Salience in Sociolinguistics

A quantitative approach

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der Philologischen Fakultät
der Albert-Ludwigs-Universität
Freiburg i. Br.

vorgelegt von

Péter Rácz
aus Budapest

Wintersemester 2012/2013
Indem wir vom Wahrscheinlichen sprechen, ist ja das Unwahrscheinliche immer schon inbegriffen und zwar als Grenzfall des Möglichen, und wenn es einmal eintritt, das Unwahrscheinliche, so besteht für unsere einen keinerlei Grund zur Verwunderung, zur Erschütterung, zur Mystifikation.

(Frisch: Homo Faber)
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>ix</td>
</tr>
<tr>
<td><strong>1 Preliminaries</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Aims and concepts</td>
<td>2</td>
</tr>
<tr>
<td>1.1.1 On the notations</td>
<td>5</td>
</tr>
<tr>
<td>1.1.2 Salience in sociolinguistics</td>
<td>5</td>
</tr>
<tr>
<td>1.1.3 Salience as low probability</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Structure of the dissertation</td>
<td>14</td>
</tr>
<tr>
<td>1.2.1 Methodology</td>
<td>14</td>
</tr>
<tr>
<td>1.2.2 Chapter structure</td>
<td>17</td>
</tr>
<tr>
<td>1.2.3 The case studies</td>
<td>21</td>
</tr>
<tr>
<td>1.3 Concluding remarks</td>
<td>28</td>
</tr>
<tr>
<td><strong>2 Defining Salience</strong></td>
<td>31</td>
</tr>
<tr>
<td>2.1 Salience as a general term</td>
<td>32</td>
</tr>
<tr>
<td>2.1.1 Salience in sociolinguistics</td>
<td>34</td>
</tr>
<tr>
<td>2.1.2 Salience in visual cognition</td>
<td>43</td>
</tr>
<tr>
<td>2.1.3 Selective attention in hearing</td>
<td>48</td>
</tr>
<tr>
<td>2.2 Operationalising sociolinguistic salience</td>
<td>49</td>
</tr>
<tr>
<td>2.2.1 Preliminaries</td>
<td>49</td>
</tr>
<tr>
<td>2.2.2 Defining salience</td>
<td>50</td>
</tr>
<tr>
<td>2.2.3 Exemplars and transitional probabilities</td>
<td>53</td>
</tr>
<tr>
<td>2.3 Concluding remarks</td>
<td>59</td>
</tr>
</tbody>
</table>
## CONTENTS

3 **Methodology** .......................................................... 61
   3.1 General considerations .............................................. 62
   3.2 Step-by-step corpus editing ........................................ 65
   3.3 Calculating transitional probabilities ............................ 71

4 **Definite Article Reduction** ........................................ 75
   4.1 Background .......................................................... 76
       4.1.1 Details of the process ....................................... 76
       4.1.2 DAR as a salient variable ................................... 80
   4.2 Analysis ............................................................. 80
       4.2.1 Methods ........................................................ 81
       4.2.2 Salience from token frequency ............................... 82
       4.2.3 Salience from transitional probability ..................... 84
       4.2.4 Further arguments for phonotactic distinctiveness .... 87
   4.3 Concluding remarks ................................................ 92

5 **Glottalisation in the South of England** ......................... 95
   5.1 Background .......................................................... 96
       5.1.1 Two recent studies ........................................... 97
       5.1.2 Salience and glottalisation .................................. 100
   5.2 Analysis ............................................................. 103
       5.2.1 Methods ........................................................ 104
       5.2.2 The London-Lund Corpus .................................... 106
       5.2.3 The Spoken Corpus of Adolescent London English .... 108
       5.2.4 Modelling results ............................................ 112
   5.3 Concluding remarks ................................................ 116

6 **Hiatus resolution in Hungarian** ................................ 119
   6.1 Background .......................................................... 119
       6.1.1 The perception of hiatus resolution: Methods .......... 123
       6.1.2 The perception of hiatus resolution: Results .......... 124
       6.1.3 Hiatus resolution and naïve linguistic awareness ...... 128
CONTENTS

6.2 Analysis .................................................. 129
  6.2.1 Corpus results ........................................ 130
  6.2.2 Main points .......................................... 132
6.3 Concluding remarks ...................................... 133

7 Derhoticisation in Glasgow ................................. 135
  7.1 Background .............................................. 136
    7.1.1 Social stratification and social awareness ......... 137
    7.1.2 Derhoticisation in Glasgow ......................... 140
    7.1.3 /r/ in Glasgow ...................................... 141
    7.1.4 Studies on coda /r/ ................................. 153
    7.1.5 Interim Summary .................................. 160
  7.2 Analysis .................................................. 162
    7.2.1 The FRED study .................................... 163
    7.2.2 Transitional probabilities in coda /r/ realisation . 165
  7.3 Concluding remarks ...................................... 169

8 Salience and language change .............................. 173
  8.1 Speaker indexation in sound change .................... 175
    8.1.1 Approaches to speaker indexation .................. 176
    8.1.2 Simulations on the role of indexation .......... 179
  8.2 Salience in the propagation of a change ............... 185
    8.2.1 Glottalisation in England ......................... 186
    8.2.2 Derhoticisation in Scotland ....................... 188
  8.3 Concluding remarks ...................................... 189

9 Conclusions .................................................. 191
  9.1 The source of salience ................................ 191
    9.1.1 From cognitive properties to language use .... 192
    9.1.2 Consequences for phonological modelling ....... 194
  9.2 The predictability of salience ........................ 196
    9.2.1 Types of phonological change .................... 196
CONTENTS

9.2.2 Consonants and vowels ............................... 198
9.2.3 Overview ............................................. 199
9.3 Concluding remarks ................................. 200

German Summary – Deutsche Zusammenfassung 221
Foreword

I used to have a little game with people in Freiburg resolved to know what I do for a living. I asked them to name a few differences between their own native dialect (usually Standard German with some Southern colouring) and Swiss German. Now, as no second language German learner would hesitate to tell you, the Alemannic dialects subsumed under the moniker of Swiss German differ from Standard German in everything. This only made it more interesting to see what speakers of the less self-conscious of the two dialects in question, Standard German, would single out.

Unsurprisingly, most people mentioned various lexical items, such as the ever-recurring Swiss greeting, Grützi – which is regarded so much as a stereotype that the Swiss supermarket chain Migros has Grützi written on its shopping bags in Germany – as well as examples like Velo for Fahrrad, Töff for Motorrad, or Depot for the centrally German notion of Pfand.

These examples are, in a way, less interesting, not only because they come from the most conscious linguistic level, the lexis, but also because they might come up in Swiss Standard German, the variety of Standard German used in German-speaking Switzerland. To a phonologist, the really exciting answers concerned the differences in how the two languages sound. Someone was able to point out, though not in such elaborate terms, that Swiss German uses a diphthong instead of a monophthong in words like müsli (müsli in Swiss German), but monophthongises Standard German /œi/ and /ɔi/ – hence the Swiss German name for the language, Schwyzerdütsch, instead of Schweizer Deutsch.
A feature that many people mentioned, though, was that Swiss German is full of ‘weird throat sounds’. The state of affairs they likely had in mind results from two independent processes. On the one hand, there is [kx], which, due to the proper implementation of the Second Germanic Consonant Shift, occurs word-initially in words where the Standard German cognate has a [k]. The realisation of this sound varies in German-speaking Switzerland: Basel German does not have it at all, and in quite a few dialects it is realised as [x]. On the other hand, /x/ lacks its (Standard German) allophone [ç]. The diachronic processes responsible for the spirantised stop and the absence of the palatal allophone proliferate velar fricatives in Swiss German dialects. (Calling all these fricatives ‘velar’ is somewhat simplifying: In Bern, for instance, /x/ is usually realised as a uvular [χ], making it even more prominent for German ears, but that is more or less irrelevant here.)

The omnipresence of the velar fricative, then, seems like a salient characteristic of Swiss German, at least to speakers of Standard German – so much so, that it is also giving rise to explicit commentary, as in the following joke: –Wie sagt man Banane in der Schweiz? –Banane-ch. An odd thing about this, though, is that Standard German itself has the weird throat sound as a regularly occurring, contrastive segment in words like Buch ([bux]) or machen ([maxen]). This certainly does not help Chuchichäschtli ([xux:ixæʃtlí]), the affectionate Swiss German term for ‘kitchen cupboard’, but we still have to say that the source of these comments on Swiss German wallowing in velar fricatives is not its absence from Standard German. It has to be sought elsewhere.

While phonetic differences, like in the Bernese example, certainly play a role, one likely cause of this perceptible difference between Standard and Swiss German is the relative frequency of the velar fricative in the latter compared to the former. It occurs more often, and, more importantly, with a distinct distribution: it might, for example, show up word-initially or after front vowels, which it would never (or hardly ever) do in Standard German. A difference in frequency, then, seems to be a good reason to find some other
dialect saliently strange.

Asking people to report on their perception of how other people speak is obviously quite hazardous. For instance, intonation might be a very good marker of a dialect, but speakers benefitting from a letter-based alphabet are more in command of the vocabulary to talk about segments than about supra-segments.

That said, a difference in consonantal distributions can be a reliable source of salience between related dialects – even if it is obviously not the only one. This work looks at cases where it is not the presence or absence of a particular sound segment that is used as a dialect marker, but the difference in the distributions of its realisations. To give another example from English, debuccalisation, glottal realisation of the fortis coronal stop /t/ is strongly stigmatised in the South of England, word-medially between vowels (as in [siʔɪ], [lɛʔɔ] for ‘city’, ‘letter’), but no-one seems to care if /t/ ends up glottalised word-medially before a consonant (as in ‘Atlantic’, [æʔlæntɪk]), or even word-finally. The glottal stop is there in both cases, but it shows up in a different pattern. Speakers seem to find it difficult to pick up on a difference in one instance, when it is phonetically fine-grained, but are strongly aware of the same difference in another, where its distribution is more bi-modal, as is the case with rhoticity in Glasgow and in New York City.

Examples like these make more sense if we assume that distributional differences of near-categorically realised segmental variants result in speaker awareness to a particular dialectal variable. This might sound nebulous at first, but the present work sets out with the precise aim of fleshing out the idea.

Despite the Swiss red herring, this book concentrates on English dialects, mostly because the relevant literature is predominantly based on evidence from English and because English dialects are indeed very well documented, making it easier to base even more literature on them. The focus is on phonological variation, which is a good thing, as most variation is actually phonological and sociolinguists do tend to pay a larger amount of attention to
sounds. My examples are almost all differences in consonants. This might be seen unfortunate, after all, a large amount of the difference people perceive between dialects is due to dissimilarities in vowel realisations. As a friend of mine from Sheffield commented on the legendary North-South divide in England, ‘it is all in the vowels’. Sociolinguists seem to have shared his sentiment, as considerable amount of attention is given to vowels in the literature. Looking at it this way, a book on consonants is, to an extent, filling a gap. By the end of this work, I will hope to have convinced the reader that consonantal differences are relevant as well.

In this book, I embrace a segmentally-based approach to salience in sociolinguistics. While such an approach clearly has its own shortcomings, it allows for the large-scale investigation of dialects and dialectal variables, and can thereby successfully extend our understanding on what constitutes the salience of markers employed in one dialectal system in relation to another.
Chapter 1

Preliminaries

Salient

adjective

1. Of material things: Standing above or beyond the general surface or outline; jutting out; prominent among a number of objects.

2. Of immaterial things, qualities, etc.: Standing out from the rest; prominent, conspicuous; often in phr. salient point. Also Psychol. standing out or prominent in consciousness.

The Oxford English Dictionary

This work is written with the aim of clutching the notion of salience in sociolinguistics. There are two main questions to tackle: (i) what sort of differences in speech are salient for the language users and (ii) how these differences are utilised by the language community, the latter question having consequences for a theory of language change. Differences in sound patterns are in the locus of the investigation. This is not only due to the background of the author, but also because most of the discussion on salience in sociolinguistics
is largely confined to the phonological domain. The examples quoted are mostly differences in consonantal pattern. As we will see, vowels seem to work rather differently, both with respect to salience and sociolinguistic variation.

1.1 Aims and concepts

Differences in speech are approached from a sociolinguistic perspective: such a difference is considered as a dialectal variable if it is able to mark social or regional differences – or, to put it more generally, group differences of any discernible form. (Weinreich et al. (1968) define a linguistic variable as alternative ways of saying the same thing, adding these ways have social significance.) A speech difference can convey a difference in meaning (where two forms have two different referents) and in social context (where two forms have the same referent but are used by different kinds of people). Here I regard the latter as incidental, as contrastiveness, in the structuralist sense, plays a minor role in the argumentation. Indeed, most of the phonological variables we find in sociolinguistics are solely employed to mark social, not semantic differences. (This is true to the extent that, for a given form, variation in rhoticity does not lead to variation in meaning. Of course, common patterns of change like phoneme splits or neutralisation can obscure the picture.)

I will argue that salience is derived from patterns that surface in language use. A sociolinguistic variable can become salient if it is unlikely or unexpected for the listener, in comparison with the listener’s own dialect or a projected linguistic norm. (The two do not necessarily have to be very far from each other.) The property of being unlikely is interpreted at the phonetic level, at the level of the speech stream: a variant is ‘surprising’ if it has a low transitional probability, i.e. a low expected occurrence in the speech stream, at least in the norm dialect, which is used as a basis of comparison. Such an approach implies a strong commitment to a segmental view of sound patterns, but we can be careful not to ignore phonetic detail and variation.

Two things need be confirmed to go on in this direction. On one hand, it
has to be shown that a particular dialectal variable indeed has a low probability of occurrence when comparing two dialects. (The two dialects would be typically a vernacular and a standard, with the investigation focussing on the salience of the vernacular realisation. To make things more confounded, things can happen the other way around: a variable can be salient because one of its realisations has a low transitional probability in the norm dialect.) On the other hand, it has to proven that language users are sensitive to this type of information. Needless to say, we have to see that the dialectal variable is actually salient for the language community, that is, speakers are, to an extent, aware of its significance in social indexation. This leads us to the second research question.

Salience, as a concept, is used straightforwardly in the cognitive sciences, as in discussions of visual perception, or even in specific, practical sub-domains, such as advertising (Itti et al., 2005). In comparison, the sociolinguistic notion is rather vague. Most of the literature on the subject does not widely part from the dictionary entry above, and regards salience as a sort of perceptual prominence. My argumentation starts with a simple definition for the concept: a variable is salient if it is consciously or unconsciously used for social indexation. Some variables consistently correlate with geographic region, social status, gender, or other group notions, but not all of them do so in an active manner. A salient variable can be recognised from hypercorrection, dependence on register, and listener attitudes.

The concept of a salient variable, articulated in such a way, bears a strong resemblance to sociolinguistic markers, also used for indexation and recognised from inter-register variation and listener attitudes. Salience, however, will not be used as a fancy synonym for a marker. First, it, as a term, connects the sociolinguistic use to the general cognitive linguistic domain. Second, as I show in the case studies, it is a property a marker has, but not equal to the marker itself.

The research plan embraced here is to list a number of detailed case studies from English phonology, though not brushing aside the concept’s general
applicability in other languages. In these studies it can – hopefully – be demonstrated that low probability of occurrence correlates with sociolinguistic salience. The thorough investigation of these variables also shows that what we observe is not mere correlation, but an instance of causation. In the case of these variables, their salience is derived, at least partly, from the low probability of occurrence of the available variants when comparing two dialects, such as a vernacular and a standard.

Linking up salience with probabilities in production is in the ethos of usage-based functionalism. This research trend has been lately gaining ground in phonology (Silverman, 2006; Wedel, 2004) and sociolinguistics (Foulkes & Docherty, 2006). The central tenet of usage-based functionalism is that linguistic structure is directly influenced by language use, and that most of the complex language patterns eventually emerge from a few basic principles of how linguistic data are produced, processed, and stored. This stands in stark contrast with structuralist theories of language. These theories are committed to a separation of language structure and language use, with the latter never influencing the former directly. Since talking about low probability of occurrence only makes sense if we take the flow of spoken and written language into consideration, the research plan outlined here is fundamentally usage-based in nature.

One important question that follows from a study of salience in variation is the role of salient versus non-salient variables in language change. This implicitly relates to the extent of social use as a factor in change, and consequently, on the possible difference in the behaviour of salient variables, as opposed to non-salient ones, in a model of language change. An evolutionary framework like Croft’s (2000) readily lends itself for the incorporation of the salient/non-salient distinction, and such an addition can relevantly contribute to our view of language change couched in social practice.

Finally, the theory of salience proposed in this work hinges on quantitative differences between dialects. This means that the relative frequency of a segment empowers a marker with the ability to carry social indexation. In
this sense, this work also contributes to the study of frequency effects on language structure.

1.1.1 On the notations

Throughout this book, oblique parenthesising marks an abstract segment type and simple parentheses a sociolinguistic variable. Square brackets refer to a particular segmental realisation and chevrons to an orthographic form of either an abstract type or a variable. For example, in Chapter 7, I talk about /r/, which is a contrastive segment in Scottish English. It can be realised in various ways, such as \([r]\), a central approximant, \([\nu]\), a retroflex approximant, or \([\text{r}]\), an alveolar tap.

If we posit a variable on /r/ realisation, we can call it (r). We can then posit all the above realisations as variants of (r), along with the vocalised, deleted variant [V] word-finally. (Note that this one is more abstract and not directly interpretable.) We can posit another variant, (rV), which only includes word-final variation in /r/ realisation, and we can distinguish other realisation variants for it in the vocalised variant group. Needless to say, all these units will be written as <r>, because English orthography is insensitive to the subtleties of rhotic variation. Finally, SMALL CAPS are used to refer to lexical sets (Wells, 1982). The notations express no commitments on the ontological status of these segments, but it is important to keep them apart.

1.1.2 Salience in sociolinguistics

If we are to consider the place of salience in sociolinguistics, we might as well start at the very beginning, with the dictionary definition of the concept in itself. Something is above the outline, jutting out. This description, interpreted in a linguistic context, begs a precise definition of the unit that possesses or lacks the ability to jut out, as well as the way this jutting out takes place. The first definition entirely depends on us. In this particular work, the focus is on phonology, so I will refer to units of segmental phonology,
suprasegments, segments, and segmental features. Bearing in mind that these are only abstractions on the recurrent patterns of the speech stream, we are able to formulate meaningful statements on their behaviour, as in ‘a /t/ is glottalised word-medially in the South of England, but not typically in the North.’

Pointing at categories established since the early structuralists is, however, not enough. It is crucial to see that the salience of a linguistic variable lies – not in its meaning or reference – but rather in its social indexation. (In the case of a contrastive sound segment, reference can be interpreted as lexical contrastiveness, that is, that [baɾ ʰ] means something else in English than [baɾ], whereas it means the same thing as [bad], but has a different social indexation.) To quote a classic paper on the social functions of phonological patterns, Foulkes & Docherty (2006), ‘no natural human utterance offers linguistic information without simultaneously indexing some social factor.’

A simple example to demonstrate this comes from the Australian language Djirbal (Dixon, 1980). Djirbal has two vocabulary sets, one for general use, and one to use in the presence of a particular kinship relation, which Dixon dubs as ‘mother-in-law’. Though the ‘mother-in-law’ set is more restricted, common words exist in two forms, one for each vocabulary set. Independent of the speech context, the Djirbal speaker has to use the ‘mother-in-law’ set in presence of such a kinship relation. What follows is that a Djirbal word has two functions: a meaning, and a social index, namely, the presence/absence of the mother-in-law.

This is a categorical example: all words either belong to one set or the other, and speakers are able to classify them upon inquiry (Silverstein, 2001). We have reasons to believe, however, that social indexation is usually not that robust. While different words usually signify different concepts (except for interjections, homonyms, etc.), not all dialectal differences index different group alignments, at least not to the layperson. There are consistent dialectal differences which are never recognised as such. Once more returning to the Djirbal example, we could imagine that while words are consistently used
differently in the two alignment contexts, this difference, in the case of some words, is not recognised by the speakers themselves. That is, they do them, but they are not aware of them.

To recapitulate on all this once again, two dialects can differ in \( n \) variables, but only a subset of these variables, \( k \), is recognised by the speakers (not necessarily consciously). The rest of the variables, while they also differ consistently, are ignored.

This fundamental dichotomy in social indexation is first established by the variationist sociolinguistic tradition (Labov, 1972b). According to Labov, an indicator is a difference in speech that systematically occurs in two lects, but it is not recognised by the language users. A marker is a difference in speech that systematically occurs in two lects, and is, consciously or unconsciously, recognised by the language community. The reason why we suppose that an indicator is not recognised as a unit of difference is that it shows social/region stratification, but no style shifting or hypercorrection. Speakers’ attitudes are neutral towards it: they will not recognise its use as a mark of a particular dialect or sociolect, and will not try to avoid it when attempting to speak the standard. One example is \([a:]\) in Norwich (Trudgill, 1986). This vowel is more fronted than the standard variety, but the speakers seem to be unaware of this difference. Other examples include different kinds of tapping (the use of \([r]\) or \([d]\) for word-medial /t/) or the interdental realisation of the dental fricatives /\(\theta/\) and /\(\delta/\) in American English dialects (Ladefoged & Maddieson, 1996).

A marker is recognised as a unit of difference, though perhaps not consciously. That is, speakers might not be able to overtly point at it as a speech characteristic, but will nonetheless have pronounced attitudes towards it. One example is the Northern \([a]\) (Wells, 1982). In the North of England, this sound is restricted to a set environments indicated by a following \(<r>\) in the orthography (e.g. ‘carton’, ‘bar’). In words like ‘dance’, ‘fast’, a fronted \([a]\) is used instead. This is a strong marker of Northern speech and a stereotype among Southern speakers.
Preliminaries

Above I follow the variationist tradition in assuming lects to be associated with broad speaker groups, defined on a regional/social basis (as well as age and gender). This is, to an extent, a simplification: reliable linguistic group identification can be found in much smaller societal units, though it is true that social indexation is the most robust and reliable when looking at larger ones.

Sadly enough, the situation with indicators and markers is not so clear-cut, not even at the broad level of social class or age distinctions. Some variables can vary with the context without proof that they are under any (even) unconscious control of the speaker. Mendoza-Denton et al. (2003) show how the use of a variant associated with African American Vernacular English (AAVE) correlates with the familiarity of the context in Oprah Winfrey’s speech. The variant is the monophthongised realisation of the AmE variable /aɪ/.

Essentially, Oprah uses the AAVE variant more when addressing the audience directly or trying to create an informal atmosphere. Monophthongisation correlates with the context, but this does not inevitably mean that it would be discarded in a formal situation, i.e. reading a word list or giving an interview. Becker (2009) illustrates a different scenario with the use of /r/ in the Lower East Side – this feature is sensitive to both the register (formal/informal) and the dialogue context.

Chopping up the linguistic variable set into indicators and markers is dangerous in the sense that it implies a complete absence of gradience. It has to be noted this early that the binary approach is only one of the possible operationalisations. As Preston (1996) shows, linguistic awareness has many levels, very few categorical. The layperson’s grasp of dialect differences can differ in availability, accuracy, and detail: one can comment on an accent without being able to describe its characteristics, one can point at a particular variant but define it inaccurately, and one might not be aware of a recurring difference at all. What is more, a variable can have a large scope of realisations showing rampant variation and fine-tuned phonetic differences, making the
listener’s job even more difficult. The concept of ‘marker’ here will refer to a variable that affects speaker attitudes and behaviour. It shows style shifting and invokes positive or negative attitudes, even if the language users are absolutely unable to pin it down.

Salience is here interpreted as a property that allows a linguistic variable to be a marker. In the case studies, attitude tests are used to determine whether it is used in the speech community for indexing purposes. As the review of the literature on this concept will show, this is a rather narrow definition. Nonetheless, it is apt to base a research methodology on it. (Of course, when a variable is an extremely strong stereotype, witnessed by overt social commentary, the importance of attitude tests diminishes – if people are able to point out which feature is weird, they probably already noticed it.)

1.1.3 Salience as low probability

Whether a variable is salient or not for the speech community can be determined by independent measures. The best tools are attitude studies, which clearly show whether listeners associate the presence/absence of a variant with a particular geographical location or social stratum. Style shifting, hypercorrection, and, of course, non-linguist comments all strongly suggest that a particular variable is a sociolinguistic marker. Which variables are chosen to serve as markers is a wholly different matter.

Originally, Labov (1972b) proposes that all variables start their life spans as indicators, and later, as the particular language change gains momentum, propelling them further, they become markers for the language users. In subsequent work, Labov (2001) discards this hypothesis, noticing that some variables never reach marker status. (This is explored in the Chapter 2 in more detail.) If this is so, we can either believe that markers are selected at random (in which case there is not much point in reading this book any further) or start to look for a possible perceptual/cognitive explanation for salience.

The hypothesis assumed here is that salience derives (at least partly)
from low probability of occurrence. An unexpected, surprising variant can be utilised to index social differences – though not necessarily. The merits of this hypothesis are that it can be operationalised easily and that it provides an empirical foundation for the concept of salience in sociolinguistics. It also has its limitations: in cases where probability of occurrence is more complicated to calculate (or simply borders on the impossible), the hypothesis is inapplicable. Vowel shifts are a clear example to this. Ongoing, socially stratified vowel shifts have a strong social significance, but this does not trivially follow from any difference in distributions – vowel qualities change, but quantities stay the same. I have a bit more to say on the issue in Chapter 9.

One concrete example of the correlation between low probability and salience is definite article reduction, a dialectal process in the North of England (Lodge, 2010; Jones, 1999). To put it very simply, the process entails the variable reduction of the definite article into a glottal stop. This reduction pattern is confined to the North of England and it is not particularly frequent, most surveys giving figures of ten to fourteen per cent of all definite articles reduced. The pattern varies with age and gender, and shows style shifting. What is more, it is an overt stereotype of Northern speech, making it an exemplary salient marker.

If we look at the distribution of glottal stops in the ‘norm’ dialect (such as Northern varieties of RP or indeed any English English dialect without definite article reduction) we see that glottal stops have much larger transitional probabilities in particular positions in the reducing dialects vis à vis the standard. For example, we find a glottal stop followed by a stressed vowel much more rarely in the standard as in a reducing dialect (where definite articles standing before vowel-initial words stand a chance to be reduced). What follows is that the reduced articles are in conspicuous positions when compared to their distributions in the standard dialect, which, we want to argue, leads to their salience. (I borrow the term ‘English English’ from Jane Stuart-Smith and Norval Smith, who rightly argue that ‘British English’, at this level of comparison, is a misleading and superfluous term)
While article reduction is relatively frequent as a morphological variable (definite articles being one of the most frequent words in corpora), it is nowhere near the robustness of phonological variables, therefore its marker status is in dire need of an additional explanation.

This was a simplified discussion, and I return to definite article reduction in Chapter 4. The principal reason of bringing it up here is to illustrate the concept of salience as low probability. It serves as a good illustration of the kind of pattern which is easily testable under our assumptions. It also helps to clarify the relationship of the concept ‘surprising’, ‘salient’, and ‘marker’. The variable’s salience follows from the variants’ low transitional probability or ‘surprisal’, which, in this case, is clearly measurable. The property of salience is essential for the variable to become a marker, that is, take up social indexation.

How surprisal is measured is crucial to the whole undertaking. I talk about modelling the salience of a variable in length later.

As stated above, the locus of discussion throughout this work is phonology, only partly because most of the literature on salience explores this domain. Phonology also offers clearcut ways of abstraction and segmentation, allowing us to make comparisons of the above kind relatively easily. Last but not least, most sociolinguistic variables are phonological. This is no surprise – slight changes in pronunciation are excellent tools to mark social indexation while preserving distinctions in semantic reference. (That is, it is easy to say a word in two slightly different ways where the difference is big enough to be noticed but small enough so that the two ways still mean the same thing.)

According to the typology of Wells (1982), phonological differences between dialects can be segmental, suprasegmental, or subsegmental. **Segmental** variation can be of four sorts. In the following, I borrow Wells’ terms to describe his categories, which will be tacitly assumed through the rest of the book.
Systemic  Systemic variation is the presence or lack of a particular phoneme in a dialect as opposed to another. Hiberno-English has the voiceless labio-velar approximant \([\text{a}]\) in words like ‘when’, ‘which’, whereas Southern English lacks it.

Phonotactic  A phonotactic distribution can also vary between dialects: Rhotic varieties of English, like most Northern American accents, have [r] everywhere, whereas non-rhotic varieties, like Southern English, restrict it to pre-vocalic position.

Lexical  Language variants can differ in the use of a particular phoneme in a particular set of words. The above example of the distribution of [a] and [æ] in Northern and Southern English is a case of lexical difference. The Southern variety uses the back vowel in more lexical items, where the Northern variety has the front vowel.

Allophonic  Allophonic variation entails that a particular pattern is restricted to a different set of environments in one dialect as opposed to another. Glottal replacement of [t] in England is such a case: while it is a norm – at least in lower registers – to use a glottal stop instead of [t] word-medially in urban accents in the South, it is a relatively unfamiliar pattern in the North.

Besides these, Wells notes suprasegmental differences, essentially differences in intonation and stress, and subsegmental differences, observable in the internal structure of the sound segments themselves. For instance, the distribution of Hiberno-English coronal stops ([t],[d]) is relatively similar to Southern English (apart from the odd glottalisation) but they are consistently produced as dental rather than alveolar. (For the sake of completeness, it has to be noted that Wells here talks about dialects, but, based on his definition, people will very likely use differing accents, the norm with a local pronunciation, when talking to users of another dialect. Since this view defines accents
rather broadly, and in any case, means strong differences only in syntax and lexis, we can ignore it.}

This typology is heavily taxonomic in nature. The reason why it serves as the basis of variable selection for us is not that we rather bear those ills we have than fly to others that we know not of. Rather, it is an effective way to abstract and operationalise phonological differences in dialects and its limits might prove to be useful in the long run. Namely, if one is to calculate probability of occurrence, one needs a tangible unit of reference. As probabilities are necessarily based on corpora, a unit of phonetic transcription is the ideal one. The presence or absence of such a unit in particular environments is easy to compute. We can straightforwardly look at the environments of realised [r]-s in a partly rhotic dialect. It is much more difficult to look at different place of articulations in a dialect which partly realises [r]-s as alveolar or retroflex. It is equally difficult to compare instances of granular reduction, or reduced articulatory gestures.

Nonetheless, examples of ongoing language change usually cross-cut all these categories with complete contempt for orderliness. One case is the loss of coda /r/ in the Southern urban belt of Scotland, where dialects differ from each other in phonotaxis, place of articulation, and the presence or absence of contrastive variants, all in a phonetically fine-grained and highly intertwined manner. As I will attempt to show, this example actually supports a segmental approach to salience, handled with due care, as even speakers seem to have difficulties in noticing gradual phonetic differences. A further argument for narrowing the scope to segmental variation is that we do have psycholinguistic evidence on listeners’ sensitivity to the frequency of sound segment-sized units (Jusczyk et al., 1994; Saffran et al., 1996b). I will return to this point in Chapter 2. For the time being, it suffices to say that the use of segmental units is not only justified by the fact that they are easier to handle, but also by the distinct attention they were given in the literature, which might ultimately stem from their status as the smallest possible units that language users pick up as salient dialectal variants – or, to put it a bit
more boldly, minimal units of speech. The influence of segmental writing systems on speaker phonology is a can of worms I would like to avoid. If vocal adherents of the non-segmental approach, like Silverman (2006), who argue that any segmentation in phonological studies is an alphabetic bias, prove to be completely right, the approach introduced here will at least be applicable to languages with segmental alphabets and a strong written tradition. This is because, at least in these languages, the combination of a segmental alphabet and high literacy allows people to rely on segment-based units for social indexation. This still leaves us with an extensive working sample.

1.2 Structure of the dissertation

1.2.1 Methodology

After establishing the sociolinguistic and phonetic-phonological correlates of salience, the procedure is to take a look at different phonological variables that arguably have the former and see if they also share the latter. That is, for a variable $V$, we have evidence that $V$ is used for indexation, and are curious whether there is a significant case of jutting out in $V$’s variants’ transitional probabilities (in an inter-dialectal comparison). The unpronounced assumptions behind this are that though a salient variant will certainly have low probability of occurrence, a variant with a low probability of occurrence need not be salient, and that variables that are not sociolinguistically salient in the first place should be excluded.

The reasons for the first assumption are twofold. First, low probability is not the sole factor in salience. A variable’s promotion to marker status depends on its position in the language community and the attitudes present in this community in general. The social evaluation of a variable (or lack thereof) is necessarily entangled with external factors such as the structure of the language community, and the thorough consideration of these factors is beyond our scope. Suffices to say that my examples got pushed to the stage
by social dynamics, and the non-linguistic forces behind them are more or less taken for granted. Second, a low probability variant can be so unlikely to occur that it does not reach the threshold of consciousness at all, in which case it is ignored completely – low probability of occurrence needs to be paired up with a certain amount of robustness in the speech signal. Putting it differently, a sequence has to be frequent enough for people to notice how rare it is.

The second assumption is a safeguard with respect to the caveats of working with indicators. As soon as a difference in speech displays in speaker attitudes, we can safely categorise it as a dialectal variable. Indicators have a status as beings duly noted by linguists but by nobody else, and while most of these are clearly distinguishable – albeit only by dialectologists – others scarcely exist at all. With these assumptions in mind, research is exclusively concentrated on salient phonological variables, mostly from the segmental domain. Constructing an argumentation in favour of the causation between low probability of occurrence and salience then proceeds as follows.

1. Identification of a variable: A dialectal variable is picked and defined narrowly. The latter step is quite useful inasmuch as the use of different terms in the literature can be rather fuzzy. For instance, ‘glottalisation’ might refer to the glottal reinforcement of fortis stops attested in dialects in England (as in [kʰæʔpʰæn] ‘captain’, the glottal replacement of /t/ in some subsets of these dialects (as in [baʔmæn] ‘Batman’), or even solely to the glottal replacement of /t/ in a number of very specific environments, such as word-medially before a vowel ([lɛʔa] ‘letter’) or before a syllabic sonorant ([baʔl] ‘bottle’). Similarly, the term ‘rhotic’ comes up both in the sense of a rhotic/non-rhotic distinction (as in Scottish Standard English versus Southern English English) or speaking of rhoticised vowels in American English, and so on.

2. Background check: The literature on the particular variable is partly surveyed. This is not only crucial to understand the mechanics of the
feature itself, but also necessary to confirm its salience. A marker shows style shifting and hypercorrection and it might be the subject of overt commentary and ridicule, which, for us, is even better. Fundamentally, it should trigger changes in attitude, if it is – perhaps covertly – associated with a particular region or social register. These attitudes can be caught in the act by sociolinguistic interviews and attitude studies (cf. e.g. Labov et al. (2006).

3. Choice of corpora: In order to find out about difference in transitional probabilities between the dialect featuring the marker and the respective standard, one must find corpora of both to work with. As written registers force standardisation, spoken material is preferred. In case the spoken material is only available in a transcribed form, but without recordings, the ratios of the marker can be extrapolated using the transcription and relevant studies on the extent of the marker’s use.

One corpus to use when working on variation in England is the FRED corpus (Kortmann et al., 2005), a collection of sociolinguistic interviews covering all major English dialect areas. The main merit of this corpus is that data were collected consistently and in a similar fashion from all major dialect areas, allowing for a straightforward comparison. It contains recordings and transcriptions, which renders it even more suitable for any investigation of phonological variation. In case more data are needed from a particular dialect area, auxiliary corpora can be put to use – to give two examples, Cheshire et al. (2008) have a large corpus of adolescent speech in London, and there are similar projects for other dialect areas, like the Newcastle Electronic Corpus of Tyneside English (Allen et al., 2007) for the Newcastle and the Tyneside.

