambridge University Press. 18-49.
- Schilling-Estes, N. 2002. Investigating stylistic variation. In Chambers et al. 2002, 375-401.
- Shockey, L. 2003. *Sound patterns of spoken English*. Oxford: Blackwell.
- Sinclair, J. (ed.) 1992. *BBC English Dictionary*. London : BBC/HarperCollins.
- Sivertsen E. 1960. *Cockney phonology*. Oslo: Oslo University Press.
- Smith, J. 1996. *An historical study of English: function, form and change*. London: Routledge.
- Spencer, J. 1957. Received Pronunciation: some problems of interpretation. *Lingua* 7, 7-29.
- Steriade, D. 2000. Paradigm uniformity and the phonetics-phonology boundary. In Broe and Pierrehumbert 2000, 313-334.
- Stoddart, J., Upton, C. and Widdowson, J. D. A. 1999. Sheffield dialect in the 1990s: revisiting the concept of NORMs. In Foulkes and Docherty 1999a, 72-89.
- Stuart-Smith, J. 1999. Glasgow: accent and voice quality. In Foulkes and Docherty 1999a, 203-222.
- Sundby, N. K. 1974. *Om normer*. Oslo: Universitetsforlaget.
- Sweet, H. 1908. *The sounds of English: an introduction to phonetics*. Oxford: Clarendon.
- Taylor, J. 1998. What, pray, is happening to dear old RP? *La Linguistique* 34: 2, 141-151.
- Timmins, C. and Stuart-Smith, J. 2005. 'A question of life': investigating the effects of television on accent variation in adolescents. Paper presented at the 5th UK Language Variation and Change Conference, Aberdeen 12-14 September 2005.

- Tollfree, L. 1999. South East London English: discrete *versus* continuous modelling of consonantal reduction. In Foulkes and Docherty 1999a, 163-184.
- Trask, R. L. *A dictionary of phonetics and phonology*. London: Routledge.
- Trim, J. L. M. 1961. English standard pronunciation. *English Language Teaching* 16: 1, 28-37.
- Trudgill, P. 1974. *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Trudgill, P. (ed.) 1984. *Language in the British Isles*. Cambridge: Cambridge University Press.
- Trudgill, P. 1986. *Dialects in contact*. Oxford: Basil Blackwell.
- Trudgill, P. 1999. Standard English: what it isn't. In Bex and Watts 1999a, 117-128.
- Trudgill, P. 2000. *Sociolinguistics: an introduction to language and society*. London: Penguin.
- Trudgill, P. 2002. *Sociolinguistic variation and change*. Edinburgh: Edinburgh University Press.
- Trudgill, P. and Hannah, J. 2002. *International English*. 4th ed. London: Arnold.
- Upton, C. 2004. Received Pronunciation. In E. Schneider et al. (eds.), *A handbook of varieties of English. Volume 1: Phonology*. Berlin: Mouton de Gruyter. 217-230.
- Upton, C., Kretzschmar, W. A. and Konopka, R. 2001. *Oxford dictionary of pronunciation for current English*. Oxford: Oxford University Press.
- Vannebo, K. I. 1980. Om språkvetenskapens norm-begrep. *Tijdschrift voor Skandinavistiek*, 1: 1, 3-23.
- Vennemann, T. 1988. *Preference laws for syllable structure and the explanation of sound change*. Berlin: Mouton de Gruyter.
- Ward, I. 1939. *The phonetics of English*. 3rd ed. Cambridge: Heffer.
- Watt, D. 2002. 'I don't speak with a Geordie accent, I speak, like, the Northern accent': Contact-induced levelling in the Tyneside vowel system. *Journal of Sociolinguistics* 6: 1, 44-63.
- Watt, D. and Milroy, L. 1999. Patterns of variation and change in three Newcastle vowels: is this dialect levelling? In Foulkes and Docherty 1999a, 25-46.
- Weinreich, U., Labov, W. and Herzog, M. 1968. Empirical foundations for a theory of language change. In W. Lehmann and Y. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press. 95-188.
- Wells, J.C. 1982. *Accents of English* (3 vols.). Cambridge: Cambridge University Press.