4. Computing probabilities: Differences in transitional probabilities for the dialect with the vernacular variant versus a standard or norm dialect (with the standard variant) are calculated. In the case of Northern English, for instance, the looming presence of RP as the national
standard gives us enough reason to compare a dialectal variation pattern with it. The internal speaker norm can be, however, inferred in most cases from attitude studies and the sociolinguistic makeup of the area. The final result is a model of the dialect and a model of the norm, but these models are reliable enough for us to make meaningful statements about them. If there is a high enough difference in probabilities between the dialect and the norm, we are safe to say that this gives the variable its unexpected, surprising quality, and, indirectly, its salience. Salience does not necessarily emerge in a standard-vernacular relationship, but our examples will mostly belong to this common type.

Apart from relying on the literature of the different variables in particular and of salience and psycholinguistics in general, this research method employs two sets of tools: corpora and computing transitional probabilities. The choice of corpora has to be carefully made, and backed up with independent studies on the reliability of the corpus, as well as the behaviour of the feature in the dialect. The use of transitional probabilities in linguistics has a long tradition dating back to Harris (1955) and the structuralists, but as putting them to this particular use is, as far as I am concerned, novel, one has to proceed with caution. Needless to say, both the question of corpora and the methods of computation are elaborated in more detail later.

1.2.2 Chapter structure

Chapter 2, a review of the literature on salience, follows this introductory chapter. I have a brief look into how the concept is used in the cognitive sciences, especially with respect to visual cognition. The similarities and differences between cognitive salience and the way sociolinguistic salience are interpreted are going to provide a useful starting point for any further discussion.

The more substantial part of the chapter deals with salience in sociolinguistics. As I wrote above, my definition of the term is rather narrow.
Sociolinguists regard the concept of salience as having a wide range of properties, from merely equating it with high token frequency to attributing an extra-linguistic property to it. A partial survey of the available studies working with this concept is handy since it provides a feasible background for the use of the term here. It is also a starting point for our discussion. This is vital as salience has to be defined, on one hand, using the general toolkit of sociolinguistics, but, on the other hand, it has to be also placed before the background of previous studies using the concept.

It is one thing to tell the reader what salience is, but one has to pay attention to what it is not. For instance, salience should not be equated with sheer token frequency. It is also dodgy to mix it up with another vaguely defined concept, markedness. Markedness is, in principle, the theory of sound segment complexity. It says that some sounds are more complex than others, and this shows in their phonetic properties, acquisition, and behaviour in sound changes. One could then simply say that the more marked a sound gets, the more salient it is. Markedness theory, however, has a huge amount of problems on its own (Harris, 2005; Ohala, 1971; Blevins, 2004), and its relationship to salience is vague to say the least. This part is important as markedness seems like a reasonable idea to anybody outside phonology (and even to some inside it), but simply treating it together with salience can be a dead-end to any research on either.

To cite one issue, a frequent example of markedness – as a principle operational in language – is the case of consonants with two places of articulation. You would not expect to find a language with a labio-velar stop [kp] but without a labial and a velar stop. Similarly, you would think that a labio-velar stop will be more salient for the listeners than a labial or a velar stop on its own. That might very well be the case, but the system easily breaks down with more complicated examples. For instance, the Southern English English /t/ has two usual realisations, a slightly affricated one ([tʰ]) before (stressed) vowels and a debuccalised one ([ʔ]) everywhere else. The affricated realisation is far more complex and should be deemed more marked in any case (as the
[?] results from lenition) Yet, listeners seem to be sensitive to the glottal stop, if anything, and, on the whole, ignore extents of affrication. This is quite at odds with our markedness-based intuitions. Similar examples are not hard to come by.

Finally, Chapter 2 elaborates on a theory of salience as low probability of occurrence. This starts with a review of the literature on transitional probabilities. If I am to argue that probabilities extracted from the speech signal are largely to blame for the selection of markers, I have to prove that listeners are sensitive to this type of information. This is far from self-evident: Zeilig Harris originally proposed reliance on transitional probabilities as a method for the field linguist.

His idea was that since word-medial segmental sequences are subject to phonotactic restrictions, while word-internal ones are not, the transitional probability of a segment X following Y (X|Y) will be higher word-medially. Consequently, low transitional probabilities suggest word boundaries. For example, English [t] can be only followed by liquids and glides in an onset, but by practically anything if it stands at the end of the word. If the field linguist is unable to segment an unknown language, they can use transitional probabilities between the segments to determine word boundaries.

It probably never occurred to Harris that the language user can take track of probabilities of any sort. The distribution of sounds in the signal (or, as we see it, in a larger corpus) belongs crucially to the parole, as it follows entirely from the phoneme inventory and morpheme structure constraints of the language, as well as from the way words are concatenated – so syntax and word order also play a minor role.

The structuralist assumption, then, is that the structure of the langue affects the shape of the parole, and the observant field linguist can exploit this to arrive at generalisations on the former. This path is shut off for the language user, however, as parole has no way to have an effect on langue directly. Consequently, since probabilities of occurrence are only available from the latter, they do not count as linguistic information. The psycholinguistic
Preliminaries

tradition, however, clearly supports that listeners are sensitive to transitional probabilities from a very early age, as they provide a vital cue for word segmentation (precisely the task Harris had in mind). If we look at cases where the only reliable cues of word segmentation are transitional probabilities, and word segmentation is better than chance-level (Saffran et al., 1996b), we have to accept that listeners somehow keep count of this part of performance. If that is so, arguments based on probability of occurrence are no longer in peril of getting outside the domain of linguistics proper.

Chapter 3 is a detailed walkthrough of modelling dialects with the purpose of discerning variant transitional probabilities in them. The aim of the chapter is to make the reader familiar with my method of building dialect models based on studies and corpora, and to make this method transparent and repeatable.

After the theory is established, it is tested in the case studies which follow in chapters 4-7. As stated before, the case studies provide phenomena which are certainly regarded as salient by the language community, and this salience can be aligned with their low probability of occurrence in a cross-dialectal setting. Essentially, the two factors are established independently. It is first proved that the pattern in question behaves like a marker, and then, that it has a low probability of occurrence in comparison with a norm dialect. The nature of the norm dialect is carefully discussed in all of the cases, as this answers the central question of salience: to whom is a dialectal feature salient – the speakers themselves, speakers of the standard, or both of them? The usual case is that to both of them, as dialect speakers are also aware of the standard, at least to an extent. The case studies are drawn from the literature on English variation. English dialects are somewhat overstudied, to say the least, but, in return, provide a wide pool of both data and theory based on said data which can be used as a context for these investigations.

In Chapter 8, a general discussion follows the case studies, where, once again, I draw conclusions, and consider the merits and faults of the theory. The merits are that, ideally, it provides an empirically falsifiable structural
ground for salience, one that it hitherto lacked. Furthermore, it might provide a new insight on the forces behind phonological patterning. The faults are mostly that it is limited in certain ways, and not universally applicable to all known dialectal variables. The final chapter provides brief conclusions.

Chapter 8 also puts the theory of salience into a broader context, and looks at its implications for a theory of language change. Chapter 9, while providing the general conclusions, also returns to an issue left in Chapter 2. Not mistakenly identifying salience with something else, like markedness, is a good start, but we have to take heed not to draw inviting parallels between salience and any phonological or social aspect either.

For example, one could argue that the phonological patterns that typically become salient are the categorical ones, where the distinction is coded at some higher level, such as front/back or lax-tense. One could also find evidence that the typically salient patterns are the low level, gradient ones, including distinction such as vowel height or consonant VOT. Apparently, salience cannot be easily reduced to be a concomitant of any broad phonological category.

1.2.3 The case studies

Definite article reduction

Definite article reduction is a dialectal phenomenon attested in the North of England, predominantly in Yorkshire and Lancashire. It is a par excellence salient feature, as it has been one of the strongest stereotypes associated with Northern speech since the Wuthering Heights (Jones, 2007; Lodge, 2010). This might be surprising given the fact that the reduced article is relatively rare. Most surveys on the subject put reduced realisations between 10 and 14 per cent of all articles. Nonetheless, it is a viable pattern, which is further corroborated by the increase in its use in the city of York (Tagliamonte & Roeder, 2009). It seems that younger York males use the reduced variant more extensively in order to mark their social and regional identity.

The basic pattern of the reduction is that the definite article can be
realised as a glottal stop before consonants and vowels, or, to a lesser extent, a voiceless dental fricative before vowels (table 4.1). This is a simplified description, but, to a large extent, accurate. It is also important to note that the reduced variant, despite the name, cannot be directly linked to the standard article. Neither its history nor its phonetic properties suggest that it is a phonetically reduced variant. Rather, it is a full-fledged allomorph in itself.

| the day [ʔdeɪ] | the inn [ʔɪn] |
| the pub [ʔpʌb] | the apple [ʔæpl] |
| the cooker [ʔkʊkə] | the order [ʔɔdə] |

Table 1.1: Reduced definite articles (examples from Fred)

What is curious about this dialectal pattern is that its frequency does not provide a sufficient explanation for its salience. It is true that definite articles are among the most frequent words in English, so even ten per cent of these articles is quite a number. From a phonological perspective, however, this number is less convincing. Since phonological features are more robustly present in the speech signal, the reduced article is in no ways outlandishly frequent in this context. The phonetic properties of the ‘classic’ reduced article are no cause for concern, either. The glottal stop, the overwhelmingly common realisation of this dialectal feature, is a common feature of dialects in England, resulting from the variable reduction of the fortis stops, typically [t], in word-final and pre-consonantal environments.

My theory of salience, however, explicitly states that the interesting part is not the what, but the where. The glottal stop realisations of the reduced article can show up in environments where one would not commonly expect a glottal stop otherwise, such as between vowels, or before a stressed vowel. While it is possible for glottal stops to occur in these positions even without definite article reduction, it is not usual, and this is precisely what makes these articles salient.
As the case study shows, the transitional probability of definite articles realised as glottal stops is always smaller than the same probability of ‘regular’ glottal stops in a dialect lacking this feature. This means that both speakers of the standard (indeed, any standard, as long as this feature is absent from it) and dialectal speakers aware of a standard will regard this feature as extremely salient.

This feature is very salient, goes the argument, no wonder it is used as a marker, for instance in the city of York. The young York males can rely on definite article reduction to support their Northern identity, since it is quite visible in the sense that it can be picked up easily. Since we have no other reasons for the feature’s salience, the probability-based explanation sounds quite plausible.

Southern English /t/ glottalisation Glottalisation, or the realisation of /t/ as a glottal stop, is a classic feature of Southern English. It was first noticed in London in the early twentieth century. Daniel Jones himself first mentions the debuccalisation of /t/ in a footnote in 1923 (Jones & Trofimov, 1923), postponing a sizeable discussion to 1932. Jones and his circle gradually recognised /t/ debuccalisation first before a syllabic nasal, including environments before liquids and glides later. During the course of the century, glottalisation gained more hold, appearing before all consonants word-medially, before consonants word-finally, and later before vowels in the same position. The present patterns include variable glottalisation in all positions, except word-initially and before stressed vowels. (For more details on the spreading of the pattern, cf. Hudson & Holloway 1977; Williams & Kerswill 1999.)

One survey exploring the distributions of /t/ glottalisation in detail is Altendorf (2003), who distinguishes three social strata (working class, middle class, upper-middle class), two registers (written and spoken) and six environments, (word-medially before a consonant, word-finally before a consonant, utterance-finally, word-medially before a vowel, before a syllabic
consonant, and word-medially before an unstressed vowel).

![Figure 1.1: Lo-Co-Ca: Glottal replacement of t in London (interview style) – from Altendorf (2003)](image)

Her main findings are that in the spoken register, and especially in working class use, [ʔ] almost categorically replaced [t]. The middle class and upper middle class use, however, shows that it is quite strongly avoided in formal registers and in certain environments, most notably before a vowel – both word-finally and word-medially (figure 1.1).

The avoidance of /t/ glottalisation before a vowel can be paired up with an attitude test performed by Fabricius (2000). She found that her upper-middle class speakers judge word-final pre-consonantal and pre-vocalic glottalisation significantly differently: roughly, while word-final pre-consonantal glottalisation (as in ‘get by’ [gɛʔbɑ̃]) is fairly okay, word-final pre-vocalic glottalisation (as in ‘get away’ [gɛʔəwei]) should be avoided. This attitude study, along with the avoidance of /t/ glottalisation in precisely this environment, strongly suggest that pre-vocalic /t/ glottalisation is a low prestige marker. (This
is corroborated by Rosewarne (1984), who explicitly mentions word-medial pre-vocalic /t/ glottalisation, the ‘letter’/‘butter’ type, as a social stereotype.)

Pronounced listener attitudes are not the sole proofs of salience. To give a more loosely related example, Foulkes & Docherty (2006) show that in their Newcastle data on child directed speech, mothers use more glottal stops when talking to their sons than when talking to their daughters. Men use more glottal stops than women in Newcastle, which is in line with the general observation that women are usually closer to the norm (Labov, 1994). The fact that this is projected to child directed speech strongly suggests that glottalisation is a salient feature in this dialect, as mothers – implicitly – try to meet the standard that women should speak more ‘properly’ to their female children.

Again, if we look at the distribution of the glottal stops, we can find good reasons for the salience of the pre-vocalic pattern. In an ‘untainted’ upper class dialect, which allows no glottalisation before a vowel, any glottal stop in this environment is highly unexpected, as they have a transitional probability which is practically zero. Even though such a dialect probably never existed, the fact remains, that if pre-vocalic glottalisation is relatively new, its transitional probability will be extremely low. It is important to note at this point that this is not the glottal stop’s fault *per se*. There are no pronounced attitudes against glottalisation pre-consonantally, where it is a phonetically natural phenomenon, and where it probably have been occurring for a while now. In sum, this feature developed salience as it spread to a new environment.

The salience of glottal stops that precede vowels as opposed to the ‘commonplace’ quality of glottal stops preceding consonants is another curious case of phonologisation pairing up with social indexation. As long as glottalisation is restricted to positions where it is phonetically motivated, it stays below the threshold of listener sensitivity to sociolinguistic variables. As soon as the pattern is generalised to new environments, where its appearance is phonetically ungrounded, it becomes perceptible and prone to incite listener
Preliminaries

judgements. The reason, however, is not phonologisation itself, rather, the phonetic sequences created as its result.

Hungarian hiatus resolution The aim of the only non-English example is to illustrate the general applicability of the theory of salience proposed in this work. Hungarian hiatus resolution is another example of a partly phonologised pattern. Alongside lexical hiatus resolution we also find post-lexical, phonetic hiatus filling with the high/mid unrounded vowels [i] and [e]. In innovative language variants, this pattern is extended to the open mid unrounded vowel [ɛ]. Hiatus resolution – always involving the glide [j] – is obligatory and unnoticed with [i]. It is variable to a certain extent with [e], but this is not subject to any linguistic awareness either. It is, however, salient, and strongly stigmatised in vowel clusters involving [ɛ] (but not [i] or [e]).

The chapter consists of two major parts. The first part is the discussion of a perception experiment aimed at establishing the salience of hiatus resolution with [ɛ] as opposed to the other patterns. Results show that listeners are more judgemental when it comes the absence or presence of hiatus-filling [j] in vowel clusters involving [ɛ], which clear effect is notably absent with the other front unrounded vowels. The second part of the chapter is a corpus study which models the ‘conservative’ standard dialect without the salient hiatus resolution pattern in order to show that the difference in salience is supported by a relative frequency difference in the Hungarian case as well. That is, while clusters of [ɛ] plus [j] are allowed in Hungarian, they are vastly outnumbered by sequences of [i] plus [j], which explains while the occurrence of the former is noted by listeners, while the occurrence of the latter is not.

Rhoticity in Scotland The distribution of /r/ is one of the most widely researched sociolinguistic features. Labov (1966/2006), in a groundbreaking study, reports on the social stratification of rhoticity in New York City. His general results are that the absence of [r] in coda position is generally associated with lower social prestige and informal registers.
Labov argues that rhoticity is a marker of New York City speech, since it shows style-shifting and hypercorrection. This would not be the case if New Yorkers were not aware of this difference, even unconsciously. The marker status of rhoticity is further supported by Becker (2009), a study conducted on rhoticity in the Lower East Side forty years later. As she notes, ‘There is much evidence that both New Yorkers and non-New Yorkers alike do identify non-rhoticity as a salient feature of NYCE, one that (in combination with other NYCE features or even alone) can index a New York persona’ (Becker, 2009, p644.).

Rhoticity is a salient feature in other English dialects as well. One other example is New Zealand. New Zealand English is a special case, since there we have a large amount of direct evidence on the de-rhoticisation of English (Hay & Sudbury, 2005). Furthermore, attitude studies on listener judgements on different dialects of New Zealand English are also available (Nielsen & Hay, 2006).

The loss of coda /r/ in Glasgow, Scotland, stands in a curious contrast with the situation observed in New York City. Though numerous studies report on the decrease of coda /r/ realisation in the city (and in the South of Scotland in general, e.g. Stuart-Smith et al. 2007; Timmins et al. 2004), it is far from evident that this would entail any social indexation.

It is true that the process in the two language areas is different: in New York, we observe an increase of rhoticity, mainly due to an external norm imposed on the language community. In Scotland, the gradual loss of [r] is probably due to a phonetic reduction process, and this is further complicated by the fact that the English English standard is non-rhotic, while the Scottish Standard is rhotic. If Scottish speakers were to follow the English English standard, we would expect upper-class speech to derhoticise, whereas the loss of coda [r] is more observed in working-class speech.

The state of affairs in Glasgow still mirrors New York City, with the working-class speakers using less coda [r] than the upper-middle class speakers. However, awareness of rhoticity is much less pronounced in Glasgow, which is
all the more surprising given that other variables are clear class markers in the city.

In my study of rhoticity I will argue that it is a phonetically fine-grained process, and the extent of variation prevents speakers from picking up on coda /r/ as a reliable carrier of social indexation. In order to prove this point, I first look at studies on English in Glasgow and other areas in the South of Scotland in detail, and then propose a corpus-based model of rhoticity to illustrate the extent of the variation.

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[R]</th>
<th>[rt]</th>
<th>[Rt]</th>
<th>[Ru]</th>
<th>[rtt]</th>
<th>[V]</th>
<th>[m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCOF</td>
<td>23.10</td>
<td>52.71</td>
<td>21.66</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>2.53</td>
<td>0.00</td>
</tr>
<tr>
<td>MCOM</td>
<td>12.37</td>
<td>64.21</td>
<td>10.37</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>13.04</td>
<td>0.00</td>
</tr>
<tr>
<td>MCYF</td>
<td>13.71</td>
<td>60.48</td>
<td>22.58</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3.23</td>
<td>0.00</td>
</tr>
<tr>
<td>MCYM</td>
<td>14.71</td>
<td>52.21</td>
<td>28.68</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>4.41</td>
<td>0.00</td>
</tr>
<tr>
<td>WCOF</td>
<td>11.00</td>
<td>39.23</td>
<td>26.79</td>
<td>2.39</td>
<td>0.00</td>
<td>0.00</td>
<td>20.57</td>
<td>0.00</td>
</tr>
<tr>
<td>WCOM</td>
<td>18.39</td>
<td>44.44</td>
<td>23.75</td>
<td>0.00</td>
<td>0.00</td>
<td>1.15</td>
<td>11.88</td>
<td>0.38</td>
</tr>
<tr>
<td>WCYF</td>
<td>14.84</td>
<td>4.52</td>
<td>20.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>60.00</td>
<td>0.65</td>
</tr>
<tr>
<td>WCYM</td>
<td>2.17</td>
<td>19.57</td>
<td>28.26</td>
<td>0.00</td>
<td>2.17</td>
<td>0.00</td>
<td>45.65</td>
<td>2.17</td>
</tr>
</tbody>
</table>

Table 1.2: (r) versus social class: spontaneous speech (n=1474)

Table 1.2, pulled from Chapter 7, demonstrates the case in point. It shows the extent of variation in the realisation of coda /r/ (that is, in cases it is not deleted) in Glasgow in a study of Timmins et al. (2004), using the variants they propose to cover (r) variation. What we observe is a wide extent of variation, with eight different categories, strongly dissimilar to the New York City case, where the realisation of /r/ is less variable.

Rhoticity in Scotland, then, provides an example of a variable not being salient due to its phonetic properties, and also shows the potential risks of the probability-based method of assessing variable salience.
1.3 Concluding remarks

This chapter gave the outline of the research programme developed in this work, which is the operationalisation of salience as a narrowly definable sociolinguistic concept. Salience was interpreted as a property in the context of cross-dialectal variation, with a clear focus on sound patterns. It is a property of phonological variation related to the way in which phonological variation conveys social information, that is, differences in the background of the speaker, not in the actual linguistic information the speaker tries to convey. The chapter set up the segmental phonological approach that is the basis of the rest of the work, acknowledging its limitations but highlighting its potential. It briefly discussed the methodology of the research programme and the course its development takes in the rest of this work, looking at individual cases of variation in English using segmental, corpus-based methods.

As we will see, salience is not a novel concept, it has been used extensively even inside sociolinguistics. The aim of this work is not to claim exclusive rights to its use as a descriptive term, rather, offers a way in which it can be operationalised as a concrete notion the presence or absence of which is measurable and testable.
Chapter 2

Defining Salience

The first chapter outlined the research programme explored in this work, which is the operationalisation of sociolinguistic salience. This chapter focuses on the history of the term both in and outside sociolinguistics. First, in section 2.1.1, I have a look at previous sociolinguistic studies of salience. Though the term is widely used, there is a fair amount of confusion with respect to its exact meaning: it is established that salience singles out certain dialectal variables to carry social indexation. It is uncertain, however, whether this difference is simply typological or if there is an external grounding to it. In the first case, the term is not much more than a useful heuristic for sociolinguists, and certainly does not deserve an entire book of its own (this was briefly touched upon in the previous chapter). I will side with the authors who believe that there is an external grounding to salience as a property. These authors usually look for it in the cognitive domain.

Therefore, in section 2.1.2, I go on to survey the term’s use in visual cognition, not only because this is the cognitive field where salience is best established, but also because visual and linguistic salience work similarly. They assign a property to a visual/linguistic unit that renders it perceptually more prominent in an array of competing units, which is crucial in cases where selective attention is useful or necessary. Section 2.1.3 has a cursory look at selective attention in hearing. In section 2.2, I attempt to give a solid
definition of salience in sociolinguistics. Since my operationalisation of salience heavily relies on theories of categorisation and human learning capacities, the respective theoretical frameworks – most notably the exemplar theory of categorisation – and the relevant psycholinguistic evidence is surveyed in section 2.2.3.

I will argue that sociolinguistic salience is based on the information-theoretic notion of surprisal: a unit is salient if it shows up with a high degree of unexpectedness, that is, a low probability of transition, in the speech signal. This definition is in tune with the ways the concept is defined and used in the cognitive sciences.

2.1 Salience as a general term

Linguists often talk about salient properties or entities. The usual interpretation of the term in phonology is quite down to earth: a salient feature is more visible or noticeable both for the linguist and the language user. In this sense, ‘salient’ is not more than a synonym for ‘conspicuous’ or ‘standing out’. One can find examples for such interpretations of this term in Vihman’s work on the acquisition of phonology. Vihman & Croft (2007) talk about ‘salient’ segmental positions and rhythmic patterns, which are acquired earlier by the child learner. These patterns are structurally more regular, occur consistently, and are easily noticeable, so the learner can procure them faster. Salience is also a tangible property of phonetic segments: a palatal glide is more salient between two low vowels than between two high vowels, since the formant transitions cueing the glide are much more visible between low vowels, these having markedly different spectra. In comparison, such an intrusive glide between two similarly high vowels will have no strong formant cues, and differs only in the lack of syllabicity. So, we might say that it is less salient.

Salience has a similar sense in morphology. Chapman (1994, 1995) establishes a distinction between morphological categories based on ‘perceptual salience’. According to Chapman, perceptually salient patterns are more
likely to serve as bases of generalisations – they are, in a way, more productive. Her example is analogical vowel lengthening in Swiss German. The core of her argument is that vowel lengthening regularly occurs with suffixation patterns that are salient, i.e. semantically and formally transparent, and that this preference cuts across the distinction between inflectional and derivational suffixes. That is to say, vowel lengthening will tend to occur with a consistent pattern, which pairs a single form with a single function. This transparency is more important than any other property of the suffix, such as whether it is inflectional or derivational.

In this interpretation, a morphological pattern is more salient if it is less marked from the point of view of complexity. The preference for less marked structures is also the central tenet of Natural Morphology (Dressler, 1999). Natural Morphology approaches morphological processes from an information-theoretical point of view: a morphological structure is optimal or maximally natural if it is iconic, uniform, and transparent. Here, salience is interpreted within the frame of morphology: the more regular the pattern, the more it adheres to the structure, the more salient it will be. This is slightly at odds with the phonetic interpretation of salience, where the maximally unlikely or out-of-place pattern gets more attention.

The common ground of the two approaches is that they do not regard salience as a special, inherent property. In phonetics, it solely refers to the physical properties a segment displays in a certain environment. In morphology, it more or less comes down to regularity. That is, it is a characteristic arising from the unit’s placement in a larger structure.

In the common linguistic usage, ‘salience’ is either to be interpreted as a simple segmental property, as in phonetics, or a descriptive term to cover a set of phenomena, as in the morphology example. It can also be associated with a general attribute that follows from the model architecture. The present example for this is Natural Morphology, where salience is associated with a more or less clearly defined notion of markedness. As noted in Chapter 1, such an approach can cause its own problems its various ways.
Defining Salience

What unites all these interpretations, at least from our point of view, is that salience is seen as a ‘hollow’ notion, either involving a straightforward stand-alone interpretation, or gaining interpretation through its connection with the rest of the model structure.

The difference in sociolinguistics, as we will see, is that the term ‘salience’ aspires for a status as a self-propelled theoretical concept. The problem is that often this theoretical status is for starters taken for granted, without any attempt to give the term an explicit definition. There exist a number of definitions of salience in the surveyed literature, but none of them covers all the phenomena that are regarded as ‘salient’ by other authors.

2.1.1 Salience in sociolinguistics

The starting point in discussions of sociolinguistic salience is the two possible behaviours typically associated with a dialectal variable. Certain variables might become targets of various processes, such as analogical levelling, dialectal borrowing, stigmatisation, style shifting, and so on. It is hard to establish, however, why some variables are prone to be targeted by such processes whilst others stay exempt. The distinction between the two types of variables is stated by Labov (1972b), who establishes two main categories to cover it.

*Indicators* vary with social stratification, but have no social interpretation. If we have a standard and a non-standard dialect, an indicator variable will be different in the two. Yet, non-standard speakers will not try to use the standard indicator variable when speaking the standard dialect, and this will not be noticed by the standard speakers. Indicators do not show style-shifting, and their use by speakers does not invoke value judgements from the members of the language community. They are not subjects of naïve linguistic awareness either.

An example is /a:/ in Norwich, mentioned in the previous chapter. Another could be the use of the velar approximant [uː] in words like *what, where* in Northern versus Southern American English – while some Southern dialects have it, it does not seem to be associated with Southern speech in general.
Salience as a general term

(Dennis Preston, p.c.). One problem with indicators is their sometimes nebulous status with respect to speaker behaviour. Ideally, an indicator does not show style shifting – it is not used less in a formal setting or when reading word lists. It does not trigger negative listener attitudes either: people will not consider its use non-standard, even if it is actually typical of a non-standard dialect. There are, however, features, which are avoided in more formal settings, but have no social stigma attached to them: listeners will not regard using them as less formal or proper, not even implicitly. Another, more serious problem is that it is sometimes hard to tell why a certain feature is an indicator rather than a marker.

Markers correlate with a sociolinguistic identity. If a marker attaches to a non-standard dialect, speakers will try to avoid it in more formal style settings and will regard its use as ‘base’ or ‘erroneous’. I already cited Northern vs Southern English English /a/ as an example. Another classic one is rhoticity in New York City, a marker heavily associated with social class and formal style settings Labov (1972b). Markers do not need to be explicit subjects of linguistic awareness. One can regard a speech variety as improper without being aware of the marker which provoked this attitude – yet, when you remove the marker, the attitude disappears along with it.

There is one case where we can be utterly certain that a variable is used as a marker. This is when it becomes so strongly associated with a particular dialect or sociolect that it turns into a subject of naïve linguistic awareness. The only rough criterion for a linguistic stereotype is overt commentary. As Rosewarne (1984) shows, Southern English English speakers are able to point at word-medial glottalisation (in words like ‘butter’), and often attribute it to slackness or carelessness, ‘having a lazy tongue’.

Preston (1996) drafts a typology of folk linguistic awareness, indicating well its complexity. Language users can be wholly able to remark on specific features, while they only become aware of others if these are identified for them. Some differences are not available for conscious discussion at all. The accuracy and detail of naïve awareness can also vary. (This has been
illustrated by the African American Vernacular realisation of /ar/ in Chapter 1.)

Labov’s distinctions, to an extent, match Silverstein’s (2003) notion of indexicality. Indexicality is put to use in the analysis of the local dialect of Pittsburgh by Johnstone et al. (2006). Linguistic units of the first, second, and third order of indexicality roughly correspond to indicators, markers, and stereotypes. These concepts are more abstract, however, due to Silverstein’s emphasis on how linguistic forms acquire social meaning.

This summary gives the unfortunate impression that a marker is something that ‘switches on’ a particular dialect image in a binary, deterministic fashion. To give an example, extensive word-medial glottalisation in ‘butter’-type words will cause the listener to categorise the speaker as ‘naff’: lower-middle class or working-class, uneducated, and socially immobile. We could conveniently skip this level of abstraction and say that word-medial glottalisation in English is naff in itself, since it always reliably triggers the same kind of social profiling. The problem with this, of course, is that the interpretation of a marker can change as a function of time and space: it can carry different indices dependent on the region, the social class of the listener, and whether it is the start- or the end-phase of an ongoing sound change (more on this in Chapter 8).

If we accept that interpretation is relative, we could still argue that recognition is not. That is to say, in overall similar dialects, the same variants are prone to become markers and the presence of a type of variation will inevitably carry social indexation in any other similar dialect.

Even this is not that simple. In an excellent paper, Campbell-Kibler (2011) shows how the interpretation of markers is dynamic. She looks at gay speech stereotypes in US English and notices that some variants do invoke particular stereotype personas, but only in conjunction with certain other variants. If this conjunction is not present, a variant might invoke a completely different persona or remain unnoticed.

This is corroborated by a study of Copenhagen Danish /s/, by Kristiansen
et al. (2011). They investigate the way differences in /s/ realisation are perceived by Danish listeners. They find that the way fronted variants are interpreted strongly depends on which area of the city the speaker is assumed to come from, which, in turn, is connected to social class – a fronted [s] can be associated with certain types of rich upper-middle class Danish speakers living in one part of the town, or immigrants living in another part. As one of their participants notes, ‘if this person comes from [rich part of the town], he must be gay, if he comes from [immigrant area], he is a foreigner.’

The present discussion necessarily simplifies this issue and talks about the perception of individual markers. This is not a fundamentally flawed approach – some sociolinguistic variables, like (ing) realisation, are reliable markers in themselves, and, in any case, one can always elaborate on the model proposed for singleton sets of variation.

What seems to be certain is that, one way or another, sociolinguistic variables can be grouped along the indicator/marker dimension, that is, whether they carry social indexation, or not. The question remains how a variable is selected to be a marker, and what the exact differences are between the two variable types.

Trudgill (1986), taking Labov’s dichotomy as a starting point, lists a number of properties that linguistic markers have (as opposed to indicators). His context is that of dialect contact situations. These properties include:

1. having an overtly stigmatised variant,
2. having a high-prestige variant (indicated in the orthography),
3. undergoing change,
4. comprising a large phonetic difference,
5. being contrastive.

He adds salience as an additional factor, with an external psycholinguistic grounding, lying in perceptual distinctiveness and speaker attitudes. A
Defining Salience

marker might have *extra strong salience* (sic.) when the language users regard it as a particularly strong feature of the target dialect. In this view, then, salience is neither the source of the distinction between indicators and markers, nor between different types of linguistic variables in general, but rather, it is an additional property with an external basis. Pointing to perceptual distinctiveness and speaker behaviour implies cognitive roots: salience, then, is a trait a variable acquires based on its perception by individual speakers.

Kerswill & Williams (2002) examine Trudgill’s features and point to two general problems: the first is that while these features might serve as guidelines in establishing the marker status of a dialectal variable, they are not particularly helpful in deciding whether the variable, as a marker, will be accepted or rejected by language users. Their second objection is that bringing forward stigmatisation as a source of marker status – which in turn leads to stigmatisation – is circular. Kerswill and Williams add that the only property they find unproblematic is the notion of contrast. This property implies, however, that the variation of a marker has to affect a phonological contrast. Nonetheless, language users can be extremely conscious of dialectal variables which do not neutralise contrasts, as in the case of /t/ glottalisation in the South of England (Fabricius, 2000). This suggests that the correlation between contrastiveness and marker status is not that clear-cut – a conclusion we ought to bear in mind despite the fact that the concept of phonetic distance, to an extent, covers contrastiveness as well.

As noted above, salience, in Trudgill’s interpretation, comes on top of all this, contributing to the marker status of a sociolinguistic variable. If we accept that, we have to look for the origins of salience in the basic phonetic/structural properties, while it also remains rather imprecise how salience promotes a variable. While we could also take a view of salience as a sum of these properties (and/or others), i.e. those that differentiate a marker from an indicator, this approach is problematic in its own particular way. It bereaves salience of any theoretical status, and reduces it to an alternative for ‘marker’. What is more, the inherent circularity remains, as the properties
of a salient variable, in this view, are precisely the ones that follow from its salience.

Kerswill and Williams, partly following Trudgill, give a provisional definition of salience as *perceptual and cognitive prominence*. This is a point to which I will later return. They go on to illustrate that the term is far from being unequivocally established. Kerswill (1985) argues that the main prerequisite is phonetic discreteness, particularly if it comes from the lexicalisation of a phonetic difference, his example being /l/ vocalisation in the South of England.

Auer et al. (1998) also start with the assumption that lexicalisation has a prominent role in the salience of a dialectal feature. In their study of East German immigrants’ adaptation of the local standard in West Germany, however, they find no strong correlation between the lexicalisation of a feature and its adoption by language users. Their starting criteria are similar to Trudgill’s, but are in fact taken from Schirmunski’s 1930 pioneering work on dialect contact situations in German speaking dialect enclaves in Russia. Schirmunski established a main difference between primary (or salient) and secondary variables in dialect contact, resembling Labov’s marker/indicator distinction. Schirmunski assumed six criteria that a salient dialectal variable will typically meet:

1. **Articulatory/perceptual distance**: the physical distance between two differing dialectal variables

2. **Lexicalisation**: Whether the change affecting the dialectal variable proceeds in a gradual or a categorical manner (see above)

3. **Categoriality**: Whether the variable constitutes a categorical or continuous phonetic difference

4. **Awareness**: The extent of naïve linguistic awareness of the variable

5. **Writing**: The variable’s occurrence in writing, informal or formal
6. Comprehension: Whether a variable is intelligible between two dialects

Auer and his colleagues note that while the first three of these criteria are objective, the latter three are subjective, and depend largely on the speaker or the speaker community. Applied to their study of East German immigrants in the West, they find that lexicalisation can override all these criteria, and that, most notably, the subjective criteria do not prove to have large predictive powers at all. This leads them to question the general usefulness of subjective salience, and, indeed, salience in general.

Kerswill & Williams (2002) note that understanding salience requires the consideration of a wide array of possible properties (including lexicalisation). They dismiss attempts to write it off as a positive correlate of high frequency (cf. Bardovi-Harlig 1987). Defining salience in such a way is also unfortunate as it goes directly against the most general interpretation, discussed at the beginning of this chapter: a phonetic feature which is salient in its environment will remain so irrespective of its frequency in a given chunk of speech.