- Wells, J.C. 1984. English accents in England. In Trudgill 1984, 55-69.
- Wells, J.C. 1985. English pronunciation and its dictionary representation. In R. Ilson (ed.), *Dictionaries, lexicography and language learning*. Oxford: Pergamon Press. 45-51.
- Wells, J.C. 1990a. *Longman Pronunciation Dictionary*. London: Longman.
- Wells, J.C. 1990b. A phonetic update on RP. *Moderna Språk* 82: 1, 3-9.
- Wells, J.C. 1990c. Syllabification and allophony. In Ramsaran 1990a, 76-86.
- Wells, J.C. 1994a. The Cockneyfication of R.P.? In G. Melchers and N.-L. Johannesson (eds.), *Non-standard varieties of language*. Stockholm Studies in English 84. Stockholm: Almqvist & Wiksell. 198-205.
- Wells, J.C. 1994b. Transcribing Estuary English: a discussion document. *Speech Hearing and Language: UCL Work in Progress* 8, 259-267.
- Wells, J.C. 1997a. Whatever happened to Received Pronunciation? In Medina and Soto (eds.), *Il Jornadas de Estudios Ingleses*. Universidad de Jaén, Spain. 19-28.
- Wells, J.C. 1997b. What is Estuary English? *English Teaching Professional* 3, 46-47.
- Wells, J.C. 1998a. Our changing pronunciation. Summary of a talk given at the British Association annual Festival of Science, Cardiff, 9th September 1998.
<http://www.phon.ucl.ac.uk/home/wells/cardiff.htm>
- Wells, J.C. 1998b. Pinning down Estuary English. Lecture given in Lund, Sweden, April 1998. <http://www.phon.ucl.ac.uk/home/estuary/est-lund.htm>
- Wells, J.C. 1999a. British English pronunciation preferences: a changing scene. *Journal of the International Phonetic Association* 29: 1, 33-50.
- Wells, J.C. 1999b. Questions and answers about Estuary English.
<http://www.phon.ucl.ac.uk/home/estuary/ee-faqs-jcw.htm>
- Wells, J.C. 2000. *Longman Pronunciation Dictionary*. New edition. London: Longman.
- Wells, J.C. and Colson, G. 1971. *Practical phonetics*. London: Pitman.
- White, M. 1999. Goodwill all round at the British-Irish summit. *The Guardian*, 18 December, 1999.
- Williams, A. and Kerswill, P. 1999. Dialect levelling: change and continuity in Milton Keynes, Reading and Hull. In Foulkes and Docherty 1999a, 141-162.
- Wold, S. H. 2003. Aspects of phonological variation in broadcast Scottish English. Unpublished MA thesis, University of Bergen.
- Wolfson, N. 1997. Speech events and natural speech. In Coupland and Jaworski 1997, 116-125.

- Woods, A., Fletcher, P. and Hughes, A. 1986. *Statistics in language studies*. Cambridge: Cambridge University Press.
- Woods, H. B. 1991. Social differentiation in Ottawa English. In J. Cheshire (ed.), *English around the world: sociolinguistic perspectives*. Cambridge: Cambridge University Press. 134-149.
- Woods, N. 1997. The formation and development of New Zealand English: interaction of gender-related variation and linguistic change. *Journal of Sociolinguistics* 1: 1, 95-125.
- Woods, N. 2001. Internal and external dimensions of language change: the great divide? Evidence from New Zealand English. *Linguistics* 39: 5, 973-1007.
- Wotschke, I. 1996. Socio-regional speech versus “Oxford accent” – trends and fashions in the pronunciation of English English. *Hermes, Journal of Linguistics* 17: 213-235.
- Wyld, H. C. 1920. *A history of modern colloquial English*. London: T. F. Unwin.
- Zue, V. W. and Laferriere, M. 1979. Acoustic study of medial /t, d/ in American English. *Journal of the Acoustical Society of America* 66: 4, 1039-1050.

Appendix: Raw data, all speakers

Linking (r): total scores

Speaker	/r/		Ø		Total	
	N	%	N	%	N	%
1	91	53.2	80	46.8	171	100
2	125	61.6	78	38.4	203	100
3	87	50.0	87	50.0	174	100
4	83	42.1	114	57.9	197	100
5	143	68.7	65	31.3	208	100
6	107	59.1	74	40.9	181	100
7	156	70.3	66	29.7	222	100
8	140	71.4	56	28.6	196	100
9	94	53.1	83	46.9	177	100
10	122	67.4	59	32.6	181	100
11	143	62.2	87	37.8	230	100
12	175	79.5	45	20.5	220	100
13	91	48.4	97	51.6	188	100
14	193	77.2	57	22.8	250	100
15	158	75.2	52	24.8	210	100
16	104	55.3	84	44.7	188	100
17	163	59.3	112	40.7	275	100
18	144	57.6	106	42.4	250	100
19	110	48.7	116	51.3	226	100
20	117	55.5	94	44.5	211	100
21	92	52.9	82	47.1	174	100
22	111	56.9	84	43.1	195	100
23	142	66.7	71	33.3	213	100
24	121	69.9	52	30.1	173	100
25	110	67.9	52	32.1	162	100
26	134	69.1	60	30.9	194	100
27	141	67.1	69	32.9	210	100
28	99	52.1	91	47.9	190	100
29	71	35.9	127	64.1	198	100
30	45	25.3	133	74.4	178	100

Intrusive (r): total scores

Speaker	/r/		Ø		Total	
	N	%	N	%	N	%
1	3	16.7	15	83.3	18	100
2	13	48.1	14	51.9	27	100
3	3	13.0	20	87.0	23	100
4	3	11.5	23	88.5	26	100
5	12	42.9	16	57.1	28	100
6	11	42.3	15	57.7	26	100
7	30	69.8	13	30.2	43	100
8	9	50.0	9	50.0	18	100
9	2	16.7	10	83.3	12	100
10	-	0	23	100	23	100
11	5	50.0	5	50.0	10	100
12	11	73.3	4	26.7	15	100
13	1	6.7	14	93.3	15	100
14	9	60.0	6	40.0	15	100
15	5	27.8	13	72.2	18	100
16	7	31.8	15	68.2	22	100
17	6	27.3	16	72.7	22	100
18	2	25.0	6	75.0	8	100
19	8	44.4	10	55.6	18	100
20	2	28.6	5	71.4	7	100
21	2	6.9	27	93.1	29	100
22	2	12.5	14	87.5	16	100
23	8	50.0	8	50.0	16	100
24	3	23.1	10	76.9	13	100
25	10	83.3	2	16.7	12	100
26	-	0	15	100	15	100
27	8	57.1	6	42.9	14	100
28	6	35.3	11	64.7	17	100
29	1	6.3	15	93.7	16	100
30	-	0	16	100	16	100

Linking (r) before proper nouns

Speaker	Before proper nouns		Other environments	
	/r/	∅	/r/	∅
1	4	17	87	63
2	1	8	124	70
3	-	17	87	70
4	4	10	79	104
5	1	7	142	58
6	2	19	105	55
7	4	23	152	43
8	2	8	138	48
9	-	13	94	70
10	-	4	122	55
11	1	5	142	82
12	1	10	174	35
13	-	11	91	86
14	3	2	190	55
15	2	8	156	44
16	-	9	104	75
17	1	25	162	87
18	-	7	144	99
19	-	6	110	110
20	2	27	115	67
21	-	2	92	80
22	-	10	111	74
23	1	2	141	69
24	7	4	114	48
25	1	3	109	49
26	4	7	130	53
27	-	9	141	60
28	-	9	99	82
29	-	12	71	115
30	-	9	45	124
Total	41	303	3571	2130