Another strong argument against blaming salience on high frequency is delivered by the study of Labov et al. (2006), who show that if listeners associate a dialect marker with low prestige, this will in turn take an effect on their judgement of speech input even when the marker is relatively rare in the input. Their case is the variable (ing), the varying velar/coronal pronunciation of the progressive suffix <-ing> (-[an])/-[an]) in US English. In their study, the coronal variant of (ing) has ten possible repetitions in an attitude test, and the results showed that listeners pick up on it after around three repetitions, and their attitudes do not change remarkably after that, irrespective of the number of further repetitions. The presence of such a threshold seems to contradict a simple matching of salience and high frequency. (Paul Foulkes [p.c.] points out that the presence of this threshold might be an artefact of the experiment design, because the third repetition is notably different from the previous ones. In such a case, the test subjects could note the presence of the variable not because of its frequency but rather due to the differences corresponding with it.)
This is by no means an exhaustive list of the properties associated with salience. For instance, Kerswill (2002), in the context of dialect contact, adds regional restrictedness to phonetic distinctiveness as a requirement.

Going back to Kerswill & Williams (2002), their conclusion is that salience is a useful term in sociolinguistics, despite the apparent haze surrounding its definition. They establish three components that any operationalisation of salience has to incorporate:

‘The presence of a linguistic phenomenon whose explanation we suspect may be due to the salience of the linguistic feature or features involved. Typically, the phenomenon will be a particular pattern observed in language change, language variation, the variable behaviour of individual speakers, or the acquisition of a linguistic feature. In cases of language change and variation, the linguistic features are items being transferred from one language variety to another through diffusion; however, diffusion-type mechanisms may hold for the other types of phenomena as well.

Language-internal explanations, such as the presence of phonological contrast, great phonetic distance, internally-defined naturalness, semantic transparency, or a particular syntactic or prosodic environment.

Language-external cognitive, pragmatic, interactional, social psychological, and sociodemographic factors. Some have a natural link with the linguistic features being adopted (e.g. that between a syntactic feature and its pragmatic function), while others have an arbitrary relationship (e.g. the favouring of one vowel quality over another).’

(Kerswill & Williams, 2002, p.91)

They stress that the language-external factors are essential in giving a
Defining Salience

circularity-free explanation to salience. They make several references to a link to a possible cognitive background; this further reinforces their earlier, provisional definition of salience as cognitive or perceptual prominence.

To sum things up, sociolinguistic salience, according to the literature, can be interpreted in two ways. First, it can be a tag applied to variables that carry social indexation, and subsequently behave differently in code switching and language contact situations. In this view, salience is no more than another term for the indicator/marker distinction.

The other possibility is to find an external basis for salience. In this case, the term is derived from properties outside sociolinguistics proper, irrespective of whether salience selects only markers, a subset thereof, or an entire cross-cut of variables, which is independent of the indicator/marker distinction. If we try to find an external source, it can either come from speaker dynamics, that is, the social setting in which language is used, or the basic cognitive capacities of the individual speaker.

Opting for the first possibility, speaker dynamics, means that we give up on finding a universal definition of salience. That is, any linguistic variable could theoretically be chosen by the community to mark social indexation, independent of the variable’s properties. This is the view embraced by Labov (1972b): he argues that all variables start out as indicators, and later become markers, when the linguistic change gains enough momentum to be noticed by the community and, as a result, they become vessels of social indexation. This view of the social life of variables can be inferred from his Martha’s Vineyard study. According to the study, the local residents at Martha’s Vineyard picked up on a shift of the realisations of the diphthongs /aw/ and /ay/ to separate themselves from the summer residents. The small difference between the local and the New England dialect became amplified to mark social identity. At the beginning of this phase, the diphthongs are only indicators of this difference, later, as they start to be used for the assertion of the local identity, they become markers.

Labov (1994) discards this simple explanation on the relationship of
Salience as a general term

indicators and markers, pointing to the fact that some variables never seem to become markers at all. This brings us back to the original question: if salience has an external basis, and it is not only social dynamics, one ought to find a general, perceptual ground for it, a frame, that prefers some features to others. Both Trudgill (1986) and Kerswill & Williams (2002) point to general cognitive capacities as a source of sociolinguistic salience. The supposition is reasonable, as some variables might be picked up because they are more highlighted in the course of acquisition or auditory perception.

If we accept that sociolinguistic salience has cognitive grounds, we might want to turn to cognitive science to look at the term’s history there. Such a comparison between sociolinguistic and cognitive salience is helpful in two ways: a look at cognitive salience can render defining sociolinguistic salience easier, while comparing the two fields can tell us more about the general properties of salience in perception. The first field to be covered is visual cognition, with a brief look at auditory cognition as well.

2.1.2 Salience in visual cognition

Salience, as a concept, is best established in visual cognition. This section discusses the interpretation and use of the term in this domain. The next section tries to operationalise this notion in sociolinguistic research relying on the cognitive background. Besides seeing, I also talk about listening in section 2.1.3, though I will contend that visual cognition is a much more useful parallel, despite the fact that our topic is phonology, an undoubtedly auditory field. (The gesture patterns of sign languages also belong to the phonological domain, but sign languages remain unaddressed in this book.) The overview is not a complete one. The contributions of Gestalt psychology and, more importantly, figure-and-ground literature (cf. e.g. Rubin 2001) remain unaddressed.

Wolfe & Horowitz (2004) argue that sensory information is an ‘embarrassment of riches’ that the brain could hardly process all at once. This is even more germane in visual attention deployment, since visual perception
Defining Salience

is responsible for most of the sensory information we obtain. The role of visual attention, then, is to filter out the relevant information and ease the processing burden. This is where salience becomes relevant: a salient object bulges out of its surroundings, and, consequently, it can become a focus of attention much easier.

Salience in visual cognition is usually interpreted as the property of objects in our visual field. Its relevance as such comes from the fact that the brain is incapable of processing the visual input in its entirety, and, therefore, is forced to select. Attention deployment, the preference of certain parts of the visual input to others, can be regarded as a spotlight, focusing on different parts of the input scene (Treisman & Gelade, 1980; Crick, 1984).

There is no complete agreement on the spotlight view of attention deployment (cf. Desimone & Duncan 1995). What seems certain is the existence of an informational bottleneck: since visual input cannot be handled as a whole, most of the information is filtered out initially, and this filtering is guided by attention. Attention gating, the selection and filtering of ‘interesting’ bits, has two broad aspects. First of all, it takes place in two steps: there is a quick, automatic sweep controlled by target detection, concentrating on the highly salient items in the input. This is followed by a slower, effortful, controlled sweep. The latency of the latter, but not the former, depends on practice as well as task difficulty (Weichselgartner & Sperling, 1987). The two separate processes of attention gating are in a strict correspondence with the two main sets of influences, the bottom-up and the top-down ones.

Bottom-up effects are largely automatic and depend on the stimuli. Salience, computed across the whole visual field, plays a weighty role here. Top-down attention deployment is volitional, task-driven, and is affected by aspects such as reward or motivation (Itti et al., 2002). In comparison with bottom-up effects, top-down ones are relatively complex. For instance, Itti & Baldi (2009) argue that surprisal has a huge part in the latter, though probably not the former. Their interpretation of surprisal, however, is not simply a mechanical one. They assume an object to be most surprising if it
causes the largest change in the observers prior beliefs about the world, which is something far more intricate than the surprisal of a red object projecting in a black field. In this connotation, the salience of a given object is a sum of possible motivations, the given task, the observer’s mental state, and so on.

In comparison, bottom-up effects are relatively straightforward. Wolfe & Horowitz (2004) give a list of attributes that might render an object salient in a visual field. Their list shows that these are very basic visual properties, such as luminescence or shape. They strongly debate whether more complex objects, even if they are frequently occurring, such as the human face, would be able to have an effect on bottom-up attention gating. Furthermore, this type of attention is overtly sensitive to the relative complexity of the visual field. The search for a salient object is easy if it differs from the rest in only one feature.

It is not very difficult to find the black pawn between the white ones in figure 2.1, since it only differs from the rest in colour.

The same is true for table 2.2. Here the difference is in a different feature, the shape, not the colour.

Figure 2.3 is another matter. Here we have a difference of two independent
features, colour and shape, so it might take a while to find the black queen. These three figures not only illustrate the role of feature complexity in visual search, but also that salience is a property calculated against the rest of the visual scene: the black pawn is not salient *per se*, but only in company of the white ones. As Itti et al. (2002) note, feature contrast is more important than
local feature strength. Parkhurst et al. (2002), in their bottom-up model of stimulus salience, go on to argue that the observer will prefer regions which are deemed to be more informative in the sense of Shannon & Weaver (1948).

The Shannonian notion of surprisal entails that informativeness is determined by the degree of uncertainty: the more unpredictable a point is, in comparison with its environment, the more salient it will be. To go back to our simple example, in Figure 2.1, a black pawn’s colour will be unpredictable based on its environment, whereas this will not be true for a white pawn. Consequently, a black pawn will be more interesting for the observer, and it will be the first target of attention.

This is an overarching property of salience, not only in visual cognition, but also in linguistics. Salience is a relative concept: it is always interpreted in a context.

Models of attention gating assume a salience map (Itti et al., 2002; Koch & Ullman, 1985), which is used to locate points of interest in the visual field. To put it very simply, a salience map is calculated in the following way: an individual salience map is calculated for each of the relevant classificatory features (in our examples, these would be colour and shape), and then these are matched in order to find the most salient point of the visual field.

There is no complete agreement on the top-down and bottom-up share of the workload. What seems to be certain is that some properties of objects are salient in the sense that they will jut out irrespective of the observers prior expectations or the search task. Itti et al. (2002) emphasise the importance of these properties: they argue that the research on salience in visual cognition can be extremely relevant to fields of application such as advertising. This is an interesting point, since the use of salience in advertisements would mean the manufacturing of visual objects that divert attention. This is slightly closer to our general topic than discussions that limit themselves on the location of salient objects in a visual field.

In advertising, one would want – for instance – to design and place a banner that sticks out and grasps the passer-by’s attention. Similarly, a
Defining Salience

language user needs linguistic units that they can use to express their social identity. These units also have to be salient, so that the other language users notice them easily and are able to pick up on the social indexation they carry.

To sum it up, salience in visual cognition is a property of individual objects in a visual field. This property, however, is calculated on the basis of comparison with other objects. It ultimately derives from informativeness, degree of unexpectedness, or surprisal: the less predictable an object is, in comparison with its environment, the more salient it will be. Salience plays a prime role in bottom-up attention deployment, which is automatic, swift, and independent of the search task.

2.1.3 Selective attention in hearing

Selective attention is also an issue in hearing. Auditory prominence belongs to a different level than sociolinguistic salience, and the two seem orthogonal. One example that is nonetheless relevant to this study is the ‘cocktail party effect’ (Pollack & Pickett, 1957). The term covers two related phenomena in auditory attention deployment: the ability to filter out a single speaker in the noise coming from the chorus of other speakers and the background, and the capacity to monitor more voices simultaneously and divert one’s attention to one of them if a prominent piece of information comes up in it. This is usually illustrated by our ability to take notice of somebody mentioning our name in a conversation at a cocktail party, even if we were paying attention to a different speaker – hence the name of the effect.

Arons (1992) gives an excellent summary of the phenomenon, from the point of view of speech processing. The task of filtering out background noise from the input of e.g. one person talking to us is trivial for humans, yet it always proposed a serious difficulty for computer-based speech processing. (This point is linked to the problem of speaker normalisation in speech perception. Exemplar-based approaches to speech perception, the advance of which in linguistics happened in the mid to late 1990s, try to tackle the issue directly. This is touched upon in section 2.2.3 below.)
One seminal paper on selective attention is Cherry (1953). Cherry performs experiments where the test subject has to listen to two separate spoken messages. His main observations are that subjects are able to separate two mixed messages – given enough repetitions – if these are distinct enough in vocabulary and content. The subjects’ ability to separate the messages deteriorates if they contain a large amount of ‘highly-probable phrases’ or prefabs, n-grams which occur together often in corpora (Bybee & Cacoullos, 2009). (Labelling these units as prefabs is somewhat anachronistic, since Cherry’s methodology is different from that of Bybee and Cacoullos, for instance. Still, the general notion behind ‘highly-probable phrases’ and prefabs is the same.)

One of Cherry’s main conclusions is that listeners rely heavily on ‘transition-probabilities’ in the subject matter, voice dynamics, and syntax, when separating two mixed messages, that is, deploying selective attention. His approach to attention deployment is information-theoretic (Shannon & Weaver, 1948). He argues that listeners seem to keep track on transitional probabilities between units at various levels of speech and divert attention using these as banners. It is implicit in his argumentation that selective attention relies on the Shannonian notion of surprisal, which is consonant with the reasoning we find in visual cognition. Needless to say, it also supports, to some extent, the theory of salience advocated here. The analogues of sociolinguistic and visual salience, such as the role of surprisal in both, will be elaborated in the next section.

2.2 Operationalising sociolinguistic salience

2.2.1 Preliminaries

At a first glance, the meanings attributed to ‘salience’ in visual cognition and sociolinguistics are completely unrelated. In the former, it is relevant to the selection of an object in a single visual field. This problem does not
Defining Salience

even come forward in the latter, as we usually regard auditory processing as sequential. Even though there are highlighted domains in the auditory input as well, such as the spectral peaks (formants) of the sound spectrum of human voice, sociolinguistics, even sociophonetics, usually takes the first part of speech processing as granted.

The second main distinction follows from this one: in visual cognition, salience is typically interpreted over a given image, whereas in sociolinguistics, we look for entities (linguistic variables) that prove salient in the long run, in the sense that speakers continuously rely on them in social indexation. It is also important to note that while visual salience seems to affect the bottom-up processes of attention gating, so it is a very low-level factor in attention deployment, linguistic salience comes in at a much higher level. It is interpreted as a property of entities that have already been identified as segments of a given dialect spoken in a certain setting (by a specific agent, etc.). This computationally demanding frame of reference is missing in the selection of a red object in front of a black background.

These differences all follow from the sociolinguistic standpoint, which assumes that salience is interpreted as a property of a dialectal variable in language change or a dialect contact situation. If we look at salience as a general characteristic, it is easy to find one main similarity. In both cases, an entity juts out due to its dissimilarity to the rest of the structure. Dissimilarity is not an inherent property. It is assessed in comparison with the rest of the structure: a red object is salient against a black background, but not against a red one. Similarly, a palatal glide is salient between open vowels, but not between closed ones.

Gauging the properties of a unit in relation to the whole structure of which it is a member is a fundamental structuralist tenet (de Saussure 1916/1995, Lévi-Strauss et al. 1955). According to structuralist reasoning, individual properties are not inherent; rather they are only interpretable within a larger system of reference. I will also rely on this in my operationalisation, and establish the following link between the use of salience in visual cognition
and sociolinguistics: in the former, it is a property an entity assumes in relation to the rest of the visual field, whereas in the latter, it is a property a variable assumes in relation to the language structure.

### 2.2.2 Defining salience

Relying on the various descriptions of salience in language and cognition, I am going to define salience in sociolinguistics as follows: *A segment is salient if it has a large surprisal value when compared to an array of language input.* Surprisal here is an information-theoretic measure – the less likely the segment is to occur in a particular sequence, the more surprising it is for the language user. The reason why we can talk about surprisal is that the language input, the speech signal, is highly structured. This follows from two properties of the phonological system, namely, a restricted set of available segments and phonotactic patterns determining the succession of segments. It also follows from the less predictable dynamics of connected speech sequences.

My main unit of measure will be transitional probability (TP). The transitional probability of segments $Y$ and $X$, $p(Y|X)$, is the probability that, given segment $X$, the following one will be $Y$ (Table 3.1). It is calculated by dividing the occurrences of $XY$ with the occurrences of $X$ in a given corpus. Its use in linguistics was first suggested by the eminent American structuralist, Zeilig S. Harris (1955).

$$p(Y|X) = \frac{\text{likelihood of pair } XY}{\text{likelihood of } X}$$

**Figure 2.4:** Probability of $Y$ following $X$

While this formulation implies that a completely unfamiliar segment will have the largest surprisal value (as, based on prior knowledge, it will not be anticipated at all), I restrict my observations to cases where surprisal is a combinatorial measure, that is, where we talk about the surprisal of segments that are regularly used in a language.
Defining Salience

The context of surprisal is always a comparison of two dialects of the same language, usually a standard variety $S$ and a non-standard variety $S'$. The typical scenario looks like this: the distribution of a segment $X$ is different (usually less restricted) in $S'$ than in $S$. What follows is that a speaker of $S$ encounters $X$ in positions in $S'$ where it is much less likely to occur in their own dialect. Consequently, $X$ will be salient to the $S$ speaker, and can carry social indexation. Mind that this correlation is not implied to be symmetrical: while I will argue that, at least in a certain set of cases, salience is the prerequisite of being a social marker, not all salient features necessarily carry social indexation. This is where social dynamics come in: salient variables constitute a pool from which language users can pick to assert social identity, but they need not use all of them.

The other important effect of the use of language in the community is that, through interactions with $S$ speakers, the $S'$ speaker can also pick up the variable $X$ and use it to assert social or geographical identity (cf. Silverstein 1998). As the sociolinguistic literature shows, the two dialects do not need to be in a standard-non-standard relation. What is important, though, is that we relate two dialects of the same language to each other. I have no intention to go into the dialect-language debate here. All I would like to stress here is that a single distinction can only be interesting if the two dialects are similar enough to be in interaction. Since this will mostly be self-evident in the case studies, this was brought up at this point only for the sake of completeness.

While the choice of words here (and, indeed, throughout the book) might imply the language user’s conscious participation, it has to be stressed that very few social variables become stereotypes, subjects of naïve linguistic awareness. Any talk of ‘jutting out’ and ‘picking up’ of variables should be put in the context that these processes usually remain entirely out of any awareness or conscious control.

In order to clarify this interpretation of salience, we have to look at two aspects of segmental distribution in the speech stream. First, we have to see how the structure of the speech stream allows domains to be salient. Second,
we have to find evidence that language users can rely on such information, e.g. information on transitional probabilities. This salience claim relies on speakers making vast generalisations based on the statistical regularities available in the speech stream, which requires the speakers to extort and process these patterns. The following section elaborates on these two aspects.

It is important to stress that this definition is usage-based. It assumes that language users rely on information extracted from the speech input. Even though abstract patterns, such as language-specific phonotactic constraints, have an influence on transitional probabilities (TP-s), they can be only acquired from language in use. This has various implications for my theory which will be addressed in the next section.

2.2.3 Exemplars and transitional probabilities

Harris (1955) was the first to suggest that linguists can use distributional cues of phones to find word boundaries. A linguist, he argued, when encountering an unknown language, can find the boundaries between words relying on the following method: when we take a string, we can count how many phones can possibly follow it. If this number is large, the string is likely to be the end of a word. If it is small, it is likely to be in a word. Languages have a limited set of lexical items, and the members of the segment inventory cannot occur freely in the words themselves, but have to adhere to phonotactic restrictions. Therefore, the probability of any given segment following another one within the word will be relatively low compared to the case of an intervening word boundary, where any segment can follow any other segment, as long as the former one can occur word-finally and the latter one can occur word-initially. (Note that this is not necessarily the case: the transitional probability of [ŋ][h] in English will always be zero, as [h] cannot appear word-finally and [ŋ] word-initially.) Consequently, low transitional probabilities (TP-s) cue a word boundary. This is illustrated by the example of [t] in Table 2.2.3: the word boundary is implied by the wider array of choice.

What Harris noticed is a regularity inherent to the phonological system,
Defining Salience

<table>
<thead>
<tr>
<th>Word-initially</th>
<th>Word-finally</th>
</tr>
</thead>
<tbody>
<tr>
<td>(lots of these)</td>
<td>(few of these)</td>
</tr>
</tbody>
</table>

\[ \begin{align*}
\#t & \{ r \, w \, j \, V \} & \ldots & t\# & \{ p \, t \, k \, t' \, s \, l \, h \, \ldots \} & \ldots 
\end{align*} \]

Figure 2.5: Possible following segment for a voiceless coronal stop in English

the asymmetry of segment combinations in word-medial versus word-final positions. A comparable asymmetry can arise if a segmental pattern has a wider distribution in the phonology of one dialect as compared to another. To give a very simple example, a rhotic dialect of English has a much higher TP for \( C[r] \) (the segment \([r]\) followed by a consonant) than a non-rhotic one. (Theoretically, the TP in the latter case would be zero, but dialects always show an extent of variation, so pre-consonantal \([r]\)-s might occur in a non-rhotic dialect, the same way as a rhotic one can delete pre-vocalic \([r]\)-s.)

If a non-rhotic speaker encounters a rhotic dialect, the occurrence of \([r]\)-s will be contrary to their expectations: the TP for \( C[r] \) will be much higher in this dialect than their native one. This unexpectedness will be a source of salience, and, consequently, rhoticity will be a possible candidate for social indexation. (As expected, rhoticity is a social marker on both sides of the Atlantic, as in New York City (Labov, 1972b) or Bristol (Wells, 1982).) Before we get carried away with the possible applications, one concern must be raised. Harris thought of TP-s as information available for the linguist, who, given a sizeable corpus, can compute it and use it to find the word boundaries.

There is abundant evidence, however, that language users are able to rely on such information (Jusczyk et al., 1994, 1999; Saffran et al., 1996a,b;
Cairns et al., 1997; Pierrehumbert, 2003; Hay, 2000). Psycholinguistic research concentrates especially on infants’ ability to segment speech based on distributional cues. Jusczyk et al. (1994) investigate 6 and 9 month-old infants’ preference of high versus low frequency phonological patterns in the native language, which is, in this case, English. Their method is to play lists of CVC shaped non-words to groups of infants. The hypothesis is that the subjects will pay attention to familiar (that is, more frequent) patterns longer. They find that the older age group shows a significant preference for patterns that occur in the native language with a higher frequency. This remains true even if the input sets are matched for vowel quality. That is, even if the vowels are identical in the non-words with the different patterns, older infants still prefer the high frequency ones. This suggests that frequency is the only cue that the subjects rely on.

Saffran et al. (1996b) perform similar tests on adult learners. Their subjects have to acquire an artificial language consisting of nonce words which were based on English, but are highly simplified. The input is a continuous stream of the nonce words read by a speech synthesiser. The only cues the subjects can use are the TP of nonce syllables, all of which have a CV form. The subjects are then tested using a list of words, some of which were built up from syllables which never occurred next to each other in the input, i.e. their TP was zero, and some which were built up from syllables which regularly occurred next to each other. The results are that performance is significantly higher than chance in the recognition of the ‘words’ of the input. These forms consist of syllables that occurred next to each other in the training input. This means that the subjects could recall the TP-s of the input phase in the test phase.

Cairns et al. (1997) provide an overview of models of speech segmentation. They note that while top-down cues, such as the knowledge of the existent word forms of the language, have a large impact on segmentation, bottom-up cues, like TP-s, play an important role as well. They duly note that there are problems with approaches exclusively relying on TP-s in segmentation. For
instance, even if we ignore top-down effects, which makes sense if the subjects are acquiring the language, there are still other strong factors in play, such as prosody: ninety per cent of English words have initial stress, which provides a reliable cue on word segmentation. Silverman (2006) argues against artificial language experiments (such as the two described above) on the grounds that by excluding all ‘noise’ (the effect of prosody and lexeme recognition, for instance) they create an unrealistic setting for language acquisition, which normally includes all possible cues, including the pragmatics and the setting.

Bearing all this in mind, one can still safely say that TP-s are available for the language user – even if the user does not depend exclusively on them. In order to put this claim in a broader context, and use it in operationalising salience, we need a theory of reference which is able to accommodate information drawn from language use – like TPs – in speech perception. This is not a trivial problem, since the classic view of speech perception has no room for such information. The classic view of speech perception regards it as a de-coding process, which reconstructs the message contained in the input signal, after stripping it from everything except the bare essentials.

This happens in the following way: let us assume that a speaker directs an utterance towards the listener. This utterance is processed as speech (and not as howling of the wind, for instance). Subsequently, it is normalised. This is necessary, since, for example, the formant values of vowels, the main cues of vowel quality, can vary dependent on the speaker’s age and gender – and the corresponding difference in the size of the vocal tract. The listener reconstructs the canonical vowel phones of the input by, for instance, establishing normalised formant values by comparing the formants of the utterance to the general formant value ranges of the speaker (Gerstman, 1968). This is called speaker normalisation. This way the listener recovers a detailed phonetic form from the highly variable input signal. This is then turned into an abstract phonemic representation by a phonological module (Chomsky & Halle, 1968). The phonemic representation is then interpreted by matching it to the entries of words in the listener’s mental lexicon. The
mental lexicon consists of similar abstract representations. In such a model, there is no toll taken of statistical patterns, since the incoming utterances are de-coded sequentially, one by one.

The view that speech perception involves speaker normalisation is challenged by Johnson (1997). Johnson uses an exemplar-based model of speech perception. In such a model, all utterances are stored by the listener in rich phonetic detail. No abstract category prototypes or underlying forms are posited. The stored utterances, the exemplars, are categorised based on auditory properties, as well as external category labels, such as identity, age, or gender of speaker, context, and so on. Categorisation involves comparing the new exemplar with each of the remembered instances of the categories, and the choice of category is based on the sum of similarity over each competing category. This leads to a self-organising behaviour in the mental lexicon, as exemplars will cluster up in distinct categories instead of spreading homogeneously in the category space (Wedel, 2004).

Exemplar-based perception is not restricted to speech. Nosofsky (1988) shows that an exemplar-based model fares better in predicting human subject behaviour in colour categorisation than alternative prototype-based models. An exemplar model’s main advantage is that they can model frequency effects better. Since categories are flexible and are able to shift under the influence of new members, the effects of different token frequencies readily follow from the model’s architecture.

The usual argument against exemplar-based categorisation is that storing all the individual instances takes up too much memory space, and is simply impossible. While exemplar-based theories involve a large amount of hierarchical structure and other measures to tackle this, Johnson (1997; 2005) also notes that human beings seem to have a huge memory capacity, as suggested by experiments on identification and recognition both for words and pictures (Goldinger, 1996; Standing et al., 1970).

Exemplar-based models give us a picture of speech perception that is radically different from the classic one. The input utterance is not de-coded,
Defining Salience

rather it is compared to already existing utterances to determine the category to which it belongs. An important consequence is that category selection is not only sensitive to the semantic content (what the speaker actually said), but also to minor details correlating with speaker age, gender, and identity. That is, there is a vast amount of information available for the language user beyond the conveyed content. This is the reason why exemplar theory enjoys widespread popularity in sociophonetics (Foulkes & Docherty, 2006) – it is the only model of the mental lexicon capable of incorporating all the observed variation. One example of sensitivity to context and speaker identity is Glaswegian (Stuart-Smith, 2007), which neatly illustrates how linguistic innovation in the working class is closely connected to social dynamics and network structures.

As I wrote in Chapter 1, an utterance carries two indices, the actual meaning and social indexation. Exemplar theory allows us to refer to the latter, as it enables us to take sub-phonemic variation into account, which usually carries social indexation and which would not fit into an abstractionist model of the mental lexicon.

A standard objection to exemplar models is that they seem to leave too much leeway for variation, which can be problematic in category formation, and it poses a difficulty in sociolinguistics as well. If people are able to use detailed phonetic information, how come they are sensitive to certain linguistic variables, but not to others? This is where the present formalisation of salience comes into the picture again.

An exemplar-based model of the lexicon allows for the presupposition of transitional probabilities as active factors in speech perception, since a detailed, episodic word storage can contain frequency information (Pierrehumbert, 2001). It comes to more than this, however, since an exemplar-based model, in turn, needs our notion of salience to explain why people pick out certain variables, but not others, despite the fact that theoretically, all are available to carry social indexation.

In this view, salience follows from frequency: a unit with low relative
frequency in the structure becomes salient, and if this unit constitutes a part of a linguistic variable, the variable itself also becomes salient.

Linguistic variables with low TP-s (in the context of comparing dialects) emerge as salient in an exemplar-based lexicon, since they constitute hotspots of attention in the range of information available for the language user. A variable with a low TP, i.e. a high surprisal value, can be reliably used as a sociolinguistic marker since it is salient irrespective of the social indexation it carries. By relying on variables which already have a lapsed distribution, i.e. behave ‘erratically’, the language users narrow down the set of possible carriers of social indexation, which makes both the deployment and recognition of social markers easier. To put it simply, speakers will assert their social identity by using features that are easily noticeable anyway.

2.3 Concluding remarks

This chapter provided an operationalisation of sociolinguistic salience, couched in an exemplar-based model of perception, the dominant framework used in sociophonetics. The argument was that linguistic variables with a high value of surprisal are salient for the language user, and can be used for social indexation. This is in line with the research on the subject in sociolinguistics, where the difference between two types of variables, indicators, and markers, is widely acknowledged. Sociolinguists also tend to attribute this difference to a notion of perceptual salience, which has an external, cognitive grounding.

We looked at the general notion of salience, which covers the sum of properties which make an entity outstanding in a given context. The sociolinguistic sense, though rather vague, is somewhat similar: certain variables are more conspicuous than others, which results in their marked use for social indexation, ranging from unconscious patterns of context-dependent style shifting to being subjects of naïve linguistic awareness. Since sociolinguists tend to look for a grounding of salience in cognition, we also briefly surveyed the term as used in visual cognition. Here again, it is a property of elements
Defining Salience

of a visual array that are sufficiently different from their background to divert attention.

Following up on all this, I established salience as low transitional probability (TP) in a given dialect, compared to a norm dialect. In this view, a variable is salient because its behaviour is different enough from the norm (or another dialect) to be surprising for the language user. Defining salience as dependent on TP-s was paired up with psycholinguistic research on listeners’ sensitivity of such statistical information and the overview of a framework, exemplar theory, which is able to accommodate this type of knowledge.

In an exemplar-based framework, almost all linguistic variation is recorded. Any difference between two dialects can be singled out to be a social marker. Salience narrows down the set of possible variables, rendering social marking more robust and reliable.

In the next chapter, I discuss my methodology of modelling variable salience based on spoken corpora. The following chapters will give examples of linguistic features which are used as social markers and which are also salient in the above sense. These case studies implement our hitherto heavily theoretical formulation of salience in sociolinguistics.
Chapter 3

Methodology

This chapter gives a walkthrough – and, partly, a defence – of the corpus-based methodology used in my case studies. It has two aims, to allow the replication of the results and to convince the reader that the methodology in itself is viable. Section 3.1 gives a broad introduction to how I use corpora to assess segmental distributions which might lead to the rise of sociolinguistic salience, while Section 3.2 is a minute, technical description of how I work with corpora, using a toy example. The chapter, and especially Section 3.2, is not necessary to understand the case studies, as these can stand in themselves. It is a supplement for those readers who are interested in the engineering, not just the broad methodology and the results.

What I do in the case studies is compare the salience of a variable to its speaker evaluation. The latter tells us whether the variable carries social indexation for the speakers. This is established by looking at attitude tests using the variable, style shifting in interviews, and popular stereotypes. Salience is modelled using studies and language corpora. These need to supply two pieces of information: the extent of the use of a variable in a dialect, and the transitional probabilities of the segments that are used to realise the variants. I will clarify this below.
Methodology

3.1 General considerations

The core idea of this work is that the source of a variable’s salience is a difference in distributions – interpreted as transitional probabilities (TP-s) – of the segmental realisation of a variant in two dialects. In most cases, the first dialect is geographically or socially restricted, and this is the dialect where the ‘marked’ variant is characteristic. The second dialect has a broader usage, and can be regarded as a standard.

Salience would be ‘generated’ at the interface of these two dialects. This interface, however, need not be interpreted narrowly, as a speaker of one dialect talking to another, because, for instance, most dialect speakers also have an abstract understanding of the standard language. For example, as Wells (1982) notes, the standard has regional varieties with an extent of local colouring in Britain, but these are still highly similar – in any case, more similar than the local non-standard varieties.

The use of a dialect corpus, such as the FRED corpus (Kortmann et al., 2005) can give a very close approximation of the dialect in use, all considerations of artificiality in speech collected by linguists notwithstanding. (Labov (1966/2006) examines this problem in detail.) It is more difficult to pin down the ‘standard’ part of the comparison: constraints on sounding mainstream are much more present in perception than in production, especially in England, where the classic standard variety, Received Pronunciation, is spoken by an small minority of the population. Besides, I just argued that a notion of the standard is present in dialect speakers as well – it is impossible to compile a language corpus of attitudes in perception.

I take time to defend my choice of corpora in the individual case studies. The aim is always to find a robust corpus for both dialects, which are roughly from the same time period (to exclude the effect of large scale sound changes) and have similar speakers in them. I interpret the dialect-standard relation, the source of salience, in one speaker group or societal unit. If this unit is the size of a city or even a country, one can be relatively flexible with the structure and temporal dimensions of the corpora used. In any case, I rank
size above anything else, because I think the amount of data is the most important factor in the significance of the results, and it can override the effect of other factors.

On the level of segmental phonology it can be assessed fairly accurately what the known processes are that can interfere with the results. This means not taking your corpus data for granted, but rather giving close attention to understanding how it is structured.

At this point, the task looks simple enough: pick a corpus, measure transitional probabilities relevant for your variable, compare with another corpus, discuss the results. The main difficulty lies in the domain of study, phonology. Phonological variation is not marked in written transcriptions of speech (in corpora consisting of conversation or interviews) and is absent from written language, which is not true for morphological or syntactic variation.

Large, phonetically transcribed corpora are rare, unsurprisingly, given the time and effort. There are notable exceptions, like the NECTE corpus (Allen et al., 2007), but the usual case is that audio recordings and accompanying written transcriptions are available, as in FRED. In some cases, even the recordings are absent. If one does not want to give up the robustness offered by available corpora, one has to use written transcriptions to approximate phonological patterns.

In sum, the best case scenario would be a huge, phonetically transcribed corpus. The second best case scenario is a written corpus in which the variant in question is marked, even though the rest is just orthographic transcription. This is the way things stand with Definite Article Reduction (see Chapter 3) in the FRED corpus. It is consistently marked in the interview texts, and since recordings are available for most of the texts, one can sample the audio material to see how exact the transcribers were.

There is, however, no information on the other phonological characteristics of the given dialect in the text, and this is where one has to resort to approximation. My method is to transcribe the material automatically, using a script which replaces the words with their citation forms in phonetic
transcription. This amounts to no more than ‘phonologising’ the English alphabet, as it disregards dialectal features and connected speech phenomena. But since I have information on the phonological characteristics of the dialect, and I can tell whether the difference between what the real thing sounds like and my dummy transcription has any bearing on the behaviour of the variant I study.

In below, I use the example of the variable (ing) in a Lancashire interview of the Fred. This variable entails the realisation of the suffix -ing as [ən] (as opposed to standard [əŋ]), is marked in the transcription. If I phonologise the transcription relying on citation forms, I will overlook all other Lancashire features, such as differences in vowel realisation. I can still formulate meaningful statements on how often I find the suffix [əŋ] before consonant- and vowel-initial words, since no known Lancashire phonological pattern would word-initially turn consonants into vowels or vice versa. (As a matter of fact, the only relatively common phonological variable that does something similar in English is /h/ dropping, the deletion of [h] in the onsets of stressed syllables ([h] is typically not realised in unstressed syllables in general). <h>-initial words, however, are not frequent enough for /h/ dropping to interfere with any other process.)

The outcome is a model of the Lancashire dialect. It is certainly incapable of reflecting all the properties of the original, but it can serve one purpose: give the TP-s for the variant we are interested in.