Linking (r) near another /r/

Speaker	Near /r/		Other environments	
	/r/	∅	/r/	∅
1	1	10	90	70
2	2	3	123	75
3	-	13	87	74
4	3	14	80	100
5	4	10	139	55
6	1	14	106	60
7	3	14	153	52
8	2	4	138	52
9	-	15	94	68
10	1	5	121	54
11	1	4	142	83
12	2	2	173	43
13	2	8	89	89
14	3	14	190	43
15	-	9	158	43
16	2	7	102	77
17	1	10	162	102
18	1	6	143	100
19	1	13	109	103
20	-	12	117	82
21	1	11	91	71
22	-	7	111	77
23	2	4	140	67
24	1	4	120	48
25	2	8	108	44
26	-	6	134	54
27	1	4	140	65
28	-	2	99	89
29	-	3	71	124
30	-	7	45	126
Total	37	243	3575	2190

Linking (r) before clause boundary

Speaker	+ boundary		÷ boundary	
	/r/	∅	/r/	∅
1	2	2	89	78
2	1	7	124	71
3	-	1	87	86
4	3	1	80	113
5	6	1	137	64
6	-	1	107	73
7	7	5	149	61
8	2	4	138	52
9	1	7	93	76
10	3	3	119	56
11	-	5	143	82
12	11	3	164	42
13	1	2	90	95
14	5	-	188	57
15	12	6	146	46
16	2	5	102	79
17	6	3	157	109
18	-	2	144	104
19	1	2	109	114
20	6	4	111	90
21	-	3	92	79
22	7	5	104	79
23	7	5	135	66
24	1	4	120	48
25	4	4	106	48
26	9	6	125	54
27	4	3	137	66
28	-	5	99	86
29	-	2	71	125
30	-	1	45	132
Total	101	102	3511	2331

Linking (r) according to the grammatical category of the surrounding words

Speaker	Between lexical words		Between function words	
	/r/	∅	/r/	∅
1	11	29	80	51
2	13	17	112	61
3	9	20	78	67
4	5	35	78	79
5	17	17	126	48
6	7	26	100	48
7	16	20	140	46
8	9	10	131	46
9	10	20	84	63
10	17	15	105	44
11	15	9	128	78
12	29	13	146	32
13	9	27	82	70
14	21	14	172	43
15	15	20	143	32
16	8	16	96	68
17	22	35	141	77
18	20	24	124	82
19	14	30	96	86
20	13	19	104	75
21	9	16	83	66
22	12	17	99	67
23	13	13	129	58
24	20	9	101	43
25	15	14	95	38
26	21	12	113	48
27	15	18	126	51
28	18	13	81	78
29	12	28	59	99
30	11	23	34	110
Total	426	579	3186	1854

Linking (r) according to the weight of the following syllable

Speaker	+ stressed V		+ unstressed V	
	/r/	Ø	/r/	Ø
1	5	22	86	58
2	24	20	101	58
3	7	34	80	53
4	7	32	76	82
5	13	25	130	40
6	12	44	95	30
7	14	35	142	31
8	11	23	129	33
9	11	26	83	57
10	13	18	109	41
11	9	31	134	56
12	29	14	146	31
13	22	34	69	63
14	24	14	169	43
15	26	18	132	34
16	17	23	87	61
17	19	38	144	74
18	32	44	112	62
19	18	40	92	76
20	9	34	108	60
21	12	28	80	54
22	13	28	98	56
23	15	17	127	54
24	27	15	94	37
25	14	19	96	33
26	26	24	108	36
27	16	41	125	28
28	2	28	97	63
29	1	28	70	99
30	4	45	41	88
Total	452	842	3160	1591

Intrusive (r) before proper nouns

Speaker	Before proper nouns		Other environments	
	/r/	Ø	/r/	Ø
1	-	4	3	11
2	-	2	13	12
3	-	1	3	19
4	-	2	3	21
5	-	-	12	16
6	-	-	11	15
7	-	5	30	8
8	-	2	9	7
9	-	2	2	8
10	-	-	-	23
11	-	-	5	5
12	-	-	11	4
13	-	1	1	13
14	1	-	8	6
15	1	3	4	10
16	-	-	7	15
17	-	-	6	16
18	-	-	2	6
19	-	-	8	10
20	-	-	2	5
21	-	4	2	23
22	-	-	2	14
23	-	1	8	7
24	1	2	2	8
25	1	-	9	2
26	-	-	-	15
27	-	1	8	5
28	-	2	6	9
29	-	2	1	13
30	-	-	-	16
Total	4	34	178	342