My results, then, are reliable and valid, even though they were reached through intermediate means, the use of a citation form-based automatic transcription. What I lost with the abstraction I gained with the size of the material I can work with, and size is especially important when one works with probabilities – the larger the sample, the more calculated probabilities will approximate the ideal. (This is precisely analogous with tossing a coin. Here we know that the probability of heads is, all other things being equal, 0.5, and the more observations we have, that is, the more we toss the coin, the more the actual ratio of heads, based on these observations,
will approximate 0.5.)

Things get considerably more difficult when the variant is not marked in the corpus. In this case, the corpus can serve as the basis for the general phonological distributions in a dialect – with the complication of automatic transcription. The extent of the use of a variant can be assessed using independent measures, such as studies on the variant in the dialect. Here, the corpus and the study complement each other: the corpus gives the general speech patterns that are typical to the dialect (such as sentence length, word length, or vocabulary), while the study serves with details on the variant’s behaviour: frequency, environments, realisation.

This approach requires even more caution, as it inserts another step of abstraction: not only is the corpus automatically transcribed, but the extent of variant usage, on which the TP-s are based, is established through an external source. Still, if one pays due attention to the properties of the corpus and the possible risks of the abstraction, it is a powerful tool in assessing segmental distributions and variable salience.

As I note above, this is just a general outline of the methodology of this work. The individual case studies hopefully answer all the concerns that the reader feels unaddressed. The next section does not elaborate this methodology further. Instead, it offers a detailed description of putting it into practice, of interest to readers who would want to know more of the particulars.

### 3.2 Step-by-step corpus editing

This section gives a detailed walkthrough of the corpus-based methodology used throughout this work. It is a technical description, suggested only for those readers who are dissatisfied with the depth of the methodology in any given chapter – in this case they can return here to have a complete view of the process.

As I write in the previous section, the main aim of the method is to get a
sensible approximation of segmental transitional probabilities (TP-s) in real speech. Real speech is approximated by the use of corpora, spoken corpora, when possible, and segmental TP-s are gauged by transcribing the corpora automatically, replacing words with their citation forms. The example I will use in this section is one sociolinguistic interview from the Fred corpus, coded LAN 001.

LAN 001 is an interview with a 76 year-old female from Barrow, Lancashire. The written transcript is a txt, with a word count of 14,400. It has a header (enclosed in square brackets), which contains recording information and transcriber’s notes, such as that the subject tends to ‘drop final -g’. This is followed by main speaker and recording data, and the interview itself. Lines 20-40 are repeated verbatim (except line breaks) below. (The first twenty lines are like ll 21-24, notes on the recording. For details, the reader should consult the Fred documentation.)

[uses singular for plural count nouns: mi first five pound]
[Informant: D1B: Mrs D; born 1899; job: domestic servant, munition worker; education: left school at fourteen; working class; father: sailor, labourer; mother: farm worker before marriage]
[Interviewer: IntER]
[transcript anonymized by ***, July 2005]
When the war started, I think you said you were working for a lady who had some Turkish gentlemen?

(\textit{v 'clock ticks throughout interview'})  

\begin{quote}
\textit{\textbf{u LanD1B}}
\end{quote}

Yes, yes, well, that was it. Well I was sixteen then well that was before the War broke out. And eh, (\textit{v 'ts'}) of course, with being a big family you know, we 're all, eh, eh, well (unclear) father (/unclear) was a navy man, they made us all do a job at home. Well, I, I would rather do a job outside, cause we were always grumbling who should do it. So anyhow, I went to eh, there was two or three of us having a walk around like, like you do, and eh, this lady came up to us lived in an, eh, em, Ramsen Dock Road. So she said eh, all I know is that she was Scotch, she said eh, Do you know of any good girl wants a job? So, I spoke up and I said, Well eh, I says, what kind of a job? She said, of course it was all domestic work then, there was no eh, factories or anything. So I says, What have you to do? She says, Well, you clean up at home, just do what you do at home, you see. She says, Anyhow, if you 'd like to come I 'll, I 'll eh, take you on. So, I said, Well, how much is it? (\textit{v 'laughs'}) Five shillin' a week, (\textit{v 'laughs'}) so -- 

As we can see, although the transcription is orthographic, interruptions, pauses, hesitations, etc. are marked. Since I am interested in the TP-s in the informant’s speech (this being representative of the dialect as a whole) I remove the header and the interviewer’s lines, as well as the bracketed notes in the text (e.g. ‘laughs’, ‘unclear’ etc.). I also remove all orthographic marks, except the dot, and I replace ‘?’ and ‘!’ with the dot as well. This way only sentence boundaries remain marked. The result has a word count of 12,121 and looks like this:

Yes yes well that was it. Well I was sixteen then well that was
before the War broke out. And eh of course with being a big family you know were all eh eh well father was a navy man they made us all do a job at home. Well I I would rather do a job outside cause we were always grumbling who should do it. So anyhow I went to eh there was two or three of us having a walk around like like you do and eh this lady came up to us lived in an eh em Ramsen Dock Road. So she said eh all I know is that she was Scotch she said eh Do you know of any good girl wants a job. So I spoke up and I said Well eh I says what kind of a job. She said of course it was all domestic work then there was no eh factories or anything. So I says What have you to do. She says Well you clean up at home just do what you do at home you see. She says Anyhow if youd like to come Ill Ill eh take you on. So I said Well how much is it. Five shillin a week so

At this point, I use a script written by Márton Sóskuthy. This script takes the above txt as an input, along with a word list that contains pairs of orthographic and phonemically transcribed forms. The script goes through the txt input, and replaces the words in it according to the word list. It writes the results into an output text file. If it does not know a word, it substitutes it with a ‘?’ and writes it into an error list. The resulting output text file looks like this:

Yes 'jEs
yes 'jEs
well 'wEl
that 'D{t
was 'wQz
it 'It
***-*** END of sentence 0
Well 'wEl
I '2
was 'wQz

68
sixteen "sIks’tin
then 'DEn
well 'wEl
that 'D{t
was 'wQz
before bI’f$R
the 'Di
War 'w$R
broke 'br5k
out '6t
***-*** END of sentence 1

This output file has two columns. The first column is the input text file (as seen above), with one word per row. The second column is the transcription. The script notes the end of sentences. These boundaries are based on the positions of the dots in the input text. This means that the script puts sentence boundaries to precisely those places where these boundaries exist in the original.

The DISC phonetic alphabet is used to transcribe the text. This alphabet does not cover all the symbols of the IPA, but it covers phonemic English, and has the major advantage of transcribing all segments as a single character, which renders calculations easier. The only suprasegmental notions are that of primary and secondary word stress. The only difference between the DISC and the IPA in terms of consonants is that [θ] is marked as T, and [ð] as D, whereas [tʃ] is marked as J and [dʒ] as __, while syllabic consonants are also marked differently (for lack of diacritics). For the vowels and syllabic consonants, see Table 3.1 below.

I cut out the second column, and convert line breaks into white spaces, which gives me a text file consisting of a single line, with transcribed words and sentence boundary-marking dots with spaces in between:

'jEs 'jEs 'wEl 'D{t 'wQz 'It . 'wEl '2 'wQz "sIks’tin 'DEn 'wEl 'D{t 'wQz bI’f$R 'Di 'w$R 'br5k '6t .
Table 3.1: The IPA and the Disc phonetic alphabet

<table>
<thead>
<tr>
<th>IPA</th>
<th>Disc</th>
<th>IPA</th>
<th>Disc</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>i</td>
<td>o</td>
<td>Q</td>
</tr>
<tr>
<td>u</td>
<td>u</td>
<td>u</td>
<td>U</td>
</tr>
<tr>
<td>a</td>
<td>#</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>ø</td>
<td>$</td>
<td>ø</td>
<td>@</td>
</tr>
<tr>
<td>æ</td>
<td>3</td>
<td>η</td>
<td>C</td>
</tr>
<tr>
<td>i</td>
<td>I</td>
<td>ι</td>
<td>F</td>
</tr>
<tr>
<td>ε</td>
<td>E</td>
<td>ι</td>
<td>H</td>
</tr>
<tr>
<td>ei</td>
<td>1</td>
<td>ai</td>
<td>2</td>
</tr>
<tr>
<td>oi</td>
<td>4</td>
<td>οu</td>
<td>5</td>
</tr>
<tr>
<td>au</td>
<td>6</td>
<td>οe</td>
<td>7</td>
</tr>
<tr>
<td>εe</td>
<td>8</td>
<td>εe</td>
<td>9</td>
</tr>
<tr>
<td>æ</td>
<td>{</td>
<td>l</td>
<td>P</td>
</tr>
<tr>
<td>r</td>
<td>R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This final output file has 11,522 words. This includes the 550 dots, which means that around 1000 word tokens are missing after the conversion. Some of this difference is simply a residue of how word counts are calculated. Some missing words, however, are due to the conversion process: unknown words are transcribed as ‘?’, and I remove these from the final output before I count the TP-s. As I note above, the script makes a record of unknown words in a separate file. When I work with larger corpora, I iterate transcriptions based on this record: I run the script, add unknown words (usually proper names, different spellings, and dialect forms) to the word list used by the script, run the script again, and so on. At some point, however, I ignore the unknown words that remain, because they are very low frequency in the original input. This are usually proper names or idiosyncratic spellings.

The very thing itself that we are looking for is, of course, also spelled in an idiosyncratic way – where the informant realises a coronal variant of (ing), the original transcriber marks it with a dropped <g> in spelling, such
as \textit{<shillin>} for [\textit{jilm}]. These are taken into consideration (cf. below).

In the present case, the final error list contains 197 tokens – this is the total number of words missing from the output because they were not transcribed. Compared to the word count of 11,522, 197 is negligible.

### 3.3 Calculating transitional probabilities

In the previous section we arrived at the converted corpus text, the basis of comparisons. It is a useful basic tool, inasmuch as it is robust and it is based on natural speech. Its main weakness is that it consists of citation forms, so any dialectal phonological process that is present in the original has to be considered separately.

As I note at the beginning of this chapter, the easiest case is a corpus where the dialectal variant in question is already marked, as in the case of Definite Article Reduction. In the present case, one example of this could be the variable (ing), the realisation of the suffix \textit{<-ing>} ([\textit{\textalpha{n}}]) as [\textit{\textalpha{n}n}]. The use of variants of (ing) is noted in the header of LAN 001, and is marked throughout the text. The accuracy of the transcriber can be verified based on audio/visual inspection of the audio recording, or, if this would take too much time, parts of the recording.

To remain at the variable (ing), this is marked in the text by omitting the final ‘g’ of the word (e.g. ‘shillin’ for ‘shilling’). Since these are not dictionary citation forms, they have to be included in the word list of the script separately. After the modification of the word list, the script replaces these forms as having a coronal nasal ending. That is, ‘shillin’ will become [\textit{j\textalpha{n}n}] (or [‘\texti{S}{\textalpha{I}{\texti{l}}}\textalpha{n}n] in the Disc alphabet). Similar systematic deviations from standard spelling (such as \textit{<t>} for the reduced definite article) are taken into consideration during transcription. After removing stress markers from the output text file (we are not interested in stress for now) one can look at bigram segment frequency and calculate transitional probability.

Since variation in (ing) changes the state of things word-finally, it makes
Methodology

sense to look at the probability of having [n] word-finally in the corpus, that
is, \( p(\#|n) \). The total number of [n]-s in the modified output text file based
on LAN 001 is 1876. There are 782 word-final [n]-s, that is, ‘n ’ sequences
(as there is white space between words). Using the equation to calculate
transitional probability (in Figure 3.1), we get the TP for \#|n\, 0.415.

\[
p(Y|X) = \frac{\text{likelihood of pair } XY}{\text{likelihood of } X}
\]

Figure 3.1: Probability of Y following X

This number is not interpretable in isolation. In order to find out whether
coronal variants of (ing) can be salient, one would have to look at other
TP-s (such as the probability of a word-final [ŋ], or even [ɔŋ]/[ɔn] in this
dialect sample as well as another dialect sample, for instance from Southern
English, or a local, Lancashire standard, to see whether there is a large
enough difference. The point here, however, is that this is the way we reach a
particular TP.

Things get more difficult if the variant we are curious about is not present
in the original transcription. In this case, I use independent research on the
use of the variable in the dialect, and base my calculations on the modified
corpus as well as the results of this research. This is the method used in
Chapter 4, with glottalisation in the South of England.

I will explain this method using a hypothetical example. I would like to
know the TP of a glottal stop preceding a vowel. Let’s say that a survey
was conducted in Barrow, with older speakers, in the 1970-s. The speakers
were recorded in standard sociolinguistic interviews. This is the place, time,
register and speaker base of our example corpus, LAN 001. This survey found
that all word-final /t/-s are realised as glottal stops, but they remain [t]-s
everywhere else. This pattern is independent of all other factors, such as
frequency or word class. (It is unlikely that such a dialect ever existed, but
this is again a hypothetical example.)

As I know that the speaker group to which the speaker in LAN 001 belongs
Calculating transitional probabilities

behaves in a certain way, I can assume that the speaker will behave in a similar way, and put forward 100% glottalisation of /t/ word-finally, and 0% elsewhere. What I would like to know is \( p(V|?) \), the probability of a vowel following a glottal stop. I look at connected speech, so basically I want to know the chance of a word starting with a vowel after a word ending in a \([t]–\) I know that the latter is glottalised, since it is word-final. The number of these sequences ([t V]) in the example corpus is 277. There are 1239 word-final [t]-s (that is, in this view, glottal stops) in the example corpus. This gives a TP of 0.223 for \( V|? \).

One important caveat of this method is that one has to pay attention to potentially intervening phonological patterns in the corpus dialect, because the results of these processes are not marked in the transcribed corpus. As noted above, /h/ dropping can potentially alter the amount of vowel-initial words, and therefore alter the TP above. One can assess whether a process is frequent enough (in terms of number of realisations) to have a strong effect.

This is the detailed outline of the basics of the method used in the case studies. In the case studies I elaborate on it and explain my choice of corpora, studies, and all the potential intervening factors.
Chapter 4

Definite Article Reduction

The previous chapters introduced the notion of salience as a variable’s surprisal, which allows its use as a sociolinguistic marker. I also promised concrete examples supporting this hypothesis. The first one of these, definite article reduction, is the subject of this chapter. Though I defined salience as a general distinction in surprisal or novelty, in this chapter, I will argue that the variable in question is not only salient because of the different segmental distributions, but also because these distributions enhance its role as a word boundary marker.

As we might recall from 2, the original purpose of calculating transitional probabilities was finding word boundaries (Harris, 1955). Boundary markers and listener sensitivity to such elements in the speech string also feature prominently in Trubetzkoy (1939). Trubetzkoy, after looking through the taxonomy of boundary markers in various languages, comes to the conclusion that these are probably the most conspicuous elements of the language input for any listener. Though my definition of sociolinguistic salience is not necessarily connected to the boundary marking function of transitional probability differences, the variable examined in this chapter, definite article reduction, happens to be a case where the variable’s salience is reinforced by its status as a good word boundary marker.

Section 4.1 goes through the properties of the variable, as well as its
Definite Article Reduction

literature, while section 4.2 explores its salience as a consequence of its probability distributions, with an emphasis on how these probabilities lead to the variable’s status as a good word boundary marker. Section 4.3 gives a summary and discusses the possible caveats of the analysis.

4.1 Background

The term ‘definite article reduction’ (DAR) refers to a dialectal variable in the North of England, where the definite article has non-standard realisations. The variable appears everywhere with the exception of Northumberland, but is generally associated with the historic counties of Yorkshire and Lancashire. Contrary to what the term implies, DAR forms are not commonly regarded as reduced variants of the standard definite article.

4.1.1 Details of the process

The process itself is outlined by Wright (1905) who argues that the reduced definite article is a single voiceless stop before consonant-initial words and a voiceless dental fricative before vowel-initial ones. Jones (1999) says that this view is oversimplified, pointing to the Survey of English Dialects (Upton et al., 1994), which lists 16 different variants. The most common ones are [t] (commonly with an extent of pre-glottalisation) and especially [ʔ], which suggests that the reduced article is a constriction, usually a glottal one. Some of the variants, like the dental fricative, only occur in certain dialect areas (from those that feature article reduction). Lodge (2010) gives the realisation of the definite article in parts of Lancashire and Yorkshire as the glottal stop both before voiced and voiceless stops (cf. table 4.1 for examples).

Various types of reduced articles can occur, along with the standard variant, in the production of one speaker. This is illustrated by an excerpt from the FRED corpus below. The speaker is an 80 year-old male from Yorkshire, the reduced article is transcribed as <t> or <th>. This transcription here refers to a [ʔ] and a [θ] realisation.
And, the handyman that lived here, S(...), they went into t’ s-, in that stable, with flames going out ten foot high above t’ stable, all full length, because we ’d put, we had a loft full of hay, we ’d just put two loads of hay up i’ t’ loft and flames coming out o’ t’ top, and all their manes, the racks was full of hay, and they were on fire, it burnt all the fringes, their fringes off, and th’ front off th’ horses.

Not only does definite article reduction show considerable variation, it also alternates with the standard variant of the article, the, with the latter occurring in certain contexts, such as in careful speech or in idioms.

The precise extent of the use of DAR is debated. On one hand, it does not occur frequently in corpora. Tagliamonte & Roeder (2009) find 13.7 per cent in their analysis of the York Corpus, and I also found a ratio of 13.8 per cent in the Yorkshire part of the FRED corpus. On the other hand, Jones (2007) notes a much larger extent of DAR use in certain studies. His observation is corroborated by Glauser’s excellent description of the dialect of Grassington, North Yorkshire (Glauser, 1984). He reports the direct opposite of the FRED results (for instance), by putting the use of the standard form to 8.5 per cent in the speech of three interviewees (cf. table 4.2).

The findings of Jones and Glauser hint at a larger extent of DAR use in a more familiar setting. Glauser’s data have to be handled with care, though, as the reported numbers are surprisingly low. In comparison, the average amount of definite articles in the FRED interviews is much higher, around four thousand per interview (For instance, it is 5900 for the second Yorkshire interview and 3200 for the third one.)

<table>
<thead>
<tr>
<th>the day</th>
<th>[θdeɪ]</th>
<th>the inn</th>
<th>[θɪn]</th>
</tr>
</thead>
<tbody>
<tr>
<td>the pub</td>
<td>[θpub]</td>
<td>the apple</td>
<td>[θæp]</td>
</tr>
<tr>
<td>the cooker</td>
<td>[θˈkuko]</td>
<td>the order</td>
<td>[θoʊɹ]</td>
</tr>
</tbody>
</table>

Table 4.1: Reduced definite articles (examples from the FRED corpus)
Rupp & Page-Verhoeff (2005) point out that DAR forms seem to be favored over standard forms in case of familiar referents, connecting the use of DAR with information structure. In the York study of Tagliamonte & Roeder (2009), young male speakers use DAR to an unexpectedly large extent, which hints at its role as an identity marker. This would also suggest that speakers use DAR to a larger extent with a familiar interviewer and suppress it with a standard speaker, which would explain the variation between the attested numbers in different studies.

Rupp (2007) and Jones (2002) agree that, in spite of the name of the process, the ‘reduced’ tokens do not come from the phonetic reduction of the standard definite article. The actual diachronic process giving birth to DAR is rather unclear, but was certainly entangled with the reduction of Middle English determiners, and involved analogical levelling in favour of the [?] forms.

In sum, DAR seems to be sensitive to the phonetic environment (the following segment and maybe even the preceding one), as well as the discourse structure. It is affected by the social dynamics in the dialect areas, and has a large range of variants, some occurring only in restricted dialectal areas, others in overlapping larger ones. These aspects interact: for instance, if there is a fricative variant, it can consistently occur before vowel-initial words, confining the plosive variant to pre-consonantal environment, even though in another dialect, which lacks the fricative form, the plosive occurs everywhere.
In general, the plosive is a glottal stop, but it can show up as a glottalised coronal stop, or only as creaky voice, which is possibly the auditory correlate of a non-complete constriction of the glottis.

The reduced articles vary with the standard ones, even in the production of one speaker – largely, but perhaps not exclusively, depending on the formality of the register. Jones (2007) adds yet another aspect: glottal stops occur in standard English, as pre-consonantal and word-final voiceless stops, especially /t/, often debuccalise and become glottal stops. These glottal stops, however, differ from the segmental realisation of the reduced article. As he concludes from a three-speaker study, the latter are shorter and more variable than the glottal stops coming from consonantal lenition.

Jones links this difference to the presence or absence of a morphological boundary. His examples can be seen in table 4.3. In the first example, there are two morphological boundaries around the glottal stop (indicated by a <t> in the orthography): #they#see#t#sacks#. In the second example, the glottal stop belongs to the previous word: #they#seat#sacks#, so there is only one flanking the glottal stop. This adds yet another aspect to definite article reduction, that of word boundaries.

Jones’ other main finding is that DAR segments systematically differ from cases where there is no article at all. In the DAR case, there is always an extent of laryngeal activity, which strongly suggests that equating DAR realisations with null articles (which are also possible in the Northern dialects) would be misleading. He claims that there are systematic differences between the DAR glottal stop realisation and glottal stops that are allophonic realisations of [t], the latter being longer and more prominent. The debuccalisation of
Definite Article Reduction

/[t/] is not the only possible source of glottal stops in standard (i.e. DAR-less) English English. Post-vocalic fortis stops are usually pre-glottalised (Wells, 1982), and vowel-initial words can have an initial glottal closure (very much like in German), especially if they stand at the head of an intonation phrase (Dilley et al., 1996). How these different glottal stops behave from a perceptual-phonological point of view is a question to which I will later return.

4.1.2 DAR as a salient variable

There is a difference of opinion in the DAR literature, both on the different realisations of the reduced article and on the extent to which these are used by the speakers. What can be safely said is that in Yorkshire and Lancashire, the standard definite article displays non-standard allomorphy, occurring in 10-14 per cent of the cases, at least in certain registers, usually in the form of a glottal constriction, though fricative realisations are also possible.

DAR seems to fit most requirements of being a marker. It fits Trudgill’s extra-strong salience or Labov’s stereotype, though the two are not, strictly speaking, synonymous (cf. Chapter 2). As Jones (1999) notes, it has long been a tool of illustrating typical Northerner speech and, based on their data, Tagliamonte & Roeder (2009) reach the conclusion that it plays an important role in establishing identity in the city of York. (I have spoken to more than one people living in or coming from Yorkshire who lamented on the local custom of ‘dropping your articles’.)

Considering all this, one can safely say that DAR is a dialectal marker, used to express social or regional differences. It is a feature that distinguishes DAR dialects from the standard spoken both in the North and the South. This leads us to the question of salience: if it is a marker, it must be salient. The critical question is where this salience comes from. In section 4.2, I will try to find a possible source of the salience of definite article reduction.
4.2 Analysis

This section investigates the connection between the salience of DAR and its probabilities. I will argue that the main source of the variable’s salience is negative frequency, that is, low transitional probabilities in the signal. In the case of DAR, however, it is not self-evident that this is the only way salience can be linked to frequency. Consequently, I go through the possible other links between salience and frequency, to reach the conclusion that the most satisfactory account points to the role of transitional probabilities, and, specifically, word boundary marking.

Section 4.2.1 discusses the methodology, section 4.2.2 enumerates the possible links between frequency and salience, section 4.2.3 introduces an alternative approach, further explored by section 4.2.4, and section 4.3 provides an interim summary.

4.2.1 Methods

My main data source in this section is the Freiburg Corpus of English Dialects (Fred, Kortmann et al. 2005). It is a monolingual spoken-language dialect corpus which contains full-length sociolinguistic interviews with native speakers from the British Isles. It reflects the traditional varieties found there during the second half of the 20th century. It contains 2.5 million words and 200 hours of recorded speech. The Northern English part of the corpus, used in this paper, contains approximately five hundred thousand words. The majority of the interviews were recorded in the 1970s and 1980s. Since the corpus aims at regional rather than social variation, the majority of the speakers are NORMs – non-mobile old rural males. (65 per cent of the speakers are male and 74 per cent are over 60.) The raw (pre-editing) word count is 11,821 for the Yorkshire part (eleven interviews) and 14,801 for the Lancashire part (twenty-three interviews).

The interviews of the corpus were recorded by native speakers, and transcriptions were performed by trained phoneticians, which means that Fred is
Definite Article Reduction

a fairly reliable source on dialectal morpho-phonological phenomena. Nevertheless, any data on DAR use were double-checked using both the transcribers’ notes (available for each interview) and the recordings.

Other corpus-based results on DAR are also used as a reference. My main source on word frequency data is the Celex corpus (Baayen et al., 1993), which is essentially a frequency list based on 17.9 million items. The reason why I do not quote word frequencies from the Fred corpus directly is that the difference in size makes Celex more robust and reliable as a data source. It does not have anything to say on connected speech sequences or definite article reduction, so generalisations on these come either from Fred or other studies.

4.2.2 Salience from token frequency

Going back to the original question, that is, why DAR is a salient variable, as simple answer would be that there is plenty of it. This assumption is, however, problematic in two ways. First, as we have seen in Chapter 2, the link between salience and high token frequency is debated. Second, the ratio of DAR in corpora is relatively low. In the Yorkshire part of the Fred corpus, reduced forms make up 13.8 per cent of all the definite articles. This ratio is 9.2 per cent in the Lancashire part. This is, however, not a trivial issue. I have already touched upon other surveys, such as Glauser (1984) on Grassington, reporting higher extents of DAR use.

On one hand, one can stress that lower DAR use appears to be typical of contact situations (where dialect speakers are in interaction with a non-local interviewer). This is the reason for the lower numbers in the corpus studies. Salience is assessed in precisely such situations, which could justify using the lower ratios in a discussion of the variable’s salience. On the other hand, as I note above, the direct link between token frequency and salience is empirically debated, which means the differences in DAR ratios could be, in this case, ignored altogether.

One could argue that around ten per cent of all definite articles, the
numbers we see in corpora, is still quite a lot. This is true on the lexical level, but definite article reduction is a morpho-phonological pattern. As such, it should be compared to phonological variables to see whether its frequency amounts to anything. One possible parallel is the debuccalisation of /t/ before obstruents (as in *Gatwick, butler*), a variable which entails the almost complete substitution of /t/ with a glottal stop in these positions, and which carries no social indexation, at least in the South of England (Fabricius, 2000). In a sample of the South-East part of the *Fred* corpus (wc.: 80,000), there are 3290 [t] plus obstruent sequences, both word-medially and across word boundaries. In comparison, the number of definite articles in this sample is 3100. Ten per cent of this number is 310, which is much smaller than the number of [t] plus obstruent sequences. Yet, the former is salient, while the latter is not. (The South-East part was chosen because it contains no definite article reduction.) It is interesting to note that the odds do not change in DAR’s favour even if we assume its widespread use – if all definite articles are realised as reduced articles, they still cannot outweigh the non-salient /t/ glottalisation pattern.

Another hypothesis would be to assume that DAR is not salient by virtue of its large numbers, but rather because it prefers to co-occur with more frequent (preceding or following) collocates. If we assume an exemplar-theoretic lexical representation, all tokens are registered in the memory. Consequently, types with larger frequency of occurrence have stronger representations, which are easier to access. When, in turn, DAR goes more often with types that have strong representations in the lexicon, it will be accessed more easily and therefore it will be more salient.

Siewerska & Hollmann (2011) look at the frequency of the collocates of the reduced article and find a tendency of DAR occurring with more frequent collocates on its left. DAR tends to prefer preceding prepositions, which are usually frequent in themselves. They admit, however, that this correlation is not clear-cut, as the standard article also occurs often in prepositional phrases, and suggest that a better explanation would be to link the preference
of DAR to the prepositional schema.

I found no correlation between DAR and the frequency of the right hand collocate in the Yorkshire part of the Fred corpus. This right hand collocate is overwhelmingly the noun of the determiner phrase. The methodology was as follows: 134 DAR collocate types and 676 standard definite article collocate types were extracted from the Yorkshire part of the corpus (wc.: 80,000). Their frequency was gained from the Celex corpus, and a t-test was performed on the two sets, with the hypothesis that the DAR collocates are, on the whole, more frequent. The results refute this hypothesis, as there is no significant difference between the frequency of the DAR collocates versus the standard definite article collocates ($t = -1.5509$, df = 595, p-value = 0.9393). This result can be accepted with two provisos: first, types, not tokens were used – since the token frequencies are given in Celex, no detail was lost, except for cases were one type occurs both with DAR and the standard article, yet with overwhelmingly different proportions. The effect of this difference in variations, however, seems negligible. Second, frequencies come from Celex, not the Fred corpus, and, as Siewerska & Hollmann (2011) also note, some words are used much more often in the Northern dialects than in general. Then again, frequency data from the former is still more reliable than from the latter, due to the relatively small size of the Fred corpus.

4.2.3 Salience from transitional probability

After looking at other possible frequency effects, I return to my original claim, namely, that DAR’s salient status comes generally from its transitional probabilities and specifically from its status as a good word boundary marker. For the sake of simplicity, I will assume the segmental realisation of DAR to be a glottal stop. This move is warranted by the available data. For example, only 2.4 per cent of the reduced articles are transcribed as fricatives in the Lancashire part of the Fred corpus. This ratio is below one per cent in the Yorkshire part. However, assuming that all glottal DAR allomorphs are facsimile glottal stops is severely abstracting, and I will provide additional
detail wherever possible.

The quirk in the analysis is that we do not look at the probabilities of DAR at word boundaries. This would not lead very far for two reasons. On one hand, if DAR is regarded as a separate segment, absent from the standard, it will always be surprising, simply because the standard lacks it completely. On the other hand, it will be an excellent boundary marker, since it will always stand at a boundary, as articles are wont to do. (Though it is useful to stress the absence of article reduction in the standard, the statement is something of a tautology, as the term *standard* in this paper is usually employed to refer to a DAR-less dialect, even though it will be elaborated below.)

The assumption behind the use of TP-s in grounding the salience of DAR is that it is similar to segments already present in the standard, namely, glottal stops that result from various phonetic-phonological processes, detailed below. DAR as phonological segment, then, will not be totally surprising for the listener; the major difference lies in its distributions, not in its existence.

As stated above, salience comes forth in a dialect contact context, when a speaker of the standard encounters the speaker of a DAR dialect. This leads us to the question of defining the standard in the context of definite article reduction. A trivial choice in England would be the Received Pronunciation (RP), or at least a regional variety thereof (Wells, 1982). While RP does not necessarily have much respect in the North of England, its use can be justified as it is a dialect without definite article reduction, and ‘standard’, in the DAR case, refers precisely to such a dialect. That is to say, the quality of the Strut vowel, for instance, is immaterial to the perception of DAR – though, of course, not to the perception of the Northern dialect that contains it.

On the face of it, assuming RP as a starting point makes the job very easy: if the DAR allomorph is regarded as a glottal stop, it will be a perfect boundary marker word-initially, as RP has no word-initial glottal stops. Therefore, if a standard speaker encounters a word-initial glottal stop, it will have perceptual prominence, as it certainly marks a word boundary.
This claim has to be elaborated, however, as word-initial glottal stops do occur in English, even if not phonemically. Vowel-initial words can have an initial glottal stop, much like in German. Dilley et al. (1996), in a study of five American English speakers in a radio speech corpus, puts the extent of initial glottal stops in vowel-initial words between 13 and 44 per cent. If our hypothetical standard speaker displays even the lower end of this variation (around 10 per cent), word-initial glottal stops ‘cease to be interesting’.

The main point of comparison is the standard definite article. Its part adjacent to the word is either [ə] before consonant-initial words or [i] before vowel-initial ones. DAR can outperform the standard article as a boundary marker if it is less likely to appear in this position in the standard dialect as opposed to the DAR dialect. This is the case if we assume no word-initial glottalisation for the standard, but this difference wanes if we put it around ten per cent, as seen in table 4.4.

<table>
<thead>
<tr>
<th>Segment</th>
<th>DAR dialect</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>[ə]</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>[ʔ]</td>
<td>0.007</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Table 4.4: Glottal stops word-initially

The numbers in this table come from the Fred corpus. They are probabilities for the occurrence of the segments in the first column, word-initially. The DAR numbers reflect a dialect with definite article reduction, while the Standard numbers come from the South-East part of the corpus, which has no DAR (the use of this part is justified further below). As can be seen, in the DAR dialect, the probability that a word starts with [i] or [ə] is 0.07, which is one order of magnitude greater than the probability of a word starting with a glottal stop. This means that DAR is a good boundary marker word-initially, both in the dialect itself and in comparison with a DAR-less standard, where the vowel probabilities are roughly the same, but there are no (phonemic)
glottal stops word-initially.

The situation drastically changes, however, if we assume word-initial glottalisation for even as little as ten per cent of the vowel-initial words in the standard (column three): now there is no strong difference between the DAR dialect and the standard, and, therefore, the status of DAR as a boundary marker diminishes.

Even so, by looking at both Dilley et al. (1996) and Jones’ study on DAR glottal stops and glottal stops resulting from [t]-glottalisation, we can assume that the three glottal stop realisations are different. This could mean that word-initial DAR glottal stops remain prominent for a standard speaker irrespective of the extent of word-initial glottalisation in the speaker’s own output, simply due to its differing phonetic characteristics.

The only other process giving rise to glottal stops (or, at least, glottal constrictions) in English is the pre-glottalisation of fortis stops (Giegerich, 1992). Since this process is confined to the syllable coda (or, at most, to ambi-syllabic positions), it is irrelevant when it comes to the word-initial distributions. The situation is different word-finally: debuccalisation, the loss of place of pre-glottalised fortis consonants, is fairly common word-finally, so the occurrence of a glottal stop would not be unusual in this position.

This is, of course, a simplified account. Word-initial glottalisation only occurs with vowel-initial words, so DAR could remain a good boundary marker before consonant-initial ones despite extensive glottalisation of vowel-initial words. In the same way, word-final debuccalisation is, at least in higher registers, avoided before a following vowel-initial word (Fabricius, 2000), which would, again, render DAR more prominent word-finally if a vowel-initial word follows.
4.2.4 Further arguments for phonotactic distinctiveness

So far I looked at phonological arguments for and against the prominence of DAR word-initially and word-finally. As we could see, its status as a boundary marker largely depends on the extent of other processes in the standard, namely, word-initial and word-final glottalisation, as these create similar sequences. Now I turn to the behaviour of connected speech sequences involving DAR, to see, if these can prove salient in comparison with a standard. This section relies less on phonological arguments and more on statistical methods. DAR is still assumed to be a glottal stop, which, again, is phonetically interpreted as a glottal constriction of a varying extent.

Depending on the previous and the following word, the connected speech sequences involving DAR can come in four forms:

- \( C[?]C \) (preceding consonant-final and following consonant-initial word)
- \( C[?]V \) (preceding consonant-final and following vowel-initial word)
- \( V[?]C \) (preceding vowel-final and following consonant-initial word)
- \( V[?]V \) (preceding vowel-final and following vowel-initial word)

The two aspects of the markedness of these clusters come from (i) the likelihood that the above clusters might occur in the speech stream if the speaker’s dialect has DAR and (ii) their general phonotactic markedness (low probability or outright non-existence). The first question can be answered by looking at the general probability of words ending or starting in vowels or consonants. I used the ratios extracted from the Yorkshire part of the Fred corpus to get the probabilities of DAR occurring with the different environments (cf. table 4.5).

The way to find out the probability of DAR in a particular environment is to calculate the probability of that environment in the first place. In the case of \( _C|C_ \), the likelihood of finding this sequence is 0.65*0.66. The probability
of a consonant-initial word following a consonant-final one is, then, 0.43, so given the null-hypothesis that DAR has the same chance to occur anywhere, these sequences will be the most abundant. Needless to say, a cluster of three consonants with a glottal stop in the middle is extremely marked in phonotactic terms. It rarely occurs word-internally, and if the consonants in question are stops, it does not occur at all, which also increases the salience of such sequences existing by virtue of DAR.

It has to be immediately added, though, that the null-hypothesis of DAR equally occurring anywhere is refuted by the results of the multivariate analysis performed by Tagliamonte & Roeder (2009) on parts of their York Corpus. They show that DAR has a larger probability to occur after vowels (cf. table 4.6).

One aspect of the multivariate analysis is that DAR is more easily distinguishable when standing before a vowel-initial words. This might lead to a bias in the data collection, which, in turn, comes up in the results, which might lead to the preference of the pre-vocalic position being an artefact of the analysis. Still, even if we take these findings into account, the weight of C[?]C clusters does not become much smaller in comparison with the others. Although DAR – supposedly – occurs more after vowel-final words, these words are relatively rare, and consonant-initial words are as preferred as vowel-initial ones (treating approximants as vowels), if not more.

Going back to the original question of connected speech sequences, the possible environments are listed in table 4.7. The first column lists the
Definite Article Reduction

<table>
<thead>
<tr>
<th>Environment</th>
<th>factor weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>After vowel</td>
<td>0.65</td>
</tr>
<tr>
<td>After liquid</td>
<td>0.63</td>
</tr>
<tr>
<td>After nasal</td>
<td>0.58</td>
</tr>
<tr>
<td>After stop</td>
<td>0.25</td>
</tr>
<tr>
<td>After fricative/affricate</td>
<td>0.10</td>
</tr>
<tr>
<td>Before vowel</td>
<td>0.38</td>
</tr>
<tr>
<td>Before approximant</td>
<td>0.50</td>
</tr>
<tr>
<td>Before liquid</td>
<td>0.56</td>
</tr>
<tr>
<td>Before nasal</td>
<td>0.58</td>
</tr>
<tr>
<td>Before stop</td>
<td>0.51</td>
</tr>
<tr>
<td>Before fricative/affricate</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Table 4.6: DAR and phonetic environment

<table>
<thead>
<tr>
<th>Environment</th>
<th>DAR Probability</th>
<th>Markedness</th>
<th>(example)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C[?]C</td>
<td>0.43</td>
<td>very high</td>
<td>feed ? cat</td>
</tr>
<tr>
<td>C[?]V</td>
<td>0.22</td>
<td>high</td>
<td>eat ? apple</td>
</tr>
<tr>
<td>V[?]C</td>
<td>0.22</td>
<td>low</td>
<td>free ? dolphin</td>
</tr>
<tr>
<td>V[?]V</td>
<td>0.12</td>
<td>low</td>
<td>throw ? orange</td>
</tr>
</tbody>
</table>

Table 4.7: DAR & connected speech sequences 1

environments, the second one their probabilities to occur in a DAR dialect. The third lists phonological markedness levels. Column four gives examples. Markedness is interpreted here as the probability of these sequences in a standard dialect.

Table 4.8 supplements the previous one with reference data on the frequency of such sequences in a dialect without reduced definite articles, thereby giving substance to claims on markedness. I used the South-East part of the FRED corpus (wc.: 80,000). The South-East was chosen because it is probably the closest to the linguistic norm in England, and mainly because it
is a dialect area without definite article reduction. What the numbers actually say, is, in fact, the number of the given sequences with an intervening [t], rather than a glottal stop. The reason for this is that /t/ is the most liable to lenite into a glottal stop. It is especially prone to do so pre-consonantally, but also to an extent intervocalically. (For the extent of /t/ glottalisation in the linguistic standard, cf. Fabricius 2000.)

Here, column three gives the overall counts of the sequences in the corpus, while column four gives the word-medial counts. The former give a hint on the general surprisal levels: a C[t]C sequence is considerably rarer than a V[t]C one (we expect /t/ glottalisation in both of these environments), while the former is more likely to occur through DAR than the latter. This, again, means, that definite article reduction will be the most salient for a standard speaker in the environment where it is most frequent.

Column four connects DAR probabilities with word boundary marking. Since, again, C[t]C (practically, C[?]C) is the least likely sequence within a word, it is the most likely one to mark a word boundary. Glauser (1984) discusses the phonetic realisation of these sequences in DAR dialects, and comes to the conclusion that a reduced article, even if it would have been realised as a [t], which is also possible, loses most of its place cues if there is an adjacent consonant, which gives us a placeless glottal constriction.

In sum, the two sequences which have a larger probability of occurrence, C[?]V, and, especially, C[?]C, are heavily marked: they have a relatively small chance to occur in a dialect without definite article reduction. This
is corroborated by the reference numbers from the South-East part of the Fred corpus. The two sequences which are more likely to occur even in a dialect without DAR, V[?]C and V[?]V, have a low probability to show up anyway. This means that, even in a larger context, DAR will stick out: it will be perceptually prominent due to the frequencies of its various realisations.

There are three good reasons to use the number of fortis coronal stops as a reference. As noted above, these stops are the most prone to lenite into glottal stops in English. This follows from their phonetic characteristics: coronal stops are the most variable and have the weakest place cues in the stop series, which means they are easily interpreted without place in a pre-consonantal setting (Paradis & Prunet, 1991). This is backed up by Glauser’s observations on the behaviour of reduced articles realised as single fortis coronal stops in Grassington. He notes that coronal realisations flanked by stops are more inclined to lose parts of their realisation, which can easily lead to debuccalisation. As it can be seen in table 4.9, a stop will lose one of its three components (obstruction, stop/closure phase, release) when flanked on either side by another stop – this can be extended to any other consonant as well. If it is flanked on both sides, all that remains is the closure phase, which is then likely perceived as a glottal stop.

<table>
<thead>
<tr>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>obstruction, stop, release</td>
<td>stop, release</td>
<td>stop, release</td>
</tr>
<tr>
<td>obstruction, stop</td>
<td>stop</td>
<td>stop, release</td>
</tr>
</tbody>
</table>

Table 4.9: stop clusters

Again, the salience of DAR seems to have two different sources. The consonant-vowel ratios of connected speech sequences will result in its frequent occurrence in certain positions as opposed to others. These happen to be the positions in which similar sequences are the least likely to arise in a dialect without DAR. Statistics are not solely responsible though, as the
phonological patterning of English also plays an important role in marking these environments.

4.3 Concluding remarks

In the previous sections I proposed that the reason behind the salience of definite article reduction is its unlikelihood in particular positions in the speech string. The reduced allomorph is typically a glottal stop, and glottal stops have extremely low probability of occurrence word-initially. The same is true for clusters that might result from the insertion of a reduced article. As low transitional probability is exploited by the listeners when processing the speech signal, reduced articles are perceptually more prominent and therefore more salient. What is more, they can enhance the discovery of word boundaries, which adds to this prominence.

The argumentation was supported by models of Northern speech and the comparison of probabilities – which can be seen as surprisal values for the listeners – between dialects with and without DAR.

We ought to bear in mind a number of caveats. First, there is the problem of frequency. If we follow descriptions like Glauser (1984), and assume much larger extents of DAR use, the argument for DAR’s position as a good word boundary marker still stands. As a matter of fact, it is strengthened, since the more reduced articles are attested, the more robustly they can mark boundaries. The problem is, then, that the salience of DAR could stem simply from its token frequency. There are two arguments against this: first, at least in certain registers, DAR seems to have a low frequency – again, around ten per cent of all articles. Second, the seemingly intuitive connection between high token frequency and salience is far from unequivocally established. If one insists on the role of high frequency in the variable’s salience, it remains true that such a view is compatible with the analysis presented here.

Second, this account – and the whole dissertation – takes a segmental phonological approach, which necessarily entails abstracting away from large
Definite Article Reduction

amounts of phonetic detail. Abstraction is a good thing, because it allows establishing wider generalisations on the data, without getting lost in the particulars. Still, whenever it felt needed, I explored the phonetic details of the environments involved, so as not to be ignorant of characteristics that would have otherwise been below the level of abstraction.

Third, low transitional probabilities are certainly not the only source of language user attention on the variable. Other possible reasons include the phonetic prominence of the reduced articles, as well as their possible interpretation as a null article. Null articles, though they feature in the Northern dialects, were consciously omitted from this discussion. It is possible, however, that a standard speaker, on encountering the reduced articles, would not regard them as articles at all, and, in turn, parse the speech stream without definite articles. The resulting input string would lack definite articles, resulting in syntactic sequences that are ill-formed in the standard. This would certainly be conspicuous and contribute to the salience of the variable in general.

Fourth, while I monopolised the concept of ‘transitional probability’ for the segmental level, and talked exclusively about transitions between segmental units, it certainly takes hold supra-segmentally as well. The use of the reduced variant creates different rhythmic patterns, which can lead to the increased salience of the variable. This remains true despite the fact that transitional probabilities, from Harris (1955) on, have been usually discussed in a segmental context. Then again, all of these factors are compatible with the segmental probability-based analysis explored above.

The fifth and main caveat, in my view, is that the salience of the variable can be derived from its status as a word boundary marker. One might argue that the role of TP-s is, in this case, indirect, since the primary source of listener attention is enhanced boundary marking. Indeed, this was my first suspicion when approaching article reduction.

Though boundary marking certainly adds to the variable’s salience, I believe its main source is still the difference in TP-s – in comparison with
a standard dialect, lacking definite article reduction. Since the issue of boundaries and transitional probabilities is intermingled in the DAR case, a possible refutation of the argument connecting salience to boundary marking would be to take an example where word boundaries clearly play a lesser role. In the following chapter, I turn to glottalisation in the South of England, which is such a phenomenon. Though the question of salience, as we will see, comes forward at a word boundary, this example is more about transitional probabilities, especially since the most stereotypical glottalisation environment, the word-medial pre-vocalic one, has nothing to do with word boundaries at all.
Chapter 5

Glottalisation in the South of England

The glottalisation of /t/ is a well-studied sociolinguistic pattern that occurs almost everywhere in England and Scotland. In this chapter I look at its behaviour with an emphasis on its sociolinguistic salience. The scope of this discussion is the South of England, and the glottalisation of /t/ word-finally. For the sake of simplicity, I ignore utterance-final environments (which, in any case, occur with a small frequency) and the glottalisation of other fortis stops.

Section 5.1 gives a brief overview of the development and distributions of the pattern in the South of England and argues for its sociolinguistic salience on the basis of external evidence. Section 5.2 describes the methods of the analysis, Section 5.2.4 gives the analysis itself, while Section 5.3 provides a short summary.

I use the term ‘/t/ glottalisation’ to refer to the glottal replacement, or loss of place, of the fortis coronal stop /t/ in the South of England. Glottalisation variably occurs in all environments with the exception of a following word-medial stressed vowel or before a vowel word-initially. The process is phonetically motivated pre-consonantly. The phonetic ground is the glottal reinforcement of stops in English English (Giegerich, 1992) –
in coda positions, the oral closure is reinforced by an accompanying glottal one. When we pair this up with the lack or release typical of pre-consonantal positions (Ladefoged & Maddieson, 1996), we end up with a stop that lost most of its place cues, such as the formant transitions into and out of the closure phase. The result is the audible gesture of glottal replacement – a glottal stop. Coronal stops, being shorter and more variable in duration, are more prone to fall victim to glottal replacement (Paradis & Prunet, 1991). Stops, however, can be subject to glottalisation before a vowel as well, a phenomenon both phonetically unmotivated and, as argued below, salient. (John Harris (p.c.) points out that the place cues of a pre-glottalised stop will be impaired even before a vowel, which can also lead to loss of place of articulation. What is certainly true is that this impairment is much smaller in extent than in the case of a pre-consonantal stop.)

While /t/ glottalisation is a phonetically gradient phenomenon, stretching from glottally reinforced coronal stops to complete loss of the coronal place of articulation (or ‘debuccalisation’), I will allow for the abstraction of regarding it as categorical, with word-final coronal stops realised either as coronal or as glottal. This is in line with the existing studies (discussed below), which also operate with this level of abstraction, and still allows for a valid conclusion on the pattern’s salience.

5.1 Background


There are a number of older and more recent surveys on the pattern in the South. An earlier example is Hudson & Holloway (1977), a study of London English conducted in 1977. The subjects were working class or middle class
and between 14 and 15 years of age. The results are that while Cockney, the traditional working class variant spoken in London, glottalised in all of the above environments (in 40 per cent of the cases in the intervocalic one), Popular London – spoken mainly by the middle classes in London – scarcely had signs of intervocalic debuccalisation, while the suburban accents had none at all. What is more, these accents shunned pre-syllabic liquid/nasal environments as well, only showing utterance-final and word-final pre-consonantal debuccalisation in 20 per cent of the cases. (Near-RP and other high registers only glottalised word-finally, before a consonant, as in [gɛ?redI] ‘get ready’.)

The differences between the social groups remain in Tollfree’s study (1999), recorded between 1990-94, but there is a spectacular difference in numbers. Tollfree worked with two age groups (15-30 and 54-89) and two accent classes, roughly corresponding to a middle class and a working class register. Unsurprisingly, the working class idiolect has more debuccalisation, while the intervocalic environment is still absent from the middle class variant. There is a great disparity between the 1977 and the 1999 study, however, in the case of word-final pre-vocalic and pre-pause environments. Tollfree’s numbers give five times more frequent debuccalisation in these environments in the speech of the younger age group than in the data of the Hudson & Holloway study.

The studies I am going to mainly rely on are Fabricius’ survey of upper-middle class speech (2000) and Altendorf’s work on London English (2003). These are the most recent comprehensive studies on /t/ glottalisation in England.

5.1.1 Two recent studies

Altendorf’s scope is London and the Home Counties. In London, she looks at three different school types, Comprehensive School, Grammar School, and Public School, which she roughly matches with working class, middle class, and upper-middle class language use. Her results show that /t/ glottalisation happens across-the-board in some environments, while it shows variation,
Glottalisation in the South of England

heavily related to speaker social class and register, in others. When it comes to the word-final environments, on one hand, we can clearly see that working class speech glottalises almost all final coronal stops irrespective of whether it is recorded in a formal or an informal setting, while upper-middle class speech avoids pre-vocalic glottalisation, especially in a formal setting. On the other hand, both classes glottalise almost every /t/ word-finally, before a consonant. (Cf. Tables 5.1 and 5.2 – the columns indicate Comprehensive School, Grammar School, and Public School, respectively.) Another pattern Altendorf notes is that glottalisation generally seems to be more extended in London than elsewhere in the South, which is expected if we regard London as a source of linguistic innovation for the area (and for England in general).

Fabricius, in her dissertation, concentrates on upper-middle class speakers chiefly from London and the South (the Home Counties) to investigate the extent to which /t/ glottalisation is accepted in the speech of a younger
generation of RP speakers. Her results corroborate the findings of Altendorf in two main aspects: first, /t/ glottalisation is more frequent in London than elsewhere in the South, second, it is common pre-consonantally (independent of the register), while it is strongly avoided pre-vocally, especially in a formal setting (cf. Tables 5.3, 5.4, and 5.5).

What both studies show is that /t/ glottalisation exhibits variation to the largest extent word-finally before a pause or a vowel, as well as word-medially. Variation seems to be sensitive on both the social background of the speaker (working class speakers realising more glottal stops than middle or especially upper-middle class speakers) and the register (all speakers realising more glottal stops in the interview setting than when reading a word list).

Altendorf’s results serve – among with the corpus data – as a basis for calculations in my analysis. As can be seen, her data are fairly congruous with Fabricius’ findings on upper-middle class speech. The extent of word-final
glottalisation in working class speech in her data are supported by previous studies and the observation of Paul Kerswill [p.c.] that /t/ glottalisation was not included as a variable in the analysis of the Spoken Corpus of Adolescent London English precisely because it was ubiquitous in all affected environments.

5.1.2 Salience and glottalisation

As seen in the previous chapter, from all the possible environments, word-final pre-vocalic /t/ glottalisation is subject to most variation. Three pieces of evidence suggest that it is salient for the language users. First, both the upper-middle class and the middle class speakers glottalise less in the _#V and _## environments in a formal setting; the feature shows style shifting. Though I restrict myself to the word-final environment (for now), all environ-
ments which display variation seem to show style shifting, which is strong
evidence for the salience of glottalisation in general. Second, glottalisation,
at least in certain positions, is subject to overt commentary. As Rosewarne
(1984) notes in his seminal article on Estuary English, especially the word-
medial glottalised realisations are regarded as ‘base’ or ‘erroneous’ by native
speakers.

The third piece of evidence comes from a discrimination study performed
by Fabricius (2000). Fabricius tried to explore the attitudes of upper-middle
class English speakers towards /t/ glottalisation. Her input set consisted of
24 test sentences, recorded with a word-final glottal/coronal pronunciation.
The nature of the following environment (___#C or ___#V) was one of the
variables. It was presented to 16 participants, 8 male, 8 female.

The interviewees were asked to judge each pair according to which pro-
nunciation they considered to be standard, good, correct. They were also told
that they could choose the first or second pronunciation, both, or reply that

Figure 5.4: Glottal replacement of /t/ in reading style – from Fabricius (2000)
they didn’t know. An ANOVA analysis of the responses reveals a strong bias on the environment, i.e. whether the following word starts with a consonant or a vowel (p< 0.0001). The upper-middle class participants strongly favoured the coronal pronunciation before a vowel, but significantly less so before a consonant.

Style shifting, stigmatisation, and negative listener attitudes all hint at #V glottalisation being salient for speakers of a conservative dialect (the one used by the upper-middle class speakers of the above studies). In contrast, we do not find style shifting, or pronounced or unpronounced listener attitudes concerning #C glottalisation. This, again, indicates that glottalisation is not salient in this environment.

This gives us two variables to compare, word-final pre-vocalic and word-final pre-consonantal /t/ glottalisation. My hypothesis is that salience is based on different probability distributions in the conservative and in the innovative dialect. The former will be modelled as spoken by the upper-middle
class, the latter by the working class. The prediction, then, is that we observe different probability distributions across dialects in the case of the former, but not the latter environment: the salience of the __#V realisation corresponds to a probability difference.

While the fact that /t/ glottalisation is overtly stigmatised in some cases is an argument for its status as a marker, salience is not necessarily connected to a threshold of consciousness. A variable, such as /t/ glottalisation, can, to all intents and purposes, behave like a marker without any extent of naïve linguistic awareness. I am noting this in a continuing effort to keep the concepts of ‘salience’ and ‘marker’ separate.

If we accept that __#V glottalisation is different, we have to ask which aspects make it different from __#C glottalisation. In both cases we are dealing with the same process (loss of place), which only leaves the segmental environment as a potential culprit. In the next section I propose an explanation for why the presence of a following vowel triggers a different perception of glottalisation.

5.2 Analysis

There is no recent study which would be sufficiently large and encompassing for the purposes of researching the salience of glottalisation. Consequently, my model is based on two sources, transcriptions of spoken corpora collected in England on one hand, and studies and descriptions of glottalisation in the South on the other. The corpora were used to get an approximation of upper-middle class and working class speech patterns. I relied on the studies to establish the extent of glottalisation, both because the corpora are not directly comparable, and because using the studies leads to statistically more robust generalisations.

The aim was to establish transitional probabilities (TP-s) for the glottal stop in word-final position, depending on speech register, following environment, and social class. The hypothesis is that the salience of word-final
pre-vocalic /t/ glottalisation (as opposed to the absence thereof in the word-final pre-consonantal environment) is directly related to the difference in TP-s between upper-middle class and working class use in the pre-V case, a difference not present in the pre-C case.

A crucial element of the argumentation is that word-final glottal stop variants is not related directly and exclusively to word-final TP-s. This argumentation would be, in any case, circular: word-final pre-vocalic glottalisation is almost absent in upper-middle class speech, while fairly common in working-class speech, no wonder it is salient in the latter – from the perspective of the former. The present line of reasoning goes one step beyond that: we start with the observation that glottalisation is present in both dialects, and go on to note that, according to the model discussed below, the word-final distributions affect its general distributions in the two dialects. It is precisely this difference in general distributions that leads to the variable’s salience.

Word-final pre-vocalic glottalisation is, then, salient exactly because pre-vocalic glottalisation is salient in general. This claim predicts that all pre-vocalic incidents of glottalisation should be noted by listeners – as we will see, this is indeed the case. Below, I look at the TP-s of glottal stops being followed by consonants and vowels in the word-final environment in particular, and in all environments in general. First, however, the transitional probabilities have to be established for the two assumed dialects, the upper-middle class and the working-class one.

### 5.2.1 Methods

In order to establish these probabilities, I used two corpora, the London-Lund Corpus of Spoken English for upper-middle class speech, and the Spoken Corpus of Adolescent London English for working class speech. Note that the two compared dialects need not be strictly tied to social class. One could regard the first dialect as the conservative, educated, non-local English accent – an approximation of the Received Pronunciation – and the second one as the local, innovative one, used by young people in London and in the South.
One issue with using the two corpora is that one is simply older than the other: the London-Lund corpus, as discussed below, contains material collected since the 1950s, while the Spoken Corpus of Adolescent London English was collected in a shorter period, in the mid-2000s. Given the alacrity and momentum of glottalisation in England, one ought to be very considerate in comparing corpora (and studies) from different periods.

This, in my view, is still less of a problem: first, the older corpus is based on conservative registers and more formal settings, which tend to change less in the first place, second, in terms of the environments relevant for glottalisation in the present study, the two corpora are fairly similar, despite the expected differences. This suggests that, when it comes to the word-final distribution of /t/, class differences are less pronounced in England. In any case, the fact that genuine samples of upper-middle class and lower class speech are available for comparison is more important than the differences in the time of the data collection.

Since both corpora are fairly large – the London-Lund corpus has a size of roughly 600,000 words, while the Spoken Corpus of Adolescent London English has a size of roughly 1,000,000 – and /t/ glottalisation is not indicated in them, they were transcribed automatically, and both available information on the corpora themselves and additional studies of /t/ glottalisation (most notably Altendorf 2003) were used to estimate its extent in the texts. (The corpora are described in detail below.)

While the size of the corpora would make transcription by hand very difficult, it gives very robust distributions of [t], so the TP-s of the [?] realisation can be estimated with high accuracy. In comparison, the use of a smaller corpus, like parts of the corpora above, or the South-East part of FRED (Kortmann et al., 2005), would allow for the precise marking of glottalised realisations, at the cost of losing the robustness of the sound patterns which serve as the base of the TP-s we are looking for. (Another argument against using FRED is its focus on NORM speakers, which removes the social angle.) In other words, this method admits a higher level of abstractness in return
for a larger dataset.

As the principle view of this chapter is that certain patterns of glottalisation in the South are salient for speakers of an educated, upper-middle class standard, the two corpora were chosen to represent the two target groups, the standard speakers and the Southern dialect speakers.

5.2.2 The London-Lund Corpus

The first corpus I used is the London-Lund Corpus of Spoken English (LLC) (Svartvik, 1990). This corpus derives from two research projects, the Survey of English Usage at University College London, launched in 1959 by Randolph Quirk, who was succeeded as Director in 1983 by Sidney Greenbaum, and the Survey of Spoken English, started by Jan Svartvik at Lund University in 1975. The aim of both projects was to provide resources for accurate descriptions of adult educated English usage. It is based on spoken language, mainly on conversations, but also on spontaneous and prepared monologues. The corpus contains roughly 600,000 words.

Large tracts of the corpus are conversations between academics, characterised by short utterances and frequent interruptions. As an illustration, here is the first ten lines of the corpus in its original txt format:

```
1 1 1 10 1 1 B 11 ((of `Spanish)) . graph\ology# /
1 1 1 20 1 1 A 11 `w=ell# . /
1 1 1 30 1 1 A 11 ((if)) did `y/ou _set __that# - /
1 1 1 40 1 1 B 11 `well !J\oe and __I# /
1 1 1 50 1 1 B 11 `set it betw\een __us# /
1 1 1 60 1 1 B 11 `actually !Joe `set the :p\aper# /
1 1 1 70 1 1 B 20 and *((3 to 4 sylls))* /
1 1 1 80 1 1 A 11 *__w=ell# . /
1 1 1 90 1 1 A 11 "^m/\ay* I __ask# /
1 1 1 100 1 1 A 11 `what goes !\into that paper n/ow#
```
The various symbols in the text are marking intonational phrases and intonation patterns. These bear no relevance to my analysis, so they were removed, along with the markers indicating which text of the corpus we are looking at and speaker identification (the first eight columns above). The corpus has other markers which were removed during the machine transcription. The only grapheme preserved is the line-medial dot, indicating the end of a larger utterance.

The texts were transcribed replacing words with their citation forms in Celex (Baayen et al., 1993) (for details, cf. Chapter 3). The extent of glottalisation was established using independent measures, namely, the available studies on glottalisation (more on this below). While other possible deviations from the standard are also absent from the transcription, the effect of these is assumed to be minimal. This is especially true for the LLC, as it displays standard upper class speech. The resulting text looks like this (for the ten-line excerpt above):

'Qv 'sp{nIS . 'wEl 'If 'dId 'ju 'sEt 'D{t 'wEl '2 'sEt 'It bI’twin 'Vs '{kJ9lI 'Di 'p1p@R '{nd 'tu 'wEl 'm1 '2 '#sk 'wQt 'g5z 'Intu 'D{t 'p1p@R 'n6

As we can see, the transcription is not perfect, for instance, it misses the proper name ‘Joe’. It also removes line breaks, but since we are not interested in the utterance-final environment, and since this environment is not that frequent compared to other ones featuring [t], this is less of a problem. Citation forms have stress marked on them. Clearly, not all words are equally stressed in fluent speech. However, word stress limits the extent of /t/ glottalisation (such as that /t/ is never glottalised before stressed vowels), so it is very useful as additional information. (A separate chapter discusses all the basic details of this corpus based method.)

The raw text of the corpus has 733,000 ‘words’ (including various annotation markers), of which 407,000 were successfully transcribed in the above
fashion. This transcribed corpus is the basis for the model of upper-middle class speech. The distributions of [t]-s are given in Table 5.1.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-consonantal</td>
<td>13,500</td>
</tr>
<tr>
<td>Word-final pre-consontantal</td>
<td>25,000</td>
</tr>
<tr>
<td>Word-final pre-vocalic</td>
<td>11,875</td>
</tr>
<tr>
<td>Utterance-final</td>
<td>1400</td>
</tr>
<tr>
<td>Before syllabic sonorant</td>
<td>1400</td>
</tr>
<tr>
<td>Intervocalic, before unstressed vowel</td>
<td>5700</td>
</tr>
<tr>
<td>Intervocalic, before stressed vowel</td>
<td>800</td>
</tr>
<tr>
<td>Other</td>
<td>37,925</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>97,600</strong></td>
</tr>
</tbody>
</table>

Table 5.1: Transcription t-counts in the LLC

Table 5.2 shows the distributions of coronal and glottal stops in these environments, based on the formal (reading) data of Altendorf for London public school students. Note that, as opposed to the previous Table, these are estimated counts. The numbers are extrapolated from the Altendorf study, the ratios of which were applied to this corpus. What this Table shows, then, is an approximation on the extent of glottalisation in a conservative dialect of Southern English. The two sources of this approximation are a corpus of upper class language use and a study on upper class glottalisation – which study is corroborated by others of the similar kind, such as Fabricius (2000). These are the estimated that will allow establishing the transitional probabilities of the glottal stop in various positions – to reach its suspected salience as a variant in these positions.

Bearing in mind the results here, we can turn to the second corpus, which is used to model innovative, working class speech patterns.
## Analysis

<table>
<thead>
<tr>
<th>Environment</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>[t]</td>
<td>13,500</td>
</tr>
<tr>
<td>[?]</td>
<td>0</td>
</tr>
<tr>
<td>Pre-consonantal</td>
<td>0</td>
</tr>
<tr>
<td>Word-final pre-consontantal</td>
<td>2250</td>
</tr>
<tr>
<td>Word-final pre-vocalic</td>
<td>11,250</td>
</tr>
<tr>
<td>Utterance-final</td>
<td>1400</td>
</tr>
<tr>
<td>Before syllabic sonorant</td>
<td>1400</td>
</tr>
<tr>
<td>Intervocalic, before unstressed vowel</td>
<td>5700</td>
</tr>
<tr>
<td>Intervocalic, before stressed vowel</td>
<td>800</td>
</tr>
<tr>
<td>Other</td>
<td>37,925</td>
</tr>
<tr>
<td>Total</td>
<td>59,725</td>
</tr>
</tbody>
</table>

Table 5.2: Coronal and glottal stop counts based on the LLC

### 5.2.3 The Spoken Corpus of Adolescent London English

The second corpus is the Spoken Corpus of Adolescent London English (SCALL) collected by Paul Kerswill, Jenny Cheshire, and their colleagues (Cheshire et al., 2008). The corpus contains interviews with adolescents (age 17-19) sampled from two London boroughs, between 2004 and 2007. The aim was to get a coherent picture of the linguistic innovation present in adolescent speech in London, hence the focus on young working class speakers. The corpus contains roughly 1,015,000 words, of which 1,005,000 were successfully transcribed. Transcription followed a course similar to the case of the LLC. One notable difference is that the SCALL focusses on working class speech. It which might substantially differ from the citation forms used in the transcription process. These citation forms, taken from Celex, reflect RP. However, the effects of these differences are assumed to be minimal.

To take one example, vowel qualities will be different in popular London speech, but that is outside this discussion. To take another example, some differences in the consonant system might affect the loci of interest – word-final pre-vocalic and pre-consonantal positions. In particular, H-dropping could
change the ratio of consonant-initial words, thereby affecting the distributions, but this is a small scale problem: in the London English corpus, 305 tokens have an initial [h] – out of 1,005,000.

The SCALL has a different format from that of the LCC, consisting of standard sociolinguistic interviews with one or more participants, transcribed orthographically. Interjections, pauses, and hesitations are marked. Here is an example from the beginning of the raw material:

S(...): OK that should be it yeah erm I came over the week before last I think it was [A(...): yeah] and you’d just gone out of the room for for something [A(...): can’t remember] yeah no I don’t know where you were so but erm

A(...): how long was I gone for?

S(...): I don’t know cos I only popped in there [A(...): ah] to find people and when I couldn’t find them I just came out again so [A(...): oh] I was in there for about ten minutes or so <sniff> but er <throat clearing> yeah erm the two that I spoke to before erm they..they were in the middle of assessment everybody 1 seems to be

After the removal of all the annotations and the machine transcription, the resulting text looks like this (S(...’s first utterance):

"5’k1 'D{t 'SUd 'bi 'It 'j8 '3m '2 'k1m '5v@R 'Di 'wik bI’f$R 'l#st '2 'TINk 'It 'wQz 'j8 'n5 '2 'd5nt 'n5 'w8R 'ju 'w3R 's5 'bVt '3m

A few errors are obvious here as well. For example, the part between A(...’s two interjections got lost along the way. Still, as noted above, the vast
The majority of the corpus was transcribed successfully. The distributions of word-final [t]-s are given in Table 5.3.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-consonantal</td>
<td>29,900</td>
</tr>
<tr>
<td>Word-final pre-consontantal</td>
<td>86,600</td>
</tr>
<tr>
<td>Word-final pre-vocalic</td>
<td>36,000</td>
</tr>
<tr>
<td>Utterance-final</td>
<td>15,500</td>
</tr>
<tr>
<td>Before syllabic sonorant</td>
<td>2,400</td>
</tr>
<tr>
<td>Intervocalic, before unstressed vowel</td>
<td>12,700</td>
</tr>
<tr>
<td>Intervocalic, before stressed vowel</td>
<td>1,200</td>
</tr>
<tr>
<td>Other</td>
<td>56,400</td>
</tr>
</tbody>
</table>

**Table 5.3: Transcription t-counts in the SCALL**

Table 5.4 shows the estimation based on the corpus and Altendorf’s reading data for Comprehensive School. Note that in this case, we observe a much larger extent of glottalisation than in the Table based on the LLC. This Table shows an approximation of glottalisation in the innovative, working class dialect of Southern English.

Now that we have the basic approximations for both the conservative and the innovative dialect, we can continue by looking at the transitional probabilities for these dialects.

### 5.2.4 Modelling results

The hypothesis of salience put forward here assumes that a certain variable is noticeable (even if unconsciously) not because of the sound segments involved, but rather due to the difference in their distributions across two dialects. In the present case, innovative /t/ glottalisation is salient for a conservative dialect speaker not as a consequence of the fact that the coronal stop has a glottal alternate – it does so in the conservative dialect as well. It is the
In this section I expand on the dialect models of /t/ glottalisation in various ways in order to compare estimated transitional probabilities in various environments.

The transitional probabilities of glottal stops and following consonant and vowel-initial words in the conservative and innovative dialects are given in Table 5.5. These numbers are based on the models discussed in the previous section. As we can see, glottal stops are more expected in both environments in the innovative dialect. This is not really surprising given the fact that this dialect has a larger general extent of glottalisation. What is more noteworthy, however, is the extent of the difference in TP-s between the conservative and the innovative dialects. Pre-consonantally, the conservative TP is slightly higher. Pre-vocalically, the innovative TP is higher, and here, the difference is much wider, one order of magnitude.

What Table 5.5 shows, then, is a difference of TP-s in the two word-final environments. Pre-consonantally, the [?] TP-s are slightly different. Pre-vocally, however, the difference is one order of magnitude between the two dialects, and the TP of [?] in the innovative dialect is much larger.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Count</th>
<th>[t]</th>
<th>[?]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-consonantal</td>
<td>0</td>
<td>29,900</td>
<td></td>
</tr>
<tr>
<td>Word-final pre-consonantal</td>
<td>0</td>
<td>86,600</td>
<td></td>
</tr>
<tr>
<td>Word-final pre-vocalic</td>
<td>4,513</td>
<td>31,487</td>
<td></td>
</tr>
<tr>
<td>Utterance-final</td>
<td>700</td>
<td>14,800</td>
<td></td>
</tr>
<tr>
<td>Before syllabic sonorant</td>
<td>0</td>
<td>2,400</td>
<td></td>
</tr>
<tr>
<td>Intervocalic, before unstressed vowel</td>
<td>9,271</td>
<td>3,429</td>
<td></td>
</tr>
<tr>
<td>Intervocalic, before stressed vowel</td>
<td>1,200</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>56,400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>61,100</td>
<td>178,200</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4: Coronal and glottal stop counts based on the SCALL contrast in the glottal alternant’s distributions that makes a difference.
This is where TP-s and salience meet: since the pre-V distributions of the glottal stop are so different in the two dialects, this difference becomes a salient marker of innovative versus conservative dialect usage, which is reflected by avoidance in formal settings, social indexation, and even overt commentary. Pre-consonantally, this split is much smaller, therefore we do not expect salient behaviour from this realisation of the variable (cf. section 5.1.2).

The higher TP of a consonant following a glottal stop in the conservative dialect might seem counter-intuitive first. It is, however, entirely logical. TP-s are based on the ratio of glottal stops in this position versus glottal stops in other positions, not on the total number of glottal stops. Glottal stop realisations pre-consonantally have a higher probability in the conservative dialect than in the innovative one, since, while the general extent of glottalisation is smaller in the former, the extent of pre-consonantal glottalisation is roughly the same. Therefore, a glottal stop is more likely pre-consonantally in the conservative dialect, irrespective of the ratio of the total number of glottal stop realisations between the two dialects. (To put the findings in context, the overall probability of an orthographic $<t>$ to be realised as a glottal stop is 0.388 in the conservative dialect and 0.744 in the innovative one.)

The salience of $/t/$ glottalisation can be evaluated in a larger context. So far, we looked at the word-final pre-vocalic environment. The investigation can be extended to all pre-vocalic environments, including word-medial ones. This is a useful step, as such a larger context involves a stereotypical, highly stigmatised environment of $/t/$ glottalisation, the glottalisation of $/t/$ word-medially, preceding an unstressed vowel. Examples of this environment are

<table>
<thead>
<tr>
<th>Environment</th>
<th>Conservative</th>
<th>Innovative</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(#)C</td>
<td>$</td>
<td>0.627</td>
</tr>
<tr>
<td>$p(#)V</td>
<td>$</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Table 5.5: Word-final [?] TP-s (formal)
Glottalisation in the South of England

[ˈsiʔi], [ˈleʔə], [ˈbaʔə] (‘city, letter, butter’).