Intrusive (r) near another /r/

Speaker	Near /r/		Other environments	
	/r/	∅	/r/	∅
1	-	2	3	13
2	1	5	12	9
3	-	6	3	14
4	1	5	2	18
5	-	9	12	7
6	5	5	6	10
7	2	5	28	8
8	1	3	8	6
9	1	2	1	8
10	-	-	-	23
11	-	-	5	5
12	2	1	9	3
13	-	3	1	11
14	1	-	8	6
15	1	1	4	12
16	-	5	7	10
17	3	12	3	4
18	1	1	1	5
19	1	3	7	7
20	-	1	2	4
21	-	7	2	20
22	-	1	2	13
23	1	3	7	5
24	2	2	1	8
25	1	-	9	2
26	-	-	-	15
27	-	-	8	6
28	1	-	5	11
29	-	6	1	9
30	-	-	-	16
Total	25	88	157	288

Intrusive (r) according to the weight of the following syllable

Speaker	+ stressed V		+ unstressed V	
	/r/	∅	/r/	∅
1	-	4	3	11
2	-	6	13	8
3	-	2	3	18
4	-	3	3	20
5	1	1	11	15
6	-	1	11	14
7	-	-	30	13
8	-	2	9	7
9	-	-	2	10
10	-	-	-	23
11	-	-	5	5
12	-	-	11	4
13	-	2	1	12
14	1	-	8	6
15	-	1	5	12
16	-	-	7	15
17	-	1	6	15
18	-	-	2	6
19	1	3	7	7
20	-	1	2	4
21	-	7	2	20
22	1	2	1	12
23	1	1	7	7
24	1	3	2	7
25	1	-	9	2
26	-	-	-	15
27	-	1	8	5
28	-	3	6	8
29	-	3	1	12
30	-	-	-	16
Total	7	47	175	329

Intrusive (r) according to the quality of the preceding vowel

Speaker	After /ɔ:, ɑ:/		After /ə/	
	/r/	∅	/r/	∅
1	3	2	-	13
2	-	-	13	14
3	2	1	1	19
4	1	2	2	21
5	-	4	12	12
6	1	1	10	14
7	4	2	26	11
8	1	1	8	8
9	-	1	2	9
10	-	-	-	23
11	-	-	5	5
12	2	-	9	4
13	-	1	1	13
14	4	-	5	6
15	1	1	4	12
16	-	2	7	13
17	2	5	4	11
18	-	-	2	6
19	1	4	7	6
20	-	4	2	1
21	-	1	2	26
22	-	-	2	14
23	3	-	5	8
24	2	-	1	10
25	4	-	6	2
26	-	-	-	15
27	2	-	6	6
28	1	-	5	11
29	-	1	1	14
30	-	-	-	16
Total	34	33	148	343

Intrusive (r) before clause boundary

Speaker	+ boundary		÷ boundary	
	/r/	∅	/r/	∅
1	-	-	3	15
2	-	-	13	14
3	-	1	3	19
4	-	3	3	20
5	2	1	10	15
6	-	1	11	14
7	6	3	24	10
8	2	-	7	9
9	-	2	2	8
10	-	-	-	23
11	2	2	3	3
12	1	-	10	4
13	-	-	1	14
14	1	1	8	5
15	1	3	4	10
16	-	1	7	14
17	1	-	5	16
18	1	1	1	5
19	4	1	4	9
20	1	1	1	4
21	-	-	2	27
22	-	-	2	14
23	1	1	7	7
24	-	1	3	9
25	1	1	9	1
26	-	-	-	15
27	-	-	8	6
28	-	-	6	11
29	-	-	1	15
30	-	-	-	16
Total	24	24	158	352