In order to calculate TP-s for this environment, one has to take into account the estimated glottal realisations word-medially, before stressed and unstressed vowels. For the former environment, there are no estimated glottal realisations based on the studies. This is not surprising, since glottalisation is widely reported as absent in this position, so much so that surveys usually do not include it as an environment for the variable.

As seen in section 5.1, /t/ glottalisation is absent in this position in the conservative dialect, based on upper class speech in a formal setting. It is, however, present in the innovative one, based on working class speech in a formal setting. Going through the same procedures as in the case of the word-final environments, we get the following TP-s for vowels following glottal stops in general, without respect to word boundaries:

- $p(V|?)$ in the conservative dialect: 0.017
- $p(V|?)$ in the innovative dialect: 0.254

Again, we find a large difference between the TP-s of the sequence in the two dialects, and it is much more likely to occur in the innovative dialect. This result, again, establishes a connection between the probability distributions of the realisation of the variable and its salience: word-medial pre-vocalic glottalisation is much less frequent in the conservative dialect, hence the conservative dialect speaker attitudes towards its innovative dialect use.

The case of word-final versus all pre-vocalic environments is relevant in two ways. First, it shows that the proposed method of TP calculation works for other environments as well, and is not an arbitrary way of singling out positions which would support the hypothesis. Second, it gives an empirical ground to the salience of word-medial pre-vocalic /t/ glottalisation, another salient realisation of the variable.

Formal interview settings served as the basis of calculations throughout this chapter. If we regard Altendorf’s data collected in the formal and the
informal setting as an apparent time study, we can make predictions on the change of salience during the variable’s spreading.

Apparent time studies usually sample data from different age groups and assume that the differences between the age groups reflect the differences between the generations (Labov, 1994) in the sense that the way 60 year-olds speak in the present is the way 30 year-olds spoke 30 years ago. The same assumption can be made on the differences between the formal and informal registers. The sources quoted in section 5.1 give a picture of /t/ glottalisation as a change coming from below, affecting more and more environments to a larger extent, while always more avoided by upper-class speakers and in formal registers. Consequently, one might argue that its distributions in the formal register in the present reflect the way it was used informally a few decades ago.

Indeed, this is precisely what we find comparing Hudson & Holloway (1977), Tollfree (1999), and Altendorf (2003): Whereas in the first study, glottalisation is almost completely absent from the informal use of the upper middle classes (except before a consonant), and its use, in comparison with later studies, is moderate in Cockney, there is a sharp increase in the latter two studies. The end result can be seen in Altendorf’s data: upper middle class speakers already glottalise pre-vocally, and the working class speakers perform almost categorical glottalisation in the affected environments.

Building on these premises, we can compare TP-s extracted from the data of the informal setting with the data of the formal setting. With the procedure described above, we arrive at the TP-s for the informal setting shown in Table 5.6. The table gives the original TP-s based on the formal setting on the left. The conservative pre-vocalic TP-s are marked in bold.

What we can see here is a radical shift in the TP of the pre-vocalic realisation. The tenfold difference based on the formal setting wanes in the data based on the informal one. This suggests that if a variable’s realisation reaches similar distributions in two dialects, it will no longer be salient across these dialects, since the TP difference underlying its salience vanishes. (Based
Glottalisation in the South of England

<table>
<thead>
<tr>
<th>Environment</th>
<th>Formal setting</th>
<th>Informal setting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conservative</td>
<td>Innovative</td>
</tr>
<tr>
<td></td>
<td>p(#C</td>
<td>?)</td>
</tr>
<tr>
<td></td>
<td>p(#V</td>
<td>?)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.572</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.103</td>
</tr>
</tbody>
</table>

Table 5.6: Word-final [?] TP-s (formal & informal)

on the data collected in the informal setting, the overall probability of an orthographic t to be realised as a glottal stop is 0.448 in the conservative dialect, and 0.752 in the innovative one.

A complicating factor, however, is that a salient marker can become stigmatised, that is, subject to overt attention. Under such circumstances, its salience will not only depend on its distributions, but also on its existence as an independent conscious object for the language users. Needless to say, further factors can affect the salience of a variable’s realisation, such as phonetic prominence, the social status of dialect users, and so on.

This section gave an operationalisation of the salience of glottalisation realisations based on the corpus data and the surveys discussed in the previous one. The three environments discussed were word-final pre-C and pre-V, as well as pre-V (word-medially and word-finally). A sharp distinction was found between the TP of the realisation pre-vocically and pre-consonantally, which distinction correlates with the realisations’ respective salience.

5.3 Concluding remarks

This section used a narrow perspective to look at a dialectal variable, /t/ glottalisation in the South of England. Concentrating on the pattern’s distribution and salience in the word-final position, I argued that the salience of pre-vocalic glottalisation is due to the different distributions in the conservative and innovative dialects. The point of comparison was the pre-consonantal position, where the pattern seems to lack salience. Salience and distributions
Concluding remarks

were established based on previous studies on the subject, including an attitude test by Fabricius, and a larger-scale study by Altendorf, as well as based on corpora of upper-middle class and working class speech in the South.

Using corpus data along with results of previous studies, I reached the conclusion that there is a large frequency difference in the salient environment, but not in the non-salient one. This shows a correlation between the pattern’s transitional probabilities and the pattern’s salience. We can assume, then, that, at least in the word-final position, there is a relationship between probability distributions and the social evaluation of the variable. Acknowledging the possible effects of the variable’s stigmatisation, we can predict that the same is true for other environments, and that the spread of glottalisation eventually eradicated the difference in distributions, and, consequently, the pattern’s salience. Indeed, this is partly supported by the inexorable spread of glottalisation with respect to environments, registers, and social classes, as can be witnessed in studies on the phenomenon.

There is one question left unaddressed. I demonstrated that there is a relationship between the salience and the different distributions of the variable in working class and upper-middle class speech. I avoided, however, the discussion on the origins of these different distributions, that is, why upper-middle class speakers seem to glottalise to a much lesser extent word-finally before a vowel than their working class counterparts, while both speaker groups seem to realise the variable to a similar extent in the same position before a consonant.

This issue is inexorably related to salience itself. Speakers aiming at a higher register or style will avoid pre-vocalic glottalisation, even if inadvertently, because it is a lower prestige marker. Then again, such a statement is not very meaningful, as it only amounts to an observation of inherent circularity: certain speakers avoid glottalisation in this position, hence we arrive at different distributions, hence it is salient, hence speakers avoid it.

Another factor in the emergence of the difference in distributions, one, that is – luckily for us – independent from the social evaluation of the variable, is
the phonology of glottalisation. Pre-consonantal glottalisation is phonetically motivated, as the loss of place of the coronal stop can be a natural consequence of its flanking by the following consonant. In that sense, pre-C glottalisation, if not expected, is hardly surprising. Pre-V glottalisation, however, is different. Its rise can not be motivated that easily by the pure phonetics, and is likely influenced by the given presence of pre-C in the language. This is, to an extent, true for word-final pre-V glottalisation, and can be even more easily observed in the case of word-medial intervocalic glottalisation (the [lɛʔo] ‘letter’ type), where the phonetically motivated lenition pattern is flapping, not glottalisation. Therefore, the different distributions in the two word-final positions have something to do with the way glottalised alternants came into existence in these positions in the first place. Differences between the perception of various glottalisation patterns in English are noted in Chapter 8.
Chapter 6

Hiatus resolution in Hungarian

All examples of the link between salience and probability distributions throughout this book are taken from dialects of English spoken in the British Isles. Apart from reflecting the author’s expertise, this preference is also supported by the amount of sociolinguistic data available on variables in these dialects in general, and on the salience and social evaluation of these variables in particular. Then again, the theory of salience proposed here is by no means language specific: it predicts a general correlation between the probability of occurrence of the different realisations of a variable and its possible use for social indexation, based on general properties of attention deployment. A way of providing proof for this is looking at a similar problem in a language completely unrelated to English, Hungarian.

6.1 Background

The phenomenon under scrutiny, hiatus resolution in Hungarian, is relatively poorly researched from the sociolinguistic perspective. Essentially, Educated Colloquial Hungarian (ECH), the standard dialect spoken in Budapest and the major cities, has two distinct post-lexical hiatus resolution patterns. The first sort is obligatory and non-salient in the standard, while the second is smaller in extent and subject to variation, as well as distinctly salient
Hiatus resolution in Hungarian

(Siptár & Törkenczy, 2007). The aim of this chapter is to find a connection between the salience of the latter pattern and its probability of occurrence. As the evidence for the pattern’s salience was collected during fieldwork, a considerable amount of naïve linguistic judgements are also available, and these can shed light on the social evaluation of this particular linguistic variable.

This section introduces hiatus resolution in ECH. Section 6.1.1 discusses the results of a study on the social evaluation of hiatus resolution. Section 6.2 pairs up these results with frequencies gained from corpus data. Section 6.3 gives a concluding discussion.

The phonetics and phonology of hiatus resolution has been extensively covered (Kálmán & Rebrus, 2010; Siptár & Törkenczy, 2007; Siptár, 2003), but its social evaluation has been scarcely discussed in any depth. Siptár & Törkenczy (2007), whose description I mainly rely on, only mention the issue in passim. The basic state of affairs is as follows: On one hand, Hungarian has lexical hiatus resolution. Certain suffixes are realised with an initial consonant if the stem ends in a vowel. Two patterns, one with [j] and another one with [v], are illustrated in Table 6.1. The intrusive segment consistently occurs if the stem ends in a vowel.

<table>
<thead>
<tr>
<th>asztal + (j)a</th>
<th>asztala</th>
<th>‘table-3SG-POSS’</th>
</tr>
</thead>
<tbody>
<tr>
<td>kapu + (j)a</td>
<td>kapuja</td>
<td>‘gate-3SG-POSS’</td>
</tr>
<tr>
<td>asztal + (v)al</td>
<td>asztallal</td>
<td>‘table-INST’</td>
</tr>
<tr>
<td>kapu + (v)al</td>
<td>kapuval</td>
<td>‘gate-INST’</td>
</tr>
</tbody>
</table>

Table 6.1: Lexical hiatus resolution in Hungarian

On the other hand, post-lexical hiatus filling is also present in Hungarian – the scope of this chapter (cf. Table 6.2). It occurs obligatorily in vowel clusters containing [i] and it is quite common in clusters containing [e:]. The inserted segment is the glide [j]. These are the two close front vowels of Hungarian. The close [i], like all Hungarian vowels, has a long pair [iː]. However, in ECH, the realisation of close vowels is subject to variation, and they generally show
a tendency to be shortened in all positions except the initial syllable. This bears no importance on the present discussion. The short pair of [ɛ:] differs not only quantitatively but also qualitatively, as it is realised as an open mid [ɛ].

Vowel clusters containing [ɛ] but not [i] or [ɛ:] also show hiatus resolution, it is, however, realised less frequently, and, in any case, is subject to variation. There are no data on the extent of hiatus resolution in the three environments, namely, in clusters involving [i], [ɛ:], and [ɛ], respectively, but it is commonly assumed that, in Conservative ECH, hiatus resolution is obligatory in the first, variable in the second, and absent in the third environment. However, as Siptár and Törkenczy note, Innovative ECH does have hiatus resolution in the third environment, although to a limited extent.

<table>
<thead>
<tr>
<th>fiú</th>
<th>[fiju:]</th>
<th>‘boy’</th>
</tr>
</thead>
<tbody>
<tr>
<td>női</td>
<td>[nøːji]</td>
<td>‘female’</td>
</tr>
<tr>
<td>ráér</td>
<td>[%rajɛɾ]</td>
<td>‘to be at leisure’</td>
</tr>
<tr>
<td>büféasztal</td>
<td>[%byfeʃnʃtɔl]</td>
<td>‘buffet table’</td>
</tr>
<tr>
<td>tea</td>
<td>[%tɛʃɛp]</td>
<td>‘tea’</td>
</tr>
<tr>
<td>beakad</td>
<td>[%beʃkød]</td>
<td>‘gets stuck’</td>
</tr>
</tbody>
</table>

Table 6.2: Post-lexical hiatus resolution in Hungarian

I will argue that, compared to the first two environments, hiatus resolution in the third one is salient for the speakers of ECH. That is, while neither of the hiatus resolution processes in Conservative ECH create salient patterns, the sequences resulting from Innovative ECH hiatus resolution are noticeable for language users. We will see that a simple phonetic explanation is insufficient for the salience of [ɛ] plus [j] patterns, because these are not noticed in any way unless when resulting from hiatus resolution.

I am going to underpin my argument using the results of an attitude test on the perception of this variable. The test provides empirical evidence on the pattern’s salience. It is also a valuable resource of naïve linguistic judgements and descriptions of the variable. In order to allow the reader to interpret
some of these comments in a larger context, we have to briefly overview other phenomena in Hungarian which involve j-insertion.

One phenomenon that is directly related to hiatus resolution is that the intensity and duration of the inserted glide can vary. Though again there are no exact data on this variation, native speakers, to an extent, share a sentiment that an extended glide inserted into a cluster where hiatus resolution is obligatory is a strong substandard feature. However, it is likely that there is no substantial phonetic difference between inserted [j]-s and ‘phonemic’ [j]-s, a point to which we will return below.

A more tangible pattern is displayed by the possessive. As noted above, possessive allomorphy avoids hiatuses: if the stem ends in a vowel, the suffix is -je or -ja. If the stem ends in a sibilant/palatal consonant, however, the occurrence of the glide is subject to variation. This variation, coupled with the fact that the glide-initial allomorph is the productive one, added to recent loanwords, leads to hesitations in possessive selection, as well as to a general sentiment that there is something going on with [j] in the possessive (Rácz & Rebrus, 2012).

A third relevant variable is the realisation of the definite article. In standard ECH, it has two forms, a, az, depending on whether the following word starts with a consonant. In certain Transdanubian dialects, definite article allomorphy is also sensitive to whether the previous word ends with a consonant. If it ends with a vowel, the article has an initial glide. This is shown in Table 6.1.

<table>
<thead>
<tr>
<th>Standard</th>
<th>Dialectal</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>V#_#V</td>
<td>adja az ég</td>
<td>adja jaz ég</td>
</tr>
<tr>
<td></td>
<td>számolja a pénzt</td>
<td>számolja ja pénzt</td>
</tr>
<tr>
<td></td>
<td>lement a nap</td>
<td>lement a nap</td>
</tr>
<tr>
<td></td>
<td>finom az alma</td>
<td>finom az alma</td>
</tr>
</tbody>
</table>

Table 6.3: Definite article allomorphy
As we can see, these processes all involve [j], but are otherwise entirely different: they relate to phonetic realisation of the hiatus filler, the selection of a suffix, and the realisation of the definite article. For the time being, we will go back to the salient and non-salient forms of hiatus filling, discussed in this section above.

To recap things briefly, in conservative ECH, hiatus resolution is obligatory in vowel clusters including [i], and it is variable in clusters with [e:]. In innovative dialects, it is also possible in clusters with [e]. I claimed that the first two hiatus resolution patterns are not salient for ECH speakers, while the latter is. In order to evaluate this claim, I performed an attitude study with native speakers of ECH.

6.1.1 The perception of hiatus resolution: Methods

The test included ten ECH speakers, five female, five male, with a mean age of 22. Eight were from Budapest, and two from the surrounding Pest country. The participants listened to a recording of 30 sentence pairs, 10 with vowel plus [e] clusters, 10 with [e] plus vowel clusters, and ten control sentences, featuring V[i]/[i]V and V[e]/[e]V equally. The recordings were mono audio waveform files, sampled at 44100Hz. The participants were not paid for the experiment.

The test sentences were read by a trained phonetician, also a native speaker of ECH, once with a hiatus filler [j], once without one. The control sentences were all read with a realised hiatus filler. The background information given to participants was that a Hungarian male in his twenties is looking for their help in general linguistic and stylistic issues, as he is going to a job interview in Budapest and is unsure about the quality of his Hungarian. The participants had to evaluate the sentence pairs (with the implication that the sentences are different) on a Likert-scale from 1 to 10, depending on whether they found the first or the second sentence better (or they were unsure, etc.). The participants listened to the pairs in a random order, both in the sense that the order of pairs was randomised and that the order of the marked
sentence (the one with hiatus resolution) and the unmarked sentence (the one without it) was randomised: half the pairs had the marked sentence first. The listening test was followed by a small discussion with the participants.

The experiment has two conditions: (i) whether the marked sentence comes first or second and (ii) whether the judgements on the test sentences differ from judgements on the control sentences. The hypothesis is that hiatus resolution with [ɛ] is a salient variable that will be rejected, whereas hiatus resolution with [i] and [ɛ:] elicits no listener attitudes. This should show up in condition (i) as a larger score on the scale if the first sentence is marked in the pair and vice versa, and in (ii) as a score more divergent from the mean in the case of test sentences versus control sentences, as participants are not expected to show explicit preference for any sentence in a pair of control sentences.

6.1.2 The perception of hiatus resolution: Results

The results show a strong preference for the unmarked pattern in condition (i) and more divergent scores in condition (ii), which confirms the hypothesis that hiatus resolution in [ɛ]+V clusters is rejected, hence, salient for ECH speakers.

The results were weighted between participants. For condition (i), the resulting scores were modified in such a way that a higher score means a preference for the marked pattern. Condition (ii) is needed in the first place because condition (i) relies on the order of sentences within the pairs. Therefore, the results can be influenced not only by which sentence was marked, but also by the order itself: if the first sentence is marked, chances are, people become more aware of it. Since condition (ii) compares all the test sentences with the control sentences, the problem of ordering disappears.

The scores are shown in Figure 6.1 for the first condition and Figure 6.2 for the second one. In the Figure 6.1, the first column is the control, the second is where the marked sentence came first in the pair, and the third is when it came second. Higher scores indicate a stronger overall preference for
the second sentence. As can be seen, if the first sentence is marked, scores are higher, and if the second sentence is marked, they are lower than in the control case.

Figure 6.2 compares control and test sentences. Higher scores generally indicate a stronger overall preference for the marked sentence. Of course, the control pairs did not have a marked sentence. Scores are again weighted, which resulted in a higher score than the expected 0 for the control case. What is visible, however, is the existence of a deviation from the mean in the test case, which is absent in the control case. That is, while there was a preference for one member in the pair in the test case, this preference was absent in the control case.

The significant difference between answer rates, visible on the plots, is supported by a fitted linear mixed model for both conditions, with speakers (subject) and read sentence pairs (sentence.id) as a random effect. The lme4 package (Bates, 2005), implemented in R (R Development Core Team, 2009), was used for the mixed-effects modelling. The summary of the model for condition (i) is in (1) and the summary of the model for condition (ii) is in (2). We can see that in condition (i), whether the first or the second sentence was marked in the pair (2nd_markedno/yes) is a significant predictor of the weighted score (score), with a strong correlation of the fixed effects. In condition (ii), whether the sentence was a condition or a test sentence (conditiontest) is again a strong predictor of the weighted score (alt.score), with an observable strong correlation again.

(1) Summary of the model for condition (i)

Linear mixed model fit by REML
Formula: score ~ 2nd_marked + (1 | subject) + (1 | sentence.id)
Data: dat
   AIC   BIC  logLik deviance REMLdev
749.9 772.1  -368.9  730.1  737.9
Random effects:
Figure 6.1: Weighted scores for condition (i)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Name</th>
<th>Variance</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>sentence.id</td>
<td>(Intercept)</td>
<td>0.054274</td>
<td>0.23297</td>
</tr>
<tr>
<td>subject</td>
<td>(Intercept)</td>
<td>0.000000</td>
<td>0.00000</td>
</tr>
<tr>
<td>Residual</td>
<td></td>
<td>0.633746</td>
<td>0.79608</td>
</tr>
</tbody>
</table>

Number of obs: 300, groups: sentence.id, 30; subject, 10

Fixed effects:

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.7434</td>
<td>0.1085</td>
</tr>
<tr>
<td>2nd_markedno</td>
<td>-0.9542</td>
<td>0.1576</td>
</tr>
<tr>
<td>2nd_markedyes</td>
<td>-1.2466</td>
<td>0.1499</td>
</tr>
</tbody>
</table>

Correlation of Fixed Effects:
Figure 6.2: Weighted scores for condition (ii)

(Intr) X2nd_s_mrkdn
2nd_mrkdn -0.688
2nd_mrkdy -0.724 0.498

(2) Summary of the model for condition (ii)

Linear mixed model fit by REML
Formula: alt.score ~ condition + (1 | subject) + (1 | sentence.id)
Data: dat

AIC BIC logLik deviance REMLdev
749.5 768 -369.7 733.7 739.5

Random effects:
Groups     Name          Variance  Std.Dev.
sentence.id (Intercept) 0.065188 0.25532
subject     (Intercept) 0.000000 0.00000
Residual           0.633746 0.79608
Number of obs: 300, groups: sentence.id, 30; subject, 10

Fixed effects:

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.7434</td>
<td>0.1134</td>
</tr>
</tbody>
</table>
Hiatus resolution in Hungarian

conditiontest -1.1151  0.1389  -8.030

Correlation of Fixed Effects:
  (Intr)
conditintst -0.816

In sum, the results confirm speaker awareness of the Innovative hiatus resolution pattern *vis-à-vis* the Conservative pattern. For a Hungarian linguist, this is hardly surprising, as the pattern is overtly discussed, and some forms like *teja* are used playfully by speakers who otherwise eschew Innovative hiatus resolution. It is, however, important to stress that, until now, the pattern’s social evaluation has not been empirically tested.

6.1.3 Hiatus resolution and naïve linguistic awareness

Each listening session was followed by a short discussion where the participant was asked to guess what the experiment was about. The comments gained during these sessions cannot be assessed quantitatively, but shed light on the subjective evaluation of the variable.

First of all, most participants picked up on the fact that there is something going on with [j]. Contrary to expectations, not all of them were able to pinpoint the main difference within the pairs: two participants put emphasis on the intonation differences they perceived alongside [j], even though the differences in intonation between the sentences was negligible. One participant did not name [j] at all, even though the participant also showed significant preference for the [j]-less realisation of vowel clusters containing [ɛ].

Participants who did mention [j] as a part of the perceived difference varied in their precision in judging the nature of the difference. Some located the salient hiatuses, and one explicitly mentioned [tejo] ‘tea’ as an example. Others gave instances of non-salient hiatus resolution as an example for the perceived difference, but increased the duration and intensity of the intrusive glide. One participant mentioned the non-standard realisation of the definite article, with an initial [j] following V-final words, and another referred to
possessive variation, also including [j] (these processes are discussed in section 6.1).

The comments show that language users are unable to pin down the variable, despite its salience. Even if they themselves notice the difference, they are not capable of reporting on it. It is generally (though not always) recognised that [j] is the culprit, but the comments invoke other, unrelated phonological, morpho-phonological, and morphological phenomena, including the non-salient hiatus resolution pattern. The reason for the latter, I suspect, is the orthography. Intrusive [j]-s are not indicated in the orthography, whereas, for instance, the (post-vocally) also predictable possessive suffix is.

To summarise the findings of the attitude study, it confirms the salience of the Innovative hiatus filling pattern, compared to the Conservative one. The participants’ comments also show that while speakers are aware of the variable, they are unable to describe its behaviour accurately.

### 6.2 Analysis

This section looks at the correlation between salience and transitional probabilities in Hungarian hiatus resolution. The procedure is the following: I take a written corpus of Hungarian, and modify it in such a way that it includes [j]-s resulting from standard hiatus resolution, not marked in the orthography. I extract the frequencies of [ij] and [ji] clusters in order to gain the transitional probabilities (TP-s) \( p([j][i]) \) and \( p([i][j]) \) in the corpus. I also extract the frequencies of [ej] and [je] to gain the TP-s \( p([j][e]) \) and \( p([e][j]) \).

The hypothesis is that pairs of [j] and [i] are much more frequent than pairs of [j] and [e]. Consequently, the former are more familiar to the listeners, so when these occur as a result of hiatus resolution, the pattern is not salient. In comparison, the latter are much less familiar, so when these result from hiatus resolution, the pattern becomes salient. Again, the key point is that
clusters of, for instance, [ɛj] are not illicit and occur in Hungarian, their salience in hiatus resolution comes from the frequency difference.

### 6.2.1 Corpus results

The data are drawn from the Hungarian Webcorpus (Halácsy et al., 2004), a corpus of 1.48 billion words from 18 million pages downloaded from the .hu Internet domain, which gives the best representation of written language, and is the most faithful corpus of present-day Hungarian. A sample of 17 million words was used to establish TP-s. Hungarian orthography is relatively consistent, at least when it comes to the representation of [j], [i], and [ɛ]. It does not mark hiatus resolution, so I inserted [j]-s into vowel clusters including [i]. This step is valid inasmuch as hiatus resolution is obligatory in these clusters. It assumes, however, that hiatus filler [j]-s are equal to contrastive [j]-s in the language. This assumption is supported by authors like Kálmán & Rebrus (2010), who argue that the intrusive segment in hiatus resolution is phonetically equal to the one in the possessive. This, in turn, means that all intervocalic [j]-s are interpreted equally, as the possessive [j] is virtually indistinguishable from the contrastive ‘lexical’ one. An example to this is given in Table 6.4. We can, then, assume an identical status for the [j]-s which are lexical or arise through hiatus resolution, an issue touched upon in this chapter.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>zoknija [zoknija] ‘sock-POSS3SG’</td>
<td>Possessive suffixation</td>
</tr>
<tr>
<td>szoknia [soknija] ‘accustom-INF3SG’</td>
<td>Hiatus resolution before infinitive -a</td>
</tr>
<tr>
<td>kijavít kijovízt ‘fix-3SG’</td>
<td>Contrastive [ijn] sequence</td>
</tr>
</tbody>
</table>

Table 6.4: [j] in Hungarian

Even if we take the ontology of [j] as granted, the analysis has to cope with another difficulty, the lack of reliable data on hiatus resolution. In
Conservative ECH, it is agreed to be obligatory in vowel clusters with [i] and variable in clusters containing [ɛː]. There are no estimates on Innovative ECH. In order to tackle the scarcity of the data, I take up the approach of looking at transitional probabilities in one dialect instead of comparing two.

This dialect, Conservative Educated Colloquial Hungarian, is assumed to be represented by the Webcorpus. It has obligatory hiatus resolution with [i], but has none with [ɛː] (since it is conservative). Nonetheless, it also has ‘lexical’ instances of vowel clusters with [i]/[ɛ] and [j]. This is illustrated in Table 6.5.

<table>
<thead>
<tr>
<th>Lexical sequence</th>
<th>Hiatus</th>
</tr>
</thead>
<tbody>
<tr>
<td>kijárat [kijaːr6t] ‘exit’</td>
<td>kiárad [kijaːr6d] ‘flow-3sg’</td>
</tr>
<tr>
<td>Tejút [t3juːt] ‘Milky Way’</td>
<td>szemleút [scmlEːут] ‘field trip’</td>
</tr>
</tbody>
</table>

Table 6.5: Conservative ECH

With these presumptions, we can look at the frequency differences of non-salient and salient hiatus resolution in the corpus. The frequency of the relevant string in the corpus is given in Table 6.6. (Both word-internal clusters and clusters including a word boundary were included.) The TP-s are given in Table 6.7 (numbers are rounded to the third decimal place).

<table>
<thead>
<tr>
<th>String</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ɛj]</td>
<td>103024</td>
</tr>
<tr>
<td>[je]</td>
<td>230857</td>
</tr>
<tr>
<td>[ij]</td>
<td>480943</td>
</tr>
<tr>
<td>[ji]</td>
<td>391069</td>
</tr>
<tr>
<td>[i]</td>
<td>4424703</td>
</tr>
<tr>
<td>[j]</td>
<td>2367677</td>
</tr>
<tr>
<td>[ɛ]</td>
<td>10892098</td>
</tr>
</tbody>
</table>

Table 6.6: String frequencies in the corpus sample
The results show that there is a frequency difference of one order of magnitude between the TP of [j] following [i] versus that of [j] following [e]. There is no difference, however, when we look at the pattern the other way around, that is, between the TP of [i] following [j] versus [e] following [j]. This asymmetry can be probably blamed on the possessive suffix, which is -je after vowel-final front vowel stems. The ratios are shown in Figure.

<table>
<thead>
<tr>
<th>Segments</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(j</td>
<td>e)</td>
</tr>
<tr>
<td>e</td>
<td>j</td>
</tr>
<tr>
<td>p(j</td>
<td>i)</td>
</tr>
<tr>
<td>p(i</td>
<td>j)</td>
</tr>
</tbody>
</table>

Table 6.7: TP-s in the corpus sample

Figure 6.3: Transitional Probabilities for the different patterns involving [j]
6.2.2 Main points

What the corpus study tells us, then, is that the salience of the innovative hiatus resolution pattern shows a correlation with the relative low frequency of the string [ɛj] (when compared to [ij]), itself part of the realisation of the resolved hiatus. This supports the hypothesis that the salience of the innovative hiatus resolution pattern, confirmed by the attitude study, springs from a difference in transitional probabilities, which difference renders the pattern less familiar to the listeners.

Two questions should be addressed at this point. First, it has to be stressed that there are no data available on the Innovative ECH dialect assumed here, apart from its existence. The extent of hiatus resolution in vowel clusters with [ɛ], as well as its origin and correlation with innovative realisations of other variables remain subject to a future study. Second, one might argue that if, based on the corpus data, instances of [j] following [ɛ] are unfamiliar, we should expect the salience of any [ɛj] sequence, not just the ones arising through hiatus resolution. To put the question differently: why is teja salient, but bejárat ‘entrance’ apparently not? In my view, the difference lies in the patterning of the two types of clusters. Lexical, contrastive clusters show no variation in ECH. In this sense, one cannot talk about conservative and innovative use, or, indeed, about a linguistic variable. The rarity of [ɛj] clusters is relevant where these clusters occur variably.

6.3 Concluding remarks

This chapter has shown that, in Educated Colloquial Hungarian, a possible reason for speaker awareness, or, at least, marked speaker attitudes towards one kind of hiatus resolution but not the other can be linked to the salience of the former, but not the latter. This difference in being salient, in turn, can be linked up with the different probability of occurrence for the patterns which arise through one type of hiatus resolution as opposed to the other one. Consequently, Hungarian hiatus resolution is another example for the
relationship between salience (as expressed through, for instance, speaker attitudes) and probability of occurrence.

This study is evidently different from the previous ones in one major respect: whereas, for example, in the case of /t/ glottalisation in the South of England, I compared two dialects in which the variable behaves differently, in the present case, I looked at only one dialect, because the amount of available data made a similar comparison impossible. The change in methodology, however, did not influence the main argument. As shown, salience as a property becomes possible if the extent of use of a variable in one dialect is notably different from this extent in another dialect. If we look at the problem this way, we can say that any variability in a dialect is a possible ‘entry point’ of social evaluation, though not a necessary one. Indeed, this only amounts up to saying that linguistic variables carry social indexation, which should not come as a shock. What should be stressed, however, is that the interaction between salience and distributions can be meaningfully operationalised even when we build on data only from one dialect.

A shared trait with the glottalisation study, however, is the problem concerning the origin of the difference in distributions in the two dialects. The reason why Conservative ECH speakers – as well as upper-middle class speakers from the South of England – notice a variable is that it is realised to a different extent in a dialect dissimilar to their own. In the glottalisation chapter I hinted at a possible explanation for the precedence of the non-salient variable over the salient one, namely, phonetic grounding (or the absence thereof). Here, we can take a similar note.

The reason why hiatus resolution in vowel clusters with [i] goes unnoticed in Hungarian is that it exhibits no variation. The reason for the absence of variation, however, is probably the phonetic grounding of hiatus resolution with [j] in these environments, which grounding is absent in clusters with [ɛ]. What follows is that the latter is much less susceptible of developing naturally. The consequent priority of the phonetically grounded hiatus resolution pattern over the ungrounded one paves the way for the latter’s salience.
Chapter 7

Derhoticisation in Glasgow

In the previous chapters, I argued for a straightforward, segmental approach to social markers. If the difference in the realisations of a variable creates differences in the segmental distributions in two related dialects – in the previous examples, a conservative norm dialect and an innovative dialect – speakers become aware of the segmental difference, and, consequently, the social connotations of variable use. This view is curbed in the sense that it can only apply in cases where we can posit abstract segmental units and differences in variable realisation resulting in differences in the segmental distributions, that is, the transitional probabilities of the segments. As such, the concept of salience only has a bearing on certain consonantal variables. We do expect these variables, however, to meet our expectations and to be sociolinguistically salient when the requirements for segmental distributions hold.

If a consonantal variable shows different transitional probabilities in two related dialects, and yet fails to carry social indexation, the state of affairs becomes more complicated. Certainly, difference in distributions is not the single and necessary requirement for sociolinguistic salience. Then again, such a case should not be brushed off the table saying that it is an unfortunate, though legitimate, exception to the rule. It is much more useful to look at the variable’s larger context and ask the question whether the language system
in which the variable is defined has an effect on speaker awareness, and if so, how. After all, the approach advocated here works with severe restrictions, which is a merit, but which, in certain cases, can prove to be a disadvantage.

7.1 Background

In this chapter, I have a look at derhoticisation, the vocalisation of the coda /r/ in the Urban Scots dialects of Glasgow. I will argue that, based on available data, this variable seems to carry no social indexation, despite the fact that it meets the requirements for sociolinguistic salience, causing a difference in segmental distributions. I will go on to claim that the apparent absence of speaker awareness is due to the vast extent of variability inherent to derhoticisation, and the behaviour of /r/ in Glasgow Scots in general, and try to point to reasons why this is so.

This section contains mostly literature review. It tries to cover the available data on derhoticisation, along with the conclusions previous researchers have drawn from this data. In the second part of the chapter, Section 7.2, I test whether we would expect the loss of coda /r/ as socially salient, and look at the possible explanations available if it proves not to be salient at all.

In Section 7.1.1, I discuss the general problem tackled in the chapter, namely, the existence of indicators, cases where social stratification does not go with speaker awareness, and the consequences for a segmentally based definition of salience. Section 7.1.2 gives the concrete example, derhoticisation, or the loss of coda /r/ in Glasgow and other Scottish English dialects, and goes into a lengthy description of available data in an attempt to show the fullest scale of the variation and relevant factors of this change. Section 7.1.5 gives an interim summary. Based on the conclusions of this summary, Section 7.2.1 builds a comparative scale of distributions based on data from the FRED corpus, in order to show that phonetical gradience will lead to large variation, which, in turn, will bereave the variable of salience. Section 7.3 provides a general discussion.
7.1.1 Social stratification and social awareness

The view maintained throughout this book is that language change which affects segmental distributions in an uneven way in two related dialects will possibly come to the attention of speakers. Speaker attention does not necessarily involve stigmatisation and other forms of meta-awareness (such as the ability to imitate an accent). However, it goes together with style shifting, the preference of a conservative variant in a formal setting, and explicit attitudes towards speakers using a particular variant. Style shifting does not exclusively happen in this direction, and speaker attitudes can go beyond being simple negative or positive. That notwithstanding, the presence of style shifting or contoured speaker attitudes can be relatively easily ascertained and are reliable indices of speaker awareness.

Social stratification and speaker awareness of a variable should go hand in hand, so much so that, for socially stratified variables, the absence of sociolinguistic salience is more the exception than the rule. One classic example is Labov’s New York Department Store study (Labov, 1972b), which goes so far as to conflate social stratification and style shifting in collecting data from employees of stores favoured by a clientele of different social backgrounds. Labov’s implicit assumption is that, in case of the variable under his scrutiny, (rV), the vocalisation of coda /r/, both social stratification and sociolinguistic salience will be a factor: Any employee of a store of high standing is both more likely to have a better social status and tend to speak in a more ‘sophisticated’ manner with costumers. (During the course of this chapter we will be confronted with another variable related to /r/, namely, (r), the variable realisation of coda /r/, i.e. the variation in its realisation, not whether it is there or not.)

His Lower East Side study (Labov, 1966/2006) shows that such behaviour is not limited to department store employees: the variable (rV) shows the most fine-grained correlation with social status in the community, and the responses to it betray clear attitudes towards its use in the matched guise tests: his subjects associate a larger extent of coda /r/ realisation with a
higher socio-economic standing.

Such a response is expected if we accept the difference in segmental distributions as a source of salience. As I will try to show in section 7.2.1, any shift in the realisation of coda /r/ causes considerable changes in the distributions of vowels followed by [r] as well as [r] preceding consonants: to put it very simply, both of these are fairly rare in non-rhotic dialects of English, and relatively common in rhotic ones.

This approach is warranted when one considers Labov’s own explanation of the language change resulting in the different coda [r] ratios in the Lower East Side, and in New York City in general. He argues that an external norm, that of other rhotic American English dialects, had been introduced in New York City, which was adopted first by the upper social classes, and is pervading the language system. Labov regards rhoticisation in New York City as a classic change from above, that is, a change that happens above the threshold of consciousness. He admits that the term is somewhat unfortunate, even more so because, in this particular case, it coincides with another ‘vertical’ movement: the introduction of the norm by the upper social classes, and its moving downwards to be adopted by the lower ones. To keep the two terms separate, I will refer to a change that starts at the higher end of the social ladder, and is proceeding downwards, as top-down. A change that starts in a social group other than the elite, and typically in the working class, I will call bottom-up.

Changes from above, that is, salient changes do not necessarily proceed top-down. The glottalisation of /t/, discussed in Chapter 5, is an excellent example. In the South-East of England it started as a substandard class feature (if we ignore, for a moment, its long and turbulent history). Bottom-up spreading is a characteristic of Labov’s change from below. According to Labov (1966/2006), changes from below happen below the threshold of speaker consciousness, and all his examples originate in a lower socio-economic class. One difference, then, between glottalisation and this notion of a change from below is that while a change from below is supposed to proceed without
any speaker awareness, glottalisation quickly became a characteristic class feature in England, leading some to claim it as a defining feature of the emerging standard in the South, Estuary English (Rosewarne, 1984). In this sense, /t/ glottalisation is a bottom-up change, starting below and climbing the social ladder, but, in its later stages, certainly a change from above.

This leaves us with a fourfold typology. Changes from above proceed in a way that is conscious to the speakers, while changes from below remain unnoticed. Top-down changes are linguistic innovations starting at the top of the social scale, and are held as a norm for the rest of the community. In contrast, bottom-up changes start lower on the social scale. (One should note that this typology assumes a view of language change climbing up and down on a ladder. Details beyond classic class structure are evaded in this chapter, but are taken into consideration in Chapter 8.)

We can probably rest assured that a top-down adaption of a new linguistic norm will be a change from above, a salient change, resulting in style shifting as well as social stratification. A much more interesting question is whether all bottom-up changes end up being salient like glottalisation did. Both Labov (1966/2006) and Trudgill (1974) list variables which carry no social indexation, that is, are only indicators, not markers. Clearly, speaker awareness also depends on other factors, such as the extent and speed of the change in question (as well as its relation to the local norm system). The decisive factor, as argued in this book, is salience, emergent from the different distributions of the realisations of the variable in question – the forms affected by language change. If, in two related dialects, the difference between the realisations entails differences in the segmental distributions of these dialects, the variable has a chance to be picked up by listeners and become a social marker. What this means, in turn, is that quite a number of consonantal variables should be able to carry social indexation, (rV), coda /t/ vocalisation, being a prime example.

The reader should note that this discussion pretends that all linguistic variation is connected to language change. This is not the case (cf. Chapter
8).

As long as it shows social stratification, coda /r/ vocalisation should be an excellent candidate for a linguistic marker – all the more awkward for a theory of salience if it does not prove to be so in a particular case. The rest of the chapter discusses such a case, the emergence of a new urban working-class dialect of Glaswegian. Coda /r/ vocalisation seems to be an indicator of this dialect, as much as it shows clear social stratification in Glasgow, but it looks as if speaker awareness were absent.

I argued more than once previously that differences in segmental distributions do not necessarily lead to sociolinguistic salience. Yet, instead of filing Glasgow /r/ as an apparent counterexample, I will try to look into its general context to answer the question why it does not carry social indexation, though it could – at least based on expectations couched in this theory.

7.1.2 Derhoticisation in Glasgow

The vocalisation of coda /r/, or derhoticisation, is not a new phenomenon in Scotland. Scottish English (and Scots) are traditionally regarded as rhotic, that is, consistently realising post-vocalic /r/ in words like car and card. The most conservative view on /r/ in Scottish English dialects is that it is realised as an alveolar trill, a sound realised by vibrations between the tip of the tongue and the alveolar ridge of the palate. The standard realisation, however, is commonly regarded to be a retroflex approximant, a sound realised by a narrow constriction of the tongue body or tongue tip behind the alveolar ridge (Wells, 1982). As we will see, this traditional middle-class variant is now contested by a new one, typical of working-class male adolescents, and working-class speakers in general. This variant is the alveolar tap, a single constriction created by the tip of the tongue and the alveolar ridge. These two are not the only possible, or, indeed, prevalent realisations of /r/ in Scotland, as both central approximants and, to an extent, uvular approximants and trills are to be found (Stuart-Smith, 2003).

Furthermore, the coda /r/ can be vocalised in the Scottish Dialects of the
Urban Belt of Scotland, most notably in Edinburgh and Glasgow. It has been observed by, among others, Romaine (1978) and Johnston (1983) in Edinburgh and Macafee (1983) in Glasgow. The most recent and most thorough study was conducted in Glasgow in 1997. At a glance, /r/ vocalisation can be handled fairly well segmentally. The change, in a very basic interpretation, would consist of sequences of /Vr/ becoming either [V:] or an in-gliding diphthong (as in Southern English English), a rhoticated vowel (as in General American), or some other recognisable sound segment. In practice, however, as we will shortly see, the variation of coda /r/ in Scotland goes far beyond the limitations of any simplified segmental approach.

To sum things up very briefly, the typical middle-class realisation of coda /r/ is a retroflex approximant, slightly similar to the General American rhotic. The innovative, working-class variant is an alveolar tap, very much like in Swiss German or Hungarian. Furthermore, in the South of Scotland, coda /r/ seems to be disappearing, undergoing vocalisation, on a par with English dialects in England (at least at a first glance). Coda /r/ exhibits many other variable realisations in Scotland.

7.1.3 /r/ in Glasgow

The main data source on which I am going to rely in this section is the corpus collected by Jane Stuart-Smith, Claire Timmins, and Fiona Tweedie (Stuart-Smith et al., 2007; Timmins et al., 2004) in Glasgow in 1997. I will also look at other studies on /r/ in Glasgow to paint a fuller picture on its variation. The 1997 Glasgow study involves 32 informants, divided equally according to age (older: 40-60 years; younger: 13-14 years old), gender, and socio-economic background (middle-class and working-class). Location in the city is also a factor. Middle-class speakers were recorded in the northern suburbs of Bearsden, Milngavie, and Westerton, targets of an ongoing middle-class influx since the mid-nineteenth century. Working-class speakers come from the districts of Maryhill, the Wyndford estate, Ruchill, and Possil, all locations of tenement buildings and thorough council housing schemes. Young informants
were collected from two schools in the two areas. Recording sessions consist of spontaneous speech, a conversation of a same-sex pair of informants in the interviewer’s absence, and reading, reading a word list of around 180 words. (For details, cf. Stuart-Smith et al. 2007.)

Limitations of space do not allow presenting the results in their entirety, so I will concentrate on the two relevant consonantal variables, (rV), /r/ vocalisation and (r), /r/ realisation. The data are taken from the report on the consonantal variables, Timmins et al. (2004), after statistical analysis performed by the authors. The counts can be regarded as percentages. The authors only included differences which they found to be statistically significant, hence the absence of certain distinctions.

The first variable to look at is (rV). The following variants are posited: [r]: articulated coda /r/, [V]: long vowel or vocalised variant, [r/V]: in-between, [V˘]: velarised vowel, [Vr]: rhotacised vowel. Table 7.1 shows reading realisations depending on social class. Table 7.2 shows reading realisations depending on phonetic environment. MC stands for middle-class, WC for working-class. O refers to the older age group, and Y to the younger one. In cases where gender is a factor, F stands for female, and M for male. The results of the reading data are visualised in Figure 7.1.

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[V]</th>
<th>[r/V]</th>
<th>[V˘]</th>
<th>[Vr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCO</td>
<td>88.21</td>
<td>1.43</td>
<td>0.36</td>
<td>1.43</td>
<td>8.57</td>
</tr>
<tr>
<td>MCY</td>
<td>81.41</td>
<td>6.32</td>
<td>0.00</td>
<td>5.58</td>
<td>6.69</td>
</tr>
<tr>
<td>WCO</td>
<td>62.91</td>
<td>17.45</td>
<td>2.55</td>
<td>9.82</td>
<td>7.27</td>
</tr>
<tr>
<td>WCY</td>
<td>17.56</td>
<td>50.54</td>
<td>2.15</td>
<td>25.81</td>
<td>3.94</td>
</tr>
</tbody>
</table>

Table 7.1: (rV) versus social class: reading (n=1103)

A significant effect of class is observable in reading. Middle-class speakers are predominantly rhotic, and also use rhotacised vowels, though to a much smaller extent. Vocalised tokens are rare. Working-class speakers, especially adolescents, vocalise to a large extent, and also show a considerable use of
Background

velarised vowel variants.

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[V]</th>
<th>[r/V]</th>
<th>[V˘]</th>
<th>[Vr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-consonantal</td>
<td>51.57</td>
<td>9.1</td>
<td>19.08</td>
<td>20.24</td>
<td>0</td>
</tr>
<tr>
<td>pre-pausal(stressed)</td>
<td>69.84</td>
<td>19.94</td>
<td>0.78</td>
<td>9.18</td>
<td>0.2</td>
</tr>
<tr>
<td>pre-pausal(unstressed)</td>
<td>63.54</td>
<td>26.56</td>
<td>1.56</td>
<td>6.25</td>
<td>2.08</td>
</tr>
<tr>
<td>overall</td>
<td>61.65</td>
<td>18.53</td>
<td>7.14</td>
<td>11.89</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table 7.2: (rV) versus environment: reading (n=1103)

Reading data reveals some sensitivity to the phonetic environment. The extent of rhotic realisation is the smallest pre-consonantally, and intermediate and velarised vowel variants have a strong presence in this position, with the former featuring almost only here. This hints at the complexity following from the phonetically fine-grained nature of derhoticisation in pre-consonantal position. The extent of rhoticity, or the use of [r], however, has no strong patterning word-finally, which again hints at the change being phonetically gradual and less sensitive to the following environment. Tables 7.3 and 7.4 show spontaneous speech realisations depending on social class and phonetic environment, respectively. These tables are visualised in Figure 7.2.

The authors note that, with all main effects being significant, working-class girls are different from all others bar working-class boys and women, whilst working class boys are polarised from middle-class men and girls. We might observe that reading patterns are retained in the sense that middle-class speakers are mostly rhotic with working-class speakers and especially working-class adolescents vocalising much more. One chief difference between the effects of the phonetic environment in read and spontaneous speech is that, in the latter, indeterminate cases of [r/V] are almost absent, with more

---

1It is true that there is a notable difference between the use of [r] pre-consonantally and word-finally. However, this might be due to the fact that many more tokens were categorised as indeterminate [r/V] in the former environment, where identification of variants is generally more difficult.
Figure 7.1: Derhoticisation versus social class and environment: reading (n=1103)
Table 7.3: (rV) versus social class: spontaneous speech (n=8170)

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[V]</th>
<th>[r/V]</th>
<th>[V^]</th>
<th>[Vr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCOF</td>
<td>94.33</td>
<td>4.12</td>
<td>0.14</td>
<td>0.07</td>
<td>1.35</td>
</tr>
<tr>
<td>MCOM</td>
<td>92.02</td>
<td>3.71</td>
<td>0.00</td>
<td>2.23</td>
<td>2.04</td>
</tr>
<tr>
<td>MCYF</td>
<td>93.03</td>
<td>2.92</td>
<td>0.16</td>
<td>0.00</td>
<td>3.89</td>
</tr>
<tr>
<td>MCYM</td>
<td>89.39</td>
<td>4.22</td>
<td>0.22</td>
<td>4.65</td>
<td>1.52</td>
</tr>
<tr>
<td>WCOF</td>
<td>69.07</td>
<td>22.45</td>
<td>0.86</td>
<td>2.42</td>
<td>5.20</td>
</tr>
<tr>
<td>WCOM</td>
<td>75.38</td>
<td>12.94</td>
<td>0.51</td>
<td>6.85</td>
<td>4.32</td>
</tr>
<tr>
<td>WCYF</td>
<td>5.82</td>
<td>76.33</td>
<td>1.50</td>
<td>15.15</td>
<td>1.20</td>
</tr>
<tr>
<td>WCYM</td>
<td>22.13</td>
<td>50.50</td>
<td>2.21</td>
<td>17.51</td>
<td>7.65</td>
</tr>
</tbody>
</table>

Table 7.4: (rV) versus environment: spontaneous speech (n=8170)

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[V]</th>
<th>[r/V]</th>
<th>[V^]</th>
<th>[Vr]</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-consonantal</td>
<td>64.6</td>
<td>19.5</td>
<td>0.56</td>
<td>7.69</td>
<td>7.65</td>
</tr>
<tr>
<td>pre-pausal(stressed)</td>
<td>70.71</td>
<td>21.45</td>
<td>0.87</td>
<td>5.91</td>
<td>1.06</td>
</tr>
<tr>
<td>pre-pausal(unstressed)</td>
<td>67.84</td>
<td>27.13</td>
<td>0.90</td>
<td>3.34</td>
<td>0.79</td>
</tr>
<tr>
<td>overall</td>
<td>67.72</td>
<td>22.69</td>
<td>0.78</td>
<td>5.65</td>
<td>3.17</td>
</tr>
</tbody>
</table>

The picture of working-class adolescents spearheading the derhoticisation process emerges from these data, along with the massive scale of variation in the process itself. One main point of interest is the relatively small extent of style-shifting. That is, realisation patterns do not differ massively depending on whether the subjects read a word list or conversed between themselves. In order to highlight this, I compare the results for [r], rhotic realisation for read and spontaneous speech in Table 7.5. Again, this variant can be regarded as the conservative, traditional one in Scots. To put the extent of
Figure 7.2: Derhoticisation versus social class and environment: spontaneous speech (n=8170)
Background

style shifting into context, I include two other variables in the comparison, /l/ vocalisation and /t/ glottalisation. Both show social stratification in the data and are argued elsewhere to be markers of the emerging working-class dialect in Glasgow (Stuart-Smith et al., 2007).

\begin{table}
\begin{tabular}{cccccc}
\hline
 & (rvoc):[r] & (l):[V] & (t):[?] \\
Read S & Spont S & Read S & Spont S & Read S & Spont S \\
MCO & 88.21 & 93.17 & 0.10 & 0.73 & 56.56 \\
MCY & 81.41 & 91.21 & 0.00 & 4.65 & \\
WCO & 62.91 & 72.23 & 0.42 & 6.91 & 92.47 \\
WCY & 17.56 & 13.98 & 8.83 & 76.32 & \\
\hline
\end{tabular}
\caption{Style shifting examples in the 1997 corpus ((rV) RS n=1103, SS n=8170; (l) RS n=656, SS n=3420; (t) RS n=1212, SS n=2385)}
\end{table}

In Table 7.5 I compare one particular realisation of three consonantal variables investigated in the 1997 corpus. The first one is /r/ vocalisation. The variant I chose is realised coda /r/ (i.e./ all the non-vocalised realisations), represented by [r]. We can see that all social groups have more realised [r]-s in spontaneous speech than in formal speech, except for the working-class adolescents. The difference, however, is never more than ten per cent. For working-class adolescents, it is less than five per cent. This contrasts sharply with the two other variables.

The second one is /l/ vocalisation. The variant shown here is [V], vocalised coda /l/. We can see that middle-class speakers and working-class adolescents do not use it to a large extent in reading in the first place, and completely avoid it in spontaneous speech. The one exception is again the working-class adolescent group. They use the vocalised variant five times more often in reading than in spontaneous speech. This might come as a surprise given the fact that we expect the reverse: avoidance of the innovative vocalised variant in the more formal reading setting. Stuart-Smith et al. (2007), however, make the point that, as a strong marker of working-class identity, vocalised /l/ might be exaggerated in the reading list. In any case, the presence of style
Derhoticisation in Glasgow

shifting in any direction sharply indicates the marker status of this variable. With /t/ glottalisation, we see middle-class and older working-class speakers strongly avoiding the glottal stop variant [ʔ] in reading, even though the latter group uses it overwhelmingly in spontaneous speech. (The two data points in the last column indicate that there is no significant difference in age in the realisation of this variant.) Working-class adolescents also show style-shifting.

Even if we think that style shifting is insufficient evidence on social awareness to a particular variable, it is still striking that coda /r/ realisation shows it much less than the other two, all the more since coda /r/ realisation is a textbook example of both social stratification and social awareness in language use. One could put forward that derhoticisation in Glasgow, as opposed to ‘rhoticisation’ in New York City, evades speaker attention precisely because it fits Labov’s description of a change from below: it starts lower on the social scale and proceeds gradually, as an articulatory lenition dependent on the phonetic environment and showing massive variation. In comparison, in New York City, we have the case of a change from above, the almost conscious adoption of an external norm by the upper-middle-classes. Therefore, though the social stratification of coda /r/ realisation is fairly similar in the two cases, with people higher on the socio-economic speaking more rhotic, the comparison of the two language settings is misleading. (Though, it has to be said, the situation is different with age: in Glasgow, older speakers tend to be more rhotic, which is the other way around in the New York City case.)

This argument is refuted by the presence of massive style shifting with the other variables, two of them exemplified above. These variables also seem to participate in the carving out of a new working-class identity in Glasgow, but all of them start low on the social scale, and not all of them can be regarded as the conscious adoption of an external norm: Glasgow is the hometown of /t/ glottalisation, and while both this variable and /l/ vocalisation might be influenced by emergent working-class norms elsewhere in Britain, they clearly originate in articulatory reduction. Researchers do not agree completely on the extent of the influence English working-class standards
on Glasgow working-class speech. I do not want to arbiter on this issue here, but it is certainly true that Stuart-Smith et al. (2007) convincingly argue against blaming language change in Glasgow on English English influence (for instance, since the English English standard is non-rhotic, we would expect conscious adoption by the middle-class speakers, and yet the reverse seems to be happening). What is more, there are other examples of bottom-up changes becoming social markers, one extensively discussed in this book being /t/ glottalisation in England.

In sum, this discussion brings us back to the distinction between changes from above/below and top-down/bottom-up changes. While the two concepts are not completely orthogonal – top-down changes are likely to be always changes from above, innovations salient to the speakers – a bottom-up change may not remain below the threshold of speaker awareness.

What follows is that the cause for the absence of style shifting in /r/ vocalisation in Glasgow has to be sought elsewhere. This is in the spirit of this book, which continuously argues that the reason for a variable’s salience is not the history or social frame of the change but rather the phonetic/phonological distributions this change affects. We will look at more evidence on the (absence of) social evaluation of /r/ in Glasgow. But before we do so, we have to pay attention to another of the factors that might influence it, namely (r), variation in /r/ realisation.

Stuart-Smith and her colleagues only look at the variation of the realisation of /r/ in the first ten per cent of the conversations in the spontaneous speech register of the 1997 corpus, due to the large frequency of the variable in speech. Social class and gender, but not phonetic environment is included as factors. They define the following variant categories: [r]: central approximants, [R]: retroflex approximants, [rt]: mainly voiced alveolar taps, [Rt]: retroflex taps, [Ru]: voiced uvular fricative, [rtt]: voiced alveolar trill, [V]: vocalised variants (discussed above), [m]: miscellaneous. Table 7.6 shows the distributions of these variants in reading.

Table 7.6 re-instates a known fact about Scots dialects, namely, that
Derhoticisation in Glasgow

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[R]</th>
<th>[rt]</th>
<th>[Rt]</th>
<th>[Ru]</th>
<th>[V]</th>
<th>[m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCOF</td>
<td>13.42</td>
<td>58.87</td>
<td>21.64</td>
<td>0.87</td>
<td>0</td>
<td>5.19</td>
<td>0</td>
</tr>
<tr>
<td>MCOM</td>
<td>11.25</td>
<td>44.15</td>
<td>29</td>
<td>7.36</td>
<td>0</td>
<td>8.22</td>
<td>0</td>
</tr>
<tr>
<td>MCYF</td>
<td>6.64</td>
<td>64.6</td>
<td>23.45</td>
<td>0</td>
<td>0</td>
<td>5.31</td>
<td>0</td>
</tr>
<tr>
<td>MCYM</td>
<td>5.17</td>
<td>41.38</td>
<td>30.6</td>
<td>0</td>
<td>0</td>
<td>22.84</td>
<td>0</td>
</tr>
<tr>
<td>WCOF</td>
<td>13.85</td>
<td>20.78</td>
<td>44.59</td>
<td>4.76</td>
<td>0</td>
<td>16.02</td>
<td>0</td>
</tr>
<tr>
<td>WCOM</td>
<td>12.61</td>
<td>0.43</td>
<td>60.87</td>
<td>0</td>
<td>0</td>
<td>26.09</td>
<td>0</td>
</tr>
<tr>
<td>WCYF</td>
<td>12.66</td>
<td>0</td>
<td>29.69</td>
<td>0</td>
<td>0.44</td>
<td>56.33</td>
<td>0.87</td>
</tr>
<tr>
<td>WCYM</td>
<td>7.92</td>
<td>2.2</td>
<td>37.44</td>
<td>0</td>
<td>0.44</td>
<td>51.54</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Table 7.6: (r) versus social class: reading (n=1834)

Retroflex approximants are typical in middle-class use, with taps being more common in the working-class (Wells, 1982). Central approximants are used, to an extent, by all speakers. The table also shows the extent of /r/ vocalisation, stronger in the case of working-class speakers, and, specifically, working-class adolescents.

<table>
<thead>
<tr>
<th></th>
<th>[r]</th>
<th>[R]</th>
<th>[rt]</th>
<th>[Rt]</th>
<th>[Ru]</th>
<th>[ttt]</th>
<th>[V]</th>
<th>[m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCOF</td>
<td>23.10</td>
<td>52.71</td>
<td>21.66</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>2.53</td>
<td>0.00</td>
</tr>
<tr>
<td>MCOM</td>
<td>12.37</td>
<td>64.21</td>
<td>10.37</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>13.04</td>
<td>0.00</td>
</tr>
<tr>
<td>MCYF</td>
<td>13.71</td>
<td>60.48</td>
<td>22.58</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>3.23</td>
<td>0.00</td>
</tr>
<tr>
<td>MCYM</td>
<td>14.71</td>
<td>52.21</td>
<td>28.68</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>4.41</td>
<td>0.00</td>
</tr>
<tr>
<td>WCOF</td>
<td>11.00</td>
<td>39.23</td>
<td>26.79</td>
<td>2.39</td>
<td>0.00</td>
<td>0.00</td>
<td>20.57</td>
<td>0.00</td>
</tr>
<tr>
<td>WCOM</td>
<td>18.39</td>
<td>44.44</td>
<td>23.75</td>
<td>0.00</td>
<td>0.00</td>
<td>1.15</td>
<td>11.88</td>
<td>0.38</td>
</tr>
<tr>
<td>WCYF</td>
<td>14.84</td>
<td>4.52</td>
<td>20.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>60.00</td>
<td>0.65</td>
</tr>
<tr>
<td>WCYM</td>
<td>2.17</td>
<td>19.57</td>
<td>28.26</td>
<td>0.00</td>
<td>2.17</td>
<td>0.00</td>
<td>45.65</td>
<td>2.17</td>
</tr>
</tbody>
</table>

Table 7.7: (r) versus social class: spontaneous speech (n=1474)

Table 7.7 shows (r) in spontaneous speech. Older working-class speakers use the retroflex variant to a much larger extent in spontaneous speech than in reading, and the same is true for young working-class males. The two
tables also compare the use of the vocalised variant in read versus spontaneous speech: again, we see no strong style shifting, with the notable exception of young middle-class males, who tend to avoid it much more in spontaneous speech. Tables 7.6 and 7.7 are visualised in Figure 7.3.

Table 7.8 compares reading and spontaneous speech data for three variants in the 1997 corpus. All three variants are realisations of coda /r/: [rt] is the tap, typical of working-class speakers, [Rt] is the retroflex approximant, a characteristic of middle-class speakers, and [V] is the innovative, vocalised variant, again more typical of working-class speakers. MCO, MCY, WCO, and WCY stand for middle-class older and younger, and working-class older and younger speakers. The columns allow comparison in the use of the variables across speaker groups. What is important to see is that, for both older and younger working-class speakers, the extent of style shifting is considerably larger for /r/ realisation than for /r/ vocalisation, though it is way behind, for instance, /l/ vocalisation or /t/ glottalisation. This shows again that style shifting is much smaller here than expected for such a usually socially salient variable as /r/ vocalisation.

<table>
<thead>
<tr>
<th></th>
<th>[rt]</th>
<th>[Rt]</th>
<th>[V]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read S</td>
<td>Spont S</td>
<td>Read S</td>
</tr>
<tr>
<td>MCO</td>
<td>51.51</td>
<td>58.46</td>
<td>25.32</td>
</tr>
<tr>
<td>MCY</td>
<td>52.99</td>
<td>56.34</td>
<td>27.03</td>
</tr>
<tr>
<td>WCO</td>
<td>21.21</td>
<td>41.84</td>
<td>52.73</td>
</tr>
<tr>
<td>WCY</td>
<td>1.1</td>
<td>12.05</td>
<td>33.57</td>
</tr>
</tbody>
</table>

Table 7.8: Style shifting examples in the 1997 corpus: (r) (RS n=1834, SS n=1474)

Two strong patterns are observable in the /r/ realisation data. First, /r/ realisation, though in no simple way, is socially stratified, with middle-class speakers leaning towards the retroflex variant, and working-class ones towards the flap. This is accompanied by a fair amount of style-shifting, especially with older middle-class speakers, who prefer the retroflex approximant, usually
regarded as the middle class variant, in spontaneous speech. Male working-
class adolescents show a similar patterning. Interestingly enough, the extent of

Figure 7.3: (rV) versus social class: reading and spontaneous speech
style shifting is larger than in the case of /r/ vocalisation, as can be seen both here and above, which suggests that speakers are, in a way, more sensitive to the realisation of /r/ than to its vocalisation in the coda. Second, /r/ realisation varies greatly, with four variants being used in more than ten per cent of the cases by one group or the other (a fifth variant, the retroflex tap, is also present in the reading of older speakers).

The main take-home message of the corpus data in this section is, I believe, that Glasgow /r/ can be realised in at least ten different ways, if we include vocalised variants. From these, three rhotic realisations and three types of vocalisation (complete vocalisation, rhoticised vowels, and velarised vowels) are fairly common. The existence of an intermediate category for /r/ vocalisation and a miscellaneous category for /r/ realisation also hint at the rampant variation present in the behaviour of this sound in English as spoken in Glasgow. The absence of salience, I will try to prove, has a lot to do with the extent of observable variation.

### 7.1.4 Studies on coda /r/

In order to further elaborate on the state of coda /r/ in Glasgow, I briefly discuss three papers on the subject. The first one is an auditory-acoustic study of postvocalic /r/ by Stuart-Smith (2007). The study looks at recordings of 12 working-class speakers reading a word list of 200 words in Glasgow in 2003. The speakers belong to four age groups, three adolescent ones (10-15 years old) and one adult group (40-60 years old). The focus is on six words ending in /ar/, followed by a consonant or word-finally. Four of these have a control pair without /r/.

The two main findings that emerge from the acoustic study of the data are that the pattern of variable coda /r/ vocalisation, observed in the 1997 data, and discussed extensively above, is present, and that the rhotic has an effect on the preceding vowel irrespective of whether it is articulated, which supports the assumption that the cues of /r/ spread more to its surrounding segments, previously proposed by Plug & Ogden (2000).
Concerning variation, the data show that adolescents use vocalised variants extensively, while the adult age group much less so. Furthermore, the vocalised realisations themselves vary extensively between plain vowels, uvular/pharyngealised, and velarised variants. It is remarkable that a large number of tokens are categorised as auditorily difficult, intermediate variants, and that the three transcribers of the study, all of them Scottish English native speakers, strongly differed in their judgements. This all hints at the gradience of derhoticisation in Glasgow: as above, in the 1997 corpus, we are confronted with a fine-grained phonetic change, which lacks a clear, near-categorical, bimodal distribution. Stuart-Smith notes that there is an increase in the use of vocalised variants as compared to the 1997 corpus, which is further evidence for a change in progress.

Coda /r/ has two notable effects on the preceding vowel. First, the vowel is significantly shorter when there is a following rhotic, irrespective of whether the rhotic is realised or not. That is, heart has a longer [a] than hat, even if it is realised with a vocalised /r/ (the two words have the same vowel in Scottish Standard English). Stuart-Smith points out, however, that this pattern is strongest with the one consistently rhotic speaker in the study, and least pronounced with the almost non-rhotic adolescents. The situation is the same with the vowel quality that correlates with following /r/, the F2 of the vowel (always [a]). It is significantly lower in words with a coda /r/, but this difference is stronger for the rhotic speaker than for the non-rhotic ones.

Both the changes in vowel length and quality suggest that the tongue retraction gesture accompanying the alveolar/retroflex realisation of the rhotic is still there even if the rhotic itself is no longer audible. Then again, the acoustic nature of the study allows for a fair extent of uncertainty, and generally emphasises the variability of coda /r/ realisation.

Lawson et al. (2011) look at the articulation of coda /r/ in Scottish English. Their study is part of a larger project aimed at establishing the methodology of looking at the physical articulation of speakers of various ends of the socio-economic spectrum (Scobbie et al., 2008). They argue that while previous
Background

studies confined themselves to the articulatory descriptions of a few subjects, the development of new techniques, most notably mobile ultra-sound devices, allows the recording of a wide array of speakers. What is more, extensive, sampled articulatory recording is required, as articulation plays a relevant role in sound changes, and the only way to understand this role is to create data samples representative of age, gender, and socio-economic background.

The authors themselves point out the issue we touched above, namely, that acoustic analysis of coda /r/ in Glasgow and in Scottish English in general seems to hint at an important role played by the articulation, and seems to be, by itself, insufficient to gain a thorough understanding of the phenomenon. The corpus they use for their study of derhoticisation was compiled in 2008, a collection of ultrasound tongue imaging recordings in eastern central Scotland. A preliminary finding was that, in a number of derhoticised tokens, the tongue-tip raising gesture (essentially, the articulation of [r]) is still present, but delayed beyond the offset of voicing, and, consequently, either weak or inaudible. These results concur with pre-[r] vowel behaviour in Stuart-Smith’s acoustic analysis, and also explain why the acoustic classification of derhoticised tokens can sometimes prove to be very difficult.

The analysis in the 2011 paper is based on word list recordings of informants from an Edinburgh fee-paying school and a state sector school in an economically deprived area of Livingston, a town between Glasgow and Edinburgh in the Central Belt of Scotland. Only tokens of the form Vr# were subjected to analysis, in order to avoid variation stemming from anticipatory co-articulation. The corpus contained 147 such tokens, of which 136 were used. The tokens were subjected to auditory and articulatory classification.

The auditory classification might be familiar from the studies discussed here above. It sets up five categories, running from weakly to strongly rhotic: no [r], derhoticised [r] ([Vt]), alveolar approximant [ɹ], retroflex approximant [ɾ], and ‘schwar’ ([ɤ]).

Articulatory classification defines four categories of tongue gestures (Lawson et al. 2011, p259):
Derhoticisation in Glasgow

- TIPUP describes an articulation where the overall shape of the tongue surface is either straight and steep, or a concave shape, suggesting retroflexion.

- FRONT UP describes an articulation where the tongue surface forms a smooth convex curve. There is no distinct bunching of the tongue front or dip behind the front region.

- FRONT BUNCHED describes an articulation where the front of the tongue has a distinctly bunched configuration (the tip and blade remain lower than the rest of the tongue front). A dip in the tongues surface behind the bunched section is also apparent.

- MID BUNCHED describes an articulation where the front, blade and tip are low, while the middle of the tongue is raised towards the hard palate.

Auditory classification gives the usual picture of working-class speakers using more vocalised variants, and middle-class speakers using more rhotic ones. The set-up of the auditory categories on a scale of rhoticity allows the calculation of the auditory strength or rhoticity score of each speaker. The mean scores for gender and background are shown in table 7.9. A two-way ANOVA shows a significant effect of both gender and social class on rhoticity scores, independently of each other. This, again, is consonant with the picture of derhoticisation in Scotland suggested by the previous studies.

<table>
<thead>
<tr>
<th></th>
<th>Working class</th>
<th>Middle class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>3.9</td>
<td>6.6</td>
</tr>
<tr>
<td>Female</td>
<td>4.1</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Table 7.9: Mean auditory rhoticity scores in Lawson et al. (2011)

The articulatory classification which the authors matched with the auditory phonetic description of the extent of vocalisation used the four categories
defined above. In general, working-class and male informants (the latter especially in the working-class group) use more tip up and front up articulations than middle-class or female ones. Tip up articulations are overwhelming in the working-class male informants’ production, while middle-class speakers almost exclusively use bunched articulations. The relationship between the use of tip/front up or bunched articulations and class as well as gender is statistically significant.

Lawson et al. (2011) point out that there is a significant correlation between tongue configuration classification and acoustic scores of rhoticity, which, differences in the detailedness of the two scales notwithstanding, strongly suggests that bunched articulations are associated with a stronger auditory rhotic effect when compared to other variants. That is, to put it simply, the working-class preference for tip up articulations runs parallel to the smaller extent of rhoticity in working-class speech.

Beside giving proof for the usefulness of ultrasound tongue imaging in sociolinguistic studies, Lawson et al. (2011) corroborate and extend the auditory/acoustic data on derhoticisation in Central Belt Scots dialects. Despite the study’s narrow scope, it is able to support that both gestural timing and tongue configuration play an important role in derhoticisation, and strongly match the corresponding acoustic cues.

Just like the corpus data, these results also confront us with an extremely fine-grained process of coda /r/ vocalisation. Apparently, from an articulatory point of view, /r/ vocalisation can be seen as a result of gestural off-timing. This means that the rhotic gesture still takes place (the token is, strictly speaking, rhotic), but it simply comes too late to be heard. To put it differently, some non-rhotic tokens still have the articulatory gesture associated with rhoticity, as well as some acoustic cues on the preceding vowel. The whole thing is closely linked to variation in /r/ realisation itself: the typical working-class variant of /r/, an alveolar flap or a central approximant, is much less audibly rhotic than the classic middle-class one, the retroflex approximant. What follows is that the working-class realisation may be connected, due
to articulatory necessities, to processes which lead to /r/ vocalisation much more than the middle-class one.