R-sandhi according to style

Speaker	Interview style		Reading style	
	/r/	Ø	/r/	Ø
1	18	10	76	85
2	44	18	94	74
3	13	12	77	95
4	16	12	70	125
5	14	2	141	79
6	6	3	112	86
7	23	7	163	72
8	14	7	135	58
9	15	7	81	86
10	19	3	103	79
11	14	10	134	82
12	29	7	157	42
13	26	18	66	93
14	32	5	170	58
15	29	12	134	53
16	21	9	90	90
17	40	17	129	111
18	16	9	130	103
19	24	14	94	112
20	31	12	88	87
21	4	2	90	107
22	-	-	113	98
23	26	6	124	73
24	4	2	120	60
25	1	-	119	54
26	7	-	127	75
27	6	1	143	74
28	3	-	102	102
29	2	5	70	137
30	-	-	45	149
Total	497	210	3297	2599

T-voicing: total scores

Speaker	[t]		[t]		Total	
	N	%	N	%	N	%
1	43	26.7	118	73.3	161	100
2	83	41.1	119	58.9	202	100
3	91	48.1	98	51.9	189	100
4	64	36.6	111	63.4	175	100
5	51	27.6	134	72.4	185	100
6	64	42.4	87	57.6	151	100
7	34	18.6	149	81.4	183	100
8	40	23.1	133	76.9	173	100
9	25	12.3	179	87.7	204	100
10	39	21.5	142	78.5	181	100
11	57	31.8	122	68.2	179	100
12	106	51.7	99	48.3	205	100
13	90	46.6	103	53.4	193	100
14	74	41.8	103	58.2	177	100
15	153	58.8	107	41.2	260	100
16	89	44.3	112	55.7	201	100
17	106	55.8	84	44.2	190	100
18	56	34.8	105	65.2	161	100
19	56	27.2	150	72.8	206	100
20	44	19.6	180	80.4	224	100
21	69	41.1	99	58.9	168	100
22	66	39.3	102	60.7	168	100
23	122	56.5	94	43.5	216	100
24	94	40.3	139	59.7	233	100
25	87	50.0	87	50.0	174	100
26	23	13.9	142	86.1	165	100
27	44	26.7	121	73.3	165	100
28	58	33.5	115	66.5	173	100
29	34	16.3	175	83.7	209	100
30	14	9.7	130	90.3	144	100

T-voicing by word position

Speaker	Word-final		Word-medial	
	[t]	[t̥]	[t]	[t̥]
1	30	28	13	90
2	50	17	33	102
3	57	16	34	82
4	55	13	9	98
5	41	18	10	116
6	39	9	25	78
7	26	32	8	117
8	26	40	14	93
9	22	42	3	137
10	35	21	4	121
11	34	20	23	102
12	57	12	49	87
13	61	14	29	89
14	48	23	26	80
15	104	5	49	102
16	52	17	37	95
17	82	16	24	68
18	47	13	9	92
19	49	20	7	130
20	42	27	2	153
21	40	13	29	86
22	34	29	32	73
23	67	23	55	71
24	59	14	35	125
25	49	6	38	81
26	21	34	2	108
27	34	15	10	106
28	52	40	6	75
29	18	46	16	129
30	13	41	1	89
Total	1344	664	632	2975

Word-final t-voicing according to the length of the target word

Speaker	Monosyllabic		Polysyllabic	
	[t]	[t̥]	[t]	[t̥]
1	29	25	1	3
2	49	14	1	3
3	55	13	2	3
4	53	13	2	-
5	41	14	-	4
6	39	8	-	1
7	26	29	-	3
8	25	33	1	7
9	21	39	1	3
10	35	15	-	6
11	34	19	-	1
12	57	11	-	1
13	61	12	-	2
14	47	22	1	1
15	94	5	10	-
16	45	15	7	2
17	80	14	2	2
18	46	10	1	3
19	46	17	3	3
20	42	25	-	2
21	39	12	1	1
22	32	29	2	-
23	66	22	1	1
24	58	14	1	-
25	45	6	4	-
26	21	34	-	-
27	34	15	-	-
28	50	34	2	6
29	18	45	-	1
30	13	40	-	1
Total	1301	604	43	60

Word-final t-voicing according to stress environment

Speaker	'V__V		'V__'V		V__V		V__'V	
	[t]	[t]	[t]	[t]	[t]	[t]	[t]	[t]
1	10	3	1	1	17	20	2	4
2	17	4	1	2	23	9	9	2
3	10	5	2	1	30	9	15	1
4	15	5	-	1	32	5	8	2
5	8	5	1	-	21	11	11	2
6	8	6	3	-	18	3	10	-
7	11	10	-	-	15	19	-	3
8	7	12	-	2	15	22	4	4
9	8	10	-	1	12	23	2	8
10	7	8	1	1	19	12	8	-
11	12	4	1	3	14	8	7	5
12	23	4	-	-	27	5	7	3
13	22	3	3	2	32	5	4	4
14	7	4	4	4	27	10	10	5
15	20	4	11	1	49	-	24	-
16	15	2	2	2	27	6	8	7
17	27	11	6	1	36	3	13	1
18	11	4	3	2	24	5	9	2
19	18	8	3	2	21	9	7	1
20	13	11	2	3	23	10	4	3
21	3	4	4	3	30	6	3	-
22	4	7	1	3	23	18	6	1
23	19	12	4	1	33	7	11	3
24	22	5	2	2	24	3	11	4
25	9	2	6	3	16	1	18	-
26	3	9	-	5	16	15	2	5
27	2	7	4	2	20	5	8	1
28	6	11	1	4	39	19	6	6
29	3	9	1	6	10	22	4	9
30	3	5	-	6	9	20	1	10
Total	343	194	67	64	702	310	232	96

T-voicing according to style

Speaker	Interview style		Reading style	
	[t]	[t̥]	[t]	[t̥]
1	15	14	28	104
2	37	54	46	65
3	19	11	72	87
4	19	15	45	96
5	19	14	32	120
6	7	5	57	82
7	10	12	24	137
8	9	15	31	118
9	11	26	14	153
10	16	20	23	122
11	14	19	43	103
12	30	10	76	89
13	34	12	56	91
14	11	18	63	85
15	39	18	114	89
16	20	17	69	95
17	43	24	63	60
18	9	11	47	94
19	21	27	35	123
20	19	33	25	147
21	1	2	68	97
22	-	-	66	102
23	33	28	89	66
24	4	1	90	138
25	-	-	87	87
26	4	13	19	129
27	2	2	42	119
28	4	8	54	107
29	6	4	28	171
30	3	-	11	130
Total	459	433	1517	3206

Smoothing: total scores

Speaker	[aə]		[aɪə, aʊə]		Total	
	N	%	N	%	N	%
1	17	36.2	30	63.8	47	100
2	35	59.3	24	40.7	59	100
3	32	66.7	16	33.3	48	100
4	32	56.1	25	43.9	57	100
5	33	60.0	22	40.0	55	100
6	18	62.1	11	37.9	29	100
7	2	3.4	56	96.6	58	100
8	17	53.1	15	46.9	32	100
9	5	12.2	36	87.8	41	100
10	5	16.7	25	83.3	30	100
11	17	27.0	46	73.0	63	100
12	27	71.1	11	28.9	38	100
13	24	50.0	24	50.0	48	100
14	32	72.7	12	27.3	44	100
15	22	51.2	21	48.8	43	100
16	30	68.2	14	31.8	44	100
17	13	36.1	23	63.9	36	100
18	17	35.4	31	64.6	48	100
19	11	37.9	18	62.1	29	100
20	14	25.0	42	75.0	56	100
21	48	85.7	8	14.3	56	100
22	26	70.3	11	29.7	37	100
23	38	88.4	5	11.6	43	100
24	11	33.3	22	66.7	33	100
25	19	57.6	14	42.4	33	100
26	36	62.1	22	37.9	58	100
27	11	22.9	37	77.1	48	100
28	8	21.6	29	78.4	37	100
29	10	23.8	32	76.2	42	100
30	11	23.4	36	76.6	47	100

Smoothing according to the morphological complexity of the target word

Speaker	Complex		Simple	
	[aə]	[aɪə] [aʊə]	[aə]	[aɪə] [aʊə]
1	1	-	16	30
2	-	1	35	23
3	3	2	29	14
4	-	5	32	20
5	-	3	33	19
6	1	1	17	10
7	-	5	2	51
8	-	1	17	14
9	-	1	5	35
10	-	5	5	20
11	2	9	15	37
12	1	-	26	11
13	1	2	23	22
14	1	5	31	7
15	1	3	21	18
16	-	2	30	12
17	2	7	11	16
18	-	4	17	27
19	-	-	11	18
20	1	3	13	39
21	-	-	48	8
22	1	4	25	7
23	1	1	37	4
24	-	1	11	21
25	1	2	18	12
26	3	2	33	20
27	-	4	11	33
28	1	4	7	25
29	-	6	10	26
30	-	-	11	36
Total	21	83	600	635

Smoothing of *-r(e)* items according to the presence and absence of a following linking /r/.

Speaker	+ /r/		+ Ø	
	[aə]	[aɪə] [aʊə]	[aə]	[aɪə] [aʊə]
1	-	1	-	2
2	1	2	-	2
3	2	-	6	4
4	3	2	1	1
5	11	1	-	1
6	2	-	-	-
7	1	4	-	-
8	3	1	-	-
9	-	-	-	-
10	1	1	1	1
11	4	2	-	3
12	3	-	-	-
13	-	2	-	-
14	3	-	-	3
15	4	-	2	2
16	3	1	4	-
17	1	-	-	-
18	1	1	1	-
19	-	-	-	-
20	1	1	-	2
21	3	-	1	-
22	1	-	-	3
23	3	-	-	-
24	2	-	1	-
25	1	1	-	-
26	4	-	2	-
27	-	1	-	5
28	1	-	1	1
29	-	-	-	2
30	-	-	-	1
Total	59	21	20	33

Smoothing according to the presence and absence of a following pause.

Speaker	+ pause		÷ pause	
	[aə]	[aɪə] [aʊə]	[aə]	[aɪə] [aʊə]
1	2	15	15	15
2	4	6	31	18
3	3	3	29	13
4	5	11	27	14
5	2	7	31	15
6	7	6	11	5
7	-	14	2	42
8	7	6	10	9
9	-	4	5	32
10	-	3	5	22
11	5	16	12	30
12	7	4	20	7
13	4	2	20	22
14	3	3	29	9
15	5	12	17	9
16	2	4	28	10
17	1	7	12	16
18	6	9	11	22
19	4	4	7	14
20	3	12	11	30
21	10	5	38	3
22	2	2	24	9
23	10	1	28	4
24	1	2	10	20
25	3	1	16	13
26	2	8	34	14
27	-	7	11	30
28	-	7	8	22
29	-	8	10	24
30	-	11	11	25
Total	98	200	523	518

Yod coalescence: total scores.

Speaker	/tʃ, dʒ/		/tj, dj/		Total	
	N	%	N	%	N	%
1	10	38.5	16	61.5	26	100
2	12	70.6	5	29.4	17	100
3	14	70.0	6	30.0	20	100
4	1	5.9	16	94.1	17	100
5	10	40.0	15	60.0	25	100
6	6	46.2	7	53.8	13	100
7	2	14.3	12	85.7	14	100
8	19	63.3	11	36.7	30	100
9	8	44.4	10	55.6	18	100
10	2	11.1	16	88.9	18	100
11	12	48.0	13	52.0	25	100
12	6	60.0	4	40.0	10	100
13	2	8.7	21	91.3	23	100
14	9	52.9	8	47.1	17	100
15	5	18.5	22	81.5	27	100
16	12	50.0	12	50.0	24	100
17	12	41.4	17	58.6	29	100
18	7	36.8	12	63.2	19	100
19	5	20.8	19	79.2	24	100
20	7	70.0	3	30.0	10	100
21	5	50.0	5	50.0	10	100
22	10	58.8	7	46.2	17	100
23	7	33.3	14	66.7	21	100
24	15	78.9	4	21.1	19	100
25	15	100	0	0.0	15	100
26	5	21.7	18	78.3	23	100
27	10	30.3	23	69.7	33	100
28	24	85.7	4	14.3	28	100
29	14	60.9	9	39.1	23	100
30	20	90.9	2	9.1	22	100