Another important aspect of derhoticisation as a derivative of gestural off-timing and varied tongue configurations (as discussed above) is that speakers might still associate a vocalised coda /r/ token with a performed tongue gesture. I.e. they would think that a vocalised coda /r/ is still there. If a working-class speaker of Glasgow Urban Scots tends to produce coda /r/-s with the respective tongue gesture but without any audible rhoticity, they could register similar derhoticised input tokens as linked to an (inaudible) tongue gesture (Koen Sebregts p.c.). The speaker would ‘reason’ that if he or she speaks rhotic (even if you cannot hear it) then other people also speak rhotic, even if he or she cannot hear that either. This has a two-fold effect on the process in general: first, the change will proceed more slowly, as the inaudible tongue configuration can linger on without any supporting acoustic cues of rhoticity, second, the difference between rhotic and non-rhotic tokens will be even more muddled, which renders them more difficult to use for social indexation.

Social awareness to /r/ in Glasgow has been put to explicit test in a study by MacFarlane & Stuart-Smith (2012). The authors were generally interested in the question whether phonetic variation, amply shown to be sensitive to social class, age, and gender, is accessible to local listeners, as well as in the dimensions along which listeners might classify available phonetic information. Their starting point was the ‘Glasgow Uni’ accent, an emerging middle-class supra-local dialect often identified with students of the University of Glasgow and the Glasgow School of Art.

The study relied on a matched guise test (Lambert, 1967) using four phonological variables and two constructed social personae, one associated with working-class brands, another one with middle-class ones (cf. the article for details). The three variables relevant here were (i) vowel quality before /r/ in words like ‘nerve, pearl, earth’, (ii) realisation of onset /r/ as a tap or a retroflex approximant, and (iii) syllable length in <-er> in disyllables, in
words like ‘number, clever, cover’.

Variable (i) touches on the issue of the SQUARE-NURSE merger, the merger of [ɛə] and [əː]. This is a regular feature of certain Northern English dialects, like Liverpool English, but did not happen to a large extent in Scotland [wells1982]. Therefore the innovative variant would be the merged one, where words like ‘nerve’ or ‘pearl’ are realised with [əː]. Clearly, the presence of the following /r/ has an influence on vowel quality, so a sensitivity to this difference hints at a general sensitivity to coda /r/. Variable (ii) directly relates to the studies discusses above: the alveolar tap is the dominant working-class variant and the retroflex approximant is the middle-class variant both in Lawson et al. (2011) and Stuart-Smith et al. (2007). In the data of the latter study, the 1997 corpus, /r/ realisation shows much more style shifting than /r/ vocalisation, though not as much as other characteristic variables, such as /l/ vocalisation or /t/ glottalisation.

Variable (iii) is complicated. We can recall from Stuart-Smith (2007) that the vocalic portion is longer in words with a coda /r/ irrespective of whether this /r/ is realised or not. However, the systematic difference of vowel length between words with and without an orthographic coda /r/ is strongest with speakers who are mostly rhotic in the first place. Lawson et al. (2011) get similar results, and point to gestural re-timing as a factor in derhoticisation, which again suggests that realised coda /r/-s go with a longer vocalic portion. However, in the MacFarlane & Stuart-Smith (2012) study, the compared variants are both rhotic, the only difference is in the length of the vocalic portion.

The stimuli consisted of word pairs recorded by the same person, a native speaker of Scottish English from Glasgow. One member of the pair was the first variant of the given variable, the other one the second. The listeners had to associate the heard variant to one of the two personae, one working-class and one middle-class young male (to put it simply). The experiment had 31 participants.

The results show that listeners judged the tokens belonging to one or
Derhoticisation in Glasgow

the other persona, without strong preference for either. Furthermore, a significant majority of the tokens were judged as expected, that is, belonging to one persona or the other depending on their realisation, which supports the authors’ hypothesis on the existence of a middle-class dialect, as defined by these variables (amongst others). If you speak in a certain way, you belong to a certain social stereotype, and the ways of speaking that define this social stereotype are, at least partly, salient to the listener.

The results for the three variables were as follows: There was a statistically significant difference in variables (ii) and (iii), that is, attributing [3ː] and the longer <-er> to the middle-class persona. Such a preference, however, was absent in the case of variable (i), as speakers did not associate taps and retroflex approximants with any of the personae in particular. That is, the middle-class persona would be the one who merges SQUARE with NURSE, and they would have a more pronounced vowel length difference before a coda /r/, a property related to more rhotic accents. The tap versus approximant distinction, however, was not associated strongly to either the working-class or the middle-class persona, contrary to expectations.

In the frame of the production data surveyed above, these results are not trivial to interpret. Sensitivity to vowel quality and quantity before /r/ strongly suggests that the perception of this segment is heavily linked to the previous vowel. This warns us against a clear segmental approach to rhoticity. Interestingly enough, when we look at the 1997 corpus data, it seems that the listeners themselves failed with the segmental approach, inasmuch as sensitivity to coda /r/ is absent. That is to say, variation in coda /r/ might only be perceived in a larger local context, that of the preceding vowel. This is further supported by the finding that listeners seem to ignore differences in /r/ realisation in the onset, where (at least in the experimental stimuli) there is no preceding vowel.
7.1.5 Interim Summary

Much ink has been spilt on the detailed description of /r/ in Scotland in this chapter. Before we go on to formulate and test a hypothesis on the speaker perception of coda /r/, it is useful to recapitulate the main points. The vocalisation of coda /r/ seems to show social stratification, and indexes speaker age, gender, and background in general. It is related to variation in /r/ realisation, with working-class speakers, who show a preference for an alveolar tap realisation of /r/, vocalising more coda /r/-s in general. It is also related to vowel length and vowel quality, which correlate with the presence of orthographic /r/ and coda /r/ vocalisation: the preceding vowel is shorter in words with a written final /r/, and this length difference (and the quality difference) is more pronounced in the use of speakers who are generally rhotic. Speakers are apparently aware of the vowel length difference according to one study in Glasgow, but do not show marked style shifting in coda /r/ vocalisation and are not aware of differences in variation in /r/ realisation either.

We can establish a link between these observations if we look at the source of the change, which is articulatory lenition, mainly the off-timing of the constriction gesture, the main source of rhoticity. Speakers who delay this gesture sufficiently have no audible \([r]\) in the coda at all, which is reinforced by the preference for taps over retroflexes, since the former result in a weaker rhotic in the first place. Curiously enough, speakers can consistently produce tokens with very limited or no auditory rhoticity but with the presence of the tongue gesture. This hints at the importance of articulation in the spread and transmission of the change. This is even more so if we accept that listeners can link the non-rhotic tokens in the input to their produced tokens – which have no audible rhoticity, but a present rhotic articulatory gesture. Furthermore, the cues of \([r]\) can spread to the preceding vowel, which complicates matters even more.

This complicated situation arises from two main sources: the articulatory base of the change, as well as the property of /r/ to exhibit its cues less locally
Derhoticisation in Glasgow

(i.e. more on the preceding vowel, for instance) in the first place (Lodge, 2009; Hawkins, 2003; Simpson, 1998). This means that a strictly segmental approach to coda /r/ in Scottish English, or in Glasgow in particular, is hopeless for two reasons, namely, the gradient nature of the change, and the consequent rampant variation in /r/ realisation.

What certainly complicates matters more is that derhoticisation is a loss of a property, which is obviously harder to perceive than a new property, such as a shift in vowel quality. Indeed, both Stuart-Smith (2007) and Stuart-Smith et al. (2007) note that rhotic, native speaker phoneticians found it very difficult to transcribe vocalised variants, as these are difficult to distinguish from each other and the rhotic variants. No wonder average speakers do not notice them at all. And yet, derhoticisation exists in the sense that it has discernible (albeit with much difficulty) variants which bear a social indexing role, and therefore cannot be dismissed as a case of rampant free variation.

Such a state of affairs would no doubt lead some to question the validity of the segment in phonological descriptions in the first place (Silverman, 2006), while others might argue that the segment is a useful heuristic tool, but not when dealing with /r/. Debates on the ontology of the segment, however, are by and large irrelevant to us, since we know that coda /r/ can be a salient marker, at least under different circumstances. In Labov’s Lower East Side study, the presence or absence of coda /r/ is the single most reliable marker of social status, as amply demonstrated by style shifting in coda /r/ realisation, as well as speaker attitudes, and even overt commentary.

In the first section of this chapter I argued that the difference between Glasgow and New York City cannot be solely blamed on the first being a bottom-up, and the second being a top-down change, since we know of other bottom-up changes which became salient for the listeners, /t/ glottalisation, extensively discussed in this book, being a prime example. The glottalisation of /t/ is a good example also in the sense that it is also grounded in articulatory/acoustic lenition, and is salient to the listeners despite all that.
7.2 Analysis

This section looks at data from the Fred corpus (Kortmann et al., 2005) and argues that the reason for the absence (or limited extent) of speaker awareness to derhoticisation is its extent of variation. This variation is certainly the artefact of the origin of the change, but by no means should be confused with it. Glasgow (rV) evades speaker recognition because, due to its history, it is highly variable, but that does not mean that it cannot stabilise in the future, and hence become a social marker.

We start with the discussion of a study of coda /r/ realisation in Scotland based on the Fred corpus. We then go on to establish transitional probabilities for coda /r/ and /r/ in general, based on the corpus data. Finally, we will investigate how derhoticisation can affect these values, and, consequently, the salience of the variable.

7.2.1 The FRED study

The aim of the study was to establish the extent of derhoticisation, that is, coda /r/ vocalisation, in a data set of Scottish English. The source of my data is the East Lothian part of the Fred corpus (Kortmann et al., 2005). This part consists of fifteen interviews recorded in the late 1970s in a high school in the town of Tranent, in East Lothian in Scotland, 18 km East of Edinburgh. The interviews, excluding the interviewer’s lines, consist of 33,700 words. The interviewees are all then students of the high school, 15-18 years old. There are fifteen interviewees, ten female and five male. Three appear in at least three interviews, and five in two interviews. The interviewer is the same person in all interviews except for the last one. Recordings are available for ten interviews, however, two of these are of poor quality and therefore they were not used in my study. The remaining recordings cover 23,400 words of interviewee speech.

The target variable was word-final /r/. The reason word-medial pre-consonantal /r/-s were excluded is that both Stuart-Smith (2007) and Stuart-
Smith et al. (2007) note the difficulty in establishing whether a rhotic in this position is realised or not. This means the exclusion of 887 potential realisations in the text. Compared to the 1687 word-final realisations, this is still a minority. The recordings were stereo wav files, sampled at 44100 Hz. Realisations were judged by the ear, using Praat (Boersma & Weenink, 2005) in more dubious cases. I put word-final /r/-s in *four categories*: [R] is a rhotic realisation (a tap, an approximant, or any other possible realised variant, as in [kar] *car*), [V] is a vocalised realisation (where no audible rhotic is heard, as in [kaː] *car*), and [R/V] is an intermediate variant, e.g. with a slightly rhoticised or velarised vowel, as in [ka~] *car*. Finally, [?] is a realisation that is inaudible, either due to ambient noise or an r-initial word following it. From the 1687 tokens, 1621 were transcribed. Table 7.10 shows the token counts with the original file names of the recordings. Table 7.11 shows the percentage ratios after the exclusion of tokens transcribed as ‘?’.

<table>
<thead>
<tr>
<th>Recording</th>
<th>R</th>
<th>R/V</th>
<th>V</th>
<th>?</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELN_004</td>
<td>139</td>
<td>15</td>
<td>16</td>
<td>8</td>
<td>178</td>
</tr>
<tr>
<td>ELN_005</td>
<td>47</td>
<td>5</td>
<td>5</td>
<td>11</td>
<td>68</td>
</tr>
<tr>
<td>ELN_006</td>
<td>80</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td>ELN_007</td>
<td>103</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>121</td>
</tr>
<tr>
<td>ELN_008</td>
<td>214</td>
<td>12</td>
<td>22</td>
<td>12</td>
<td>260</td>
</tr>
<tr>
<td>ELN_009</td>
<td>386</td>
<td>14</td>
<td>26</td>
<td>10</td>
<td>436</td>
</tr>
<tr>
<td>ELN_010</td>
<td>113</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>127</td>
</tr>
<tr>
<td>ELN_011</td>
<td>123</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>135</td>
</tr>
<tr>
<td>ELN_012</td>
<td>187</td>
<td>6</td>
<td>13</td>
<td>2</td>
<td>208</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>1,392</td>
<td>75</td>
<td>104</td>
<td>54</td>
<td>1,625</td>
</tr>
</tbody>
</table>

Table 7.10: Rhoticity in FRED ELN: Tokens

As the results show, the data are overwhelmingly rhotic. This is in line with Romaine’s 1978 study, comparable with respect to both age and location. With these data, we have a written corpus with the extent of derhoticisation
<table>
<thead>
<tr>
<th>Recording</th>
<th>R</th>
<th>R/V</th>
<th>V</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELN_004</td>
<td>81.765</td>
<td>8.824</td>
<td>9.412</td>
<td>100</td>
</tr>
<tr>
<td>ELN_005</td>
<td>82.456</td>
<td>8.772</td>
<td>8.772</td>
<td>100</td>
</tr>
<tr>
<td>ELN_006</td>
<td>88.889</td>
<td>5.556</td>
<td>5.556</td>
<td>100</td>
</tr>
<tr>
<td>ELN_007</td>
<td>88.793</td>
<td>5.172</td>
<td>6.034</td>
<td>100</td>
</tr>
<tr>
<td>ELN_008</td>
<td>86.290</td>
<td>4.839</td>
<td>8.871</td>
<td>100</td>
</tr>
<tr>
<td>ELN_009</td>
<td>90.610</td>
<td>3.286</td>
<td>6.103</td>
<td>100</td>
</tr>
<tr>
<td>ELN_010</td>
<td>91.129</td>
<td>4.839</td>
<td>4.032</td>
<td>100</td>
</tr>
<tr>
<td>ELN_011</td>
<td>91.791</td>
<td>4.478</td>
<td>3.731</td>
<td>100</td>
</tr>
<tr>
<td>ELN_012</td>
<td>90.777</td>
<td>2.913</td>
<td>6.311</td>
<td>100</td>
</tr>
<tr>
<td>Average</td>
<td>88.056</td>
<td>5.409</td>
<td>6.536</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7.11: Rhoticity in Fred ELN: Ratios

![Diagram showing rhoticity in Fred ELN: Tokens](image)

Figure 7.4: Rhoticity in Fred ELN: Tokens

(not marked in the orthography) established separately. Preliminary analysis of the data suggested that there is no difference in coda /r/ vocalisation before
vowels and consonants, so this distinction was not taken into consideration. At this level of rhoticity, such a difference could not be overwhelming in any case.

7.2.2 Transitional probabilities in coda /r/ realisation

The next step is to establish the transitional probabilities of a word-final /r/ in a rhotic dialect, based on these data (word-medial pre-consonantal realisations of /r/, are, due to the difficulties they present, ignored). In order to have more robust results, I used all the transcripts of the East Lothian interviews, even though useable recordings are only available for eight of the fifteen in total. This procedure, however, is warranted by the following: first, the same speakers feature in both the transcripts with and without recordings. Second, the extent of coda /r/ vocalisation is homogeneous in the recorded interviews – a Chi Square test of independence confirms that there is no significant variation between the recordings themselves (p=0.18). This permits the assumption that the extent of coda /r/ vocalisation is more or less the same within this speaker community, therefore the extents of vocalisation gauged from the existing interviews can be applied to the other transcripts as well. Since the differences between stressed and unstressed pre-pausal /r/ vocalisation are not that pronounced in the 1997 Corpus, I ignored this distinction and simply used the average – things would get complicated later on in any case.

Table 7.13 shows transitional probabilities based on the written part of the corpus and the derhoticisation study presented above. The first number is the probability of a coda /r/ following a vowel: this is a general measure of derhoticisation, based on the frequency of <Vr> final words. The second one is the probability of a word-final vowel followed by an [r]-initial word. This measure is based on the connected speech sequences and is independent of derhoticisation, which does not affect onset /r/ realisations. The third number follows from the previous two: it gives the general transitional probability of an [r] following a vowel, irrespective of whether there is a word boundary
in between. The fourth and fifth numbers give the probability of a coda [r] or any [r] followed by a vowel. In the former case, this is based on realised coda /r/-s followed by vowels, in the latter, on any realised /r/ followed by a vowel, including onset [r]+vowel sequences. The last number is, again, a general measure of derhoticisation, as it gives the probability of a coda [r] being followed by a consonantal onset.

<table>
<thead>
<tr>
<th></th>
<th>V Cod /r/</th>
<th>V Ons /r/</th>
<th>V /r/</th>
<th>Cod /r/ V</th>
<th>/r/ V</th>
<th>Cod Ons C /r/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fred</td>
<td>0.047</td>
<td>0.010</td>
<td>0.056</td>
<td>0.329</td>
<td>0.693</td>
<td>0.575</td>
</tr>
<tr>
<td>ELN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.12: Coda [r] transitional probabilities in Fred East Lothian

These numbers are, of course, not interpretable in isolation. In order to provide a basis of comparison, I calculated the probabilities for the same environment based on the same data (the written transcriptions of Fred ELN) but drawing on the results on working-class youth behaviour in derhoticisation in the 1997 Glasgow Corpus. According to Stuart-Smith et al. (2007), young working-class speakers realise 15 per cent of the rhotics in spontaneous speech, and, crucially, this is not heavily dependent on the phonetic environment: that is, for instance, whether the coda /r/ is followed by a consonant or a vowel. This is a detail to which we will later return. If we assume, in a somewhat abstract manner, a 15 per cent general rhotic realisation, we get the set of contrasting numbers for the two dialects as seen in Table 7.13. The columns are the same categories as above, and the lines are Fred ELN, the Fred-based data from Table 7.13, and Glasgow WCY, the transitional probabilities based on the Fred corpus and the 1997 Glasgow study.

It is useful to dwell on these numbers for a while. To recapitulate, what we see here is the difference, based on my Fred data and a study in Glasgow, between two dialects which have different extents of derhoticisation. There is, however, negligible context-dependent variation. (That is, the extent of coda
Table 7.13: Coda [r] transitional probabilities in Fred East Lothian and Glasgow working-class youth conversations data

<table>
<thead>
<tr>
<th></th>
<th>V Cod /r/</th>
<th>V Ons /r/</th>
<th>V /r/</th>
<th>Cod /r/ V</th>
<th>/r/ V</th>
<th>Cod Ons C /r/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fred ELN</td>
<td>0.047</td>
<td>0.010</td>
<td>0.056</td>
<td>0.329</td>
<td>0.693</td>
<td>0.575</td>
</tr>
<tr>
<td>Glasgow WCY</td>
<td>0.007</td>
<td>0.010</td>
<td>0.017</td>
<td>0.330</td>
<td>0.919</td>
<td>0.574</td>
</tr>
</tbody>
</table>

/r/ vocalisation is identical before consonants and vowels.) The numbers are abstract in the sense that they do not come from actual speech patterns, but this is saved by the fact that all sources are instances of spontaneous speech.

What is clearly affected by the different extents of /r/ vocalisation is the distributions of vowel + [r] sequences. The probability of a vowel followed by a realised [r] (third column) is three times higher in the more rhotic dialect than in the less rhotic one, mainly because, in the latter, the only main source of V+r sequences is V-final words being followed by [r]-initial ones, unaffected by coda /r/ vocalisation. These are, however, in a minority compared to word-medial V+r sequences, most of which are absent in the less rhotic dialect (cf. the first column).

Naively one would assume that the best predictor of derhoticisation is the probability of coda [r]-s followed by a consonant. This is because, in ‘proper’ non-rhotic dialects, this is near zero, as rhotics are only realised before vowels – and not even there categorically, at least in connected speech (Sóskuthy, 2010). However, in a change such as Scottish English derhoticisation, where an articulatory lenition affects all environments more or less equally, this effect disappears.

We could then argue that the difference between the probabilities of [Vr] sequences in the two dialects is still a source of salience: more rhotic speakers (such as the middle-class informants of the Stuart-Smith et al. (2007) study) would pick up on this and promote coda /r/ vocalisation to the status of a salient marker. This is, however, where the extent of variation in /r/
vocalisation comes into play. This is a factor that is completely unseen in the table above. Essentially, in the Fred data, coda /r/ behaves, to an overwhelming extent, in one way: it gets realised. In the Glasgow working-class youth data, it also has a majority pattern: it gets vocalised. However, this pattern is contrasted with several other ones: it can realise as a rhotacised vowel, a a velarised vowel, or an intermediate variant. The main source of salience, a stark contrast between two distributions, withers away in the presence of rampant variation. The relevant scores are repeated in Table 7.14.

<table>
<thead>
<tr>
<th></th>
<th>Fred ELN</th>
<th>Glasgow WCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>88.05</td>
<td>25.04</td>
</tr>
<tr>
<td>r/V</td>
<td>5.40</td>
<td>63.42</td>
</tr>
<tr>
<td>V</td>
<td>6.53</td>
<td>1.86</td>
</tr>
<tr>
<td>V°</td>
<td>n.a.</td>
<td>16.33</td>
</tr>
<tr>
<td>Vr</td>
<td>n.a.</td>
<td>4.43</td>
</tr>
</tbody>
</table>

Table 7.14: Variation in coda /r/ vocalisation

In sum, the Fred study, with its subsequent comparison to the Glasgow data, hints at two possible causes for the lack of salience in Scottish derhoticisation. First, in the absence of context sensitivity, the established patterns will not be largely different, even with notably distinct extents of derhoticisation. Second, even this difference can be masked by a larger extent of variation in /r/ realisation, which also illustrates how transitional probabilities might be misinterpreted when taken out of their larger context.

### 7.3 Concluding remarks

This chapter looked at derhoticisation in Scottish English and the question of salience. I argued that derhoticisation, the vocalisation of coda /r/, an ongoing process in the Central Belt dialects of Scots, shows very limited (if any) speaker awareness, as shown by the lack of a large extent of style shifting in
Derhoticisation in Glasgow

data from both Glasgow (Stuart-Smith et al., 2007) and Edinburgh (Johnson, 1983), and by the apparent difficulty of middle-class Scottish transcribers to notice it. This should come as a surprise for the theory advocated in this work, namely, that the source of a dialectal variable’s salience is the difference in distributions in related dialects that the variable’s behaviour entails. At a first glance, /r/ vocalisation, which clearly relates to the distributions of /r/ realisations in the speech signal, should be a prime candidate for salience.

I went on to claim that the lack of salience in this case can be linked to the variable’s large inherent variance – this has been shown by Timmins et al. (2004); Stuart-Smith et al. (2007). Coda /r/ can surface in many different ways, which means that speakers have a hard time in noticing the differences that follow from the use of any of these variants. The reason for the variance is the source of the sound change, articulatory lenition. Articulatory and acoustic studies like Lawson et al. (2011) and Stuart-Smith (2007) strongly suggest that derhoticisation is best viewed as a gradual, phonetically fine-grained gestural reduction, which entails variation in coda /r/ realisation, small differences between the individual variants, and the dispersion of the cues of the realised /r/, with the previous vowel playing an important role in its identification.

Gestural reduction as the source of the change can be linked to its bottom-up nature: it is not a norm enforced on the speaker community by the upper classes, but rather a gradient change slowly percolating through the entire community. (It has to be stressed at this point that though real-time and apparent-time data suggests the spread of non-rhoticity, we have no solid evidence on the change sweeping through all registers, leaving Scottish dialects without coda [r]-s.)

Based on a study on the East Lothian part of the Fred corpus (Kortmann et al., 2005), as well as the studies cited above, I investigated the transitional probabilities tied to the different extents of derhoticisation observed in Scotland. The starting point was the study of derhoticisation in Fred, which, around 88 per cent, agrees with similar early studies of derhoticisation
Concluding remarks

(Romaine, 1978; Macafee, 1983). I compared this level of rhoticity with working-class youth behaviour in the Glasgow corpus.

The first main finding was that though there is a difference in transitional probabilities, mainly in Vr sequences, that speakers can rely on, it is, in general, much smaller in extent, chiefly because derhoticisation, as a gradual phonetic change, is context-independent. Therefore, typical rhotic and non-rhotic patterns of [r] distributions before consonants and vowels are much less pronounced.

The second main finding was that the extent of variation associated with larger derhoticisation effectively masks the probability difference itself. Since coda /r/ vocalisation takes place gradually, its large extent entails a large amount of variant realisations, which makes it more different to establish a comparison in distributions (for the speakers of two dialects) and, based on the difference, use the variable as a social marker.

This effectively means that segmental distributions as criteria and indicators of salience have to be used with special care for two reasons: first, segments as abstractions have to be established especially cautiously when handling a gradient phonetic change.

This need not mean that segments should be eschewed in general, particularly because a segmental approach allows the use of large corpora and leads to more robust findings. With proper heed, we can define abstract segments that give a good fit on the data. Second, it is not enough to compare the segmental distributions related to a particular variable in two dialects – it is also important to look at the competing variants in a single dialect, as the more variants are possible, the less the difference in distributions for any two of them will be noticeable.

It is at this point that two comparisons are in order. Derhoticisation in Scotland can be readily compared to rhoticisation in New York City (Labov, 1966/2006). The differences, though, are now fairly clear. The headway of coda [r] is a novel standard spreading top-down in New York City. This has two important consequences. First, /r/ realisation will be much more
Derhoticisation in Glasgow

categorical: coda /r/ will not slowly come into being, but are, rather, adopted. This does not mean that language users categorically switch from non-rhotic to rhotic behaviour, but it reduces phonetic variation. Second, the basis for the change is a different dialect, which compares better to the non-rhotic dialects of the South of England than to derhoticised Scottish English, as it has ‘linking-r’, a consistently larger extent of coda /r/ realisation before vowels than before consonants. Both these effects are absent in Scots derhoticisation, and the concomitant variability is both wide and large in extent enough to mask differences in segmental distributions and hence salience in general.

This is a proper illustration of the relationship between salience and context-sensitivity, a property that follows from the model adopted in this book. If a language change affects segmental distributions evenly, it causes smaller transitional probability differences, and is therefore much less likely to be noticed – as shown in this chapter. Indeed, we have data (Llamas et al., 2009) that the rhoticity difference is salient between Scottish and English English dialects: speakers in towns near the Scottish-English border are overtly sensitive to coda /r/ realisation, which is displayed both in their attitudes and speech behaviour.

Crucially, lack of context-sensitivity (and the subsequent lack of salience) is not linked to a change being top-down or bottom-up (contrary to Labov 1966/2006). True, a top-down change can proceed in such a way that different use in different contexts, and the following probability differences, will always be present. This is the New York City example. A bottom-up change, which is often phonetically gradual, can proceed at an even pace in all environments, hence lacking pronounced probability differences. This is the Scottish example. However, a top-down change can affect a whole category, which might lead to absence of salience, as in the Philadelphia vowel shift (cf. Labov 1994). A bottom-up change might be sensitive to phonetic environment in the first place, like glottalisation, where word-final pre-consonantal and pre-vocalic behaviour split fairly early. It can also reach an equilibrium where variation turns into alternation triggered by the phonetic context, like Southern English
non-rhoticity, where coda /r/ is realised before vowels but is absent before consonants with a near consistency.

Pertaining to the theory of salience presented in this work, a look at derhoticisation is useful in showing that a segmental approach, when handled with due care, can not only be applied to phonetically fine-grained changes, but it can also have long-range predictions on the relationship of context-sensitivity, extent of variation, and sociolinguistic salience.
Chapter 8

Salience and language change

In the previous chapters, we saw examples of the link between salience and differences in segmental distributions for variables that carry social indexation and variables that do not seem to do so. In this chapter, I go on to examine the role of social indexation (and, consequently, salience) in language change. Since studies on both sociolinguistic salience and the connection between social language use and language change generally concentrate on the phonological domain, and since all my case studies concentrate on phonological variables, this chapter focusses on sound change in particular.

This work argues that certain sound patterns are salient. In Chapter 2 I cite evidence that humans are sensitive to differences in frequency distributions – measured in transitional probabilities – and that this sensitivity is carried over to sound patterns in natural language. In Chapter 3 I introduce a method of modelling whether a certain dialectal pattern is different, ergo salient for the speaker of another dialect. The case studies in chapters 4-7 provide examples of this.

A salient variable can be a potential carrier of social indexation, or, to borrow from Labovian terminology, a marker – this is contrasted with a dialectal difference apparently neglected by the speakers, an indicator. Markers behave differently than indicators in various ways: Speakers try to manipulate them dependent on situation and register, the most typical case
being the switch to the relevant standard in a formal situation. Speakers also have implicit attitudes towards the use of markers, which can be empirically tested for, and, in extreme cases, they can point to them when describing a dialect.

The marker-indicator distinction is just one way of abstracting over listener awareness of a particular dialectal variant, which awareness is necessarily gradient in nature. Furthermore, listeners are usually alert to whole dialectal profiles (cf. the enduring geographical dialect stereotypes, such as Brummie, Scouse, Geordie, or Cockney in England), rather than individual variation patterns (Labov, 1994). Still, the marker-indicator distinction is a useful way of thinking about research carried out on listener awareness of sound variation and change, one that can be easily carried over to simulations of sound change as well.

Why is salience interesting? The way people chart social space using language is relevant to the sociologist (and the dialectologist) and the extent to which they can differentiate how other people sound can be exciting to a phonetician or a speech therapist, and to a small extent even to a phonologist. What language users make of individual differences, however, would not appeal to a theoretical linguist unless it was related to some general linguistic property. Below I will argue that this is the case, and that salience contributes to our understanding of language change. My theory of salience is based on relative differences, differences in frequency, and this entails that these frequency differences also count in language change – a case of very low level attributes (how a dialect ‘sounds’, basically) influencing very complex processes (change in language).

In this chapter, I first go through evidence on the connection between group indexation and sound change. The two main sources of this evidence are empirical studies and simulations. We will see that the available body of facts suggests a view of sound change where the individual’s attempt at indexing identity and interactions between individuals play a relevant role in shaping the course of sound change. Second, I will discuss a modelling set-up
in which an explicit distinction is made between variation with and without
indexing value. Third, I give two examples from studies in this work which
illustrate how variants that carry social indexation behave differently from
variants that do not do so in the course of sound change.

In Section 8.1 I look at the literature on the effect of speaker indexation
on sound change. We will see that this effect, altogether discarded in early
work on sound change, comes more and more into the foreground in later
studies. While Section 8.1.1 surveys theoretical work and field studies, Section
8.1.2 surveys computer simulations of change in which speaker indexation
plays a relevant role. Particular attention is given to the model of language
change by Baxter et al. (2009), an implementation of the evolutionary theory
of Croft (2000). Section 8.2 connects this model to two of the case studies in
the present work in order to give a concrete example of the role of speaker
indexation, and, indirectly, salience, in sound change. Section 8.3 gives a
summary.

8.1 Speaker indexation in sound change

The fact that salient variables which carry social indexation (that is, markers)
are used differently by the speakers does not directly imply that they behave
differently in sound change. This claim has two premises, first that salient
variables do carry social indexation, and second that social indexation shapes
sound change.

I would be brief on the first premise: On one hand, salient, socially
stratified variables are likely to become social markers. On the other hand,
social markers, wherever the probability-based theory of salience is applicable
to them (see the Scottish example in Chapter 7), should always be salient,
i.e. entail a low transitional probability in the speech signal.

The second premise, that social indexation is connected to sound change –
not to mention guides it – is less obvious.
8.1.1 Approaches to speaker indexation

The Neogrammarians see sound change as an unstoppable force working above the speakers’ heads (Paul & Wundt, 1886/1995), wholly resistant to influence by any individual who uses it and independent of language use as a whole. This tradition is upheld by the American Structuralists (Sapir, 1921; Bloomfield, 1933), who regard sound change crucially not only as untouchable but also as invisible to the speaker.

The Structuralist reluctance to acknowledge the speaker in any way partly follows from Saussure’s 1916/1995 *Langue–Parole* distinction: this distinction expects the linguist to draw a line between the linguistic structure, perceptible to all sorts of influences, and the people who actually use it to communicate. Disenfranchising language use is also partly a reaction to unfortunate earlier claims on the direct effect of the climate and social development on linguistic structure (cf. comments in Trubetzkoy 1999).

Weinreich et al. (1968) and the variationist sociolinguist school established in their wake, including works like Labov (1972b), challenge the notion of sound change as a law of nature. This, in turn, allows them to consider the speaker as an actor and social indexation as a potential factor. They are, however, cautious with the subject, to say the least. In their opinion, the social values associated with variables have an apparent influence on the course of the change, but mainly in a robust and deterministic manner. Sound change, when vulnerable to speaker influence, is steered by the prestige of certain speaker groups and the larger community’s adherence to the local norm.

The classic variationist sociolinguistic tradition, to which this work owes much, associates particular linguistic variants with membership of broadly defined speaker groups, usually being part of a clear social hierarchy. The preference of a variant, then, can signal affiliation to speaker groups of various types. Examples include a rooted, geographically outlined speaker community, as in the case of the centralisation of the diphthongs [ai] and [au] in Martha’s Vineyard (Labov, 1972b), a wide social class, as in the case of the use of a stop
variant for the standard dental fricative in New York City (Labov, 1972a), an age or a gender group, or most likely a combination of these, as in the use of the high rising terminal intonation pattern with young female speakers in New Zealand (Britain, 1992), or the young male working class speakers’ preference for the vocalised variant of /r/ in Glasgow (Stuart-Smith, 2003).

The language use of certain social groups, typically those with a better social standing, has overt prestige – it is often regarded as the standard of the language variety, which results in the overt or covert stigmatisation and avoidance of the non-standard variants – as discussed at length in the present work. Non-standard, vernacular variants, however, are attributed covert prestige (Trudgill, 1974), which would explain why they persist in spite of the pressure of the standard.

The main contributions of variationist sociolinguistics is the now well-established fact that sound change is observable in progress, along with the valuable insights to its mediation in the speech community, guided by prestige. Although there is a tacit understanding that overt prestige is associated with correctness, education, and high social standing, whereas covert prestige with solidarity, friendliness, and neighbourhood identity, Labov (2001) duly notes that evidence on the existence of covert prestige is controversial. He quotes attitude studies from Philadelphia to show that variants allegedly enjoying covert prestige never have the upper hand in attitude tests, irrespective of what aspects subjects are given to score them. That is, in these studies, a speaker voice using a vernacular variant with alleged covert prestige would score low not only on ‘educated’ but also on ‘friendly’ or ‘empathetic’.

In comparison, the existence of overt prestige is copiously supported by empirical evidence, including self-reports, popular stereotypes, and attitude tests (cf. e.g. Labov 1966/2006). Indeed, as Labov notes, the only solid evidence for covert prestige is the stubborn refusal of urban vernacular dialects to die out (cf. Williams & Kerswill 1999).

In the classic variationist view, the speaker is not much more than a vessel of sociolinguistic factors, and the personal influence over the choice of identity
Salience and language change

is, to a large extent, underestimated. (Cf. the notion of the ‘cultural dupe’ Giddens 1979.) Social practice in this case comes down to adherence to the norm.

Social structure and the role of the speaker are connected in classic variationist sociolinguistic research, which typically concentrates on urban centres (such as Trudgill (1974) on Norwich or Labov (1966/2006) on New York City). To put it simply, society is a big ladder that speakers try to climb. The use of the terms ‘overt’ and ‘covert’ prestige mirror this, as does the typology of sound change proposed in these works. Change from above is a conscious change that is introduced by the leading group, who hold onto the overt prestige form. Change from below creeps in undercover, originating in a lower, non-prestigious social group. Change from above will then be robustly categorical in structure, while change from below is gradient and phonetically fine-grained. This issue has been discussed in detail in Chapter 7.

Potential inherent weaknesses of the variationist tradition are highlighted by the work of authors like Milroy (1980) and Eckert (2000). They both develop on the classic model based on the primacy of social class-based speaker categorisation and put a larger emphasis on speaker organisation into social networks. According to both Milroy and Eckert, language use is better seen as guided by identity formation in the field of communities of practice, social grouplets organised around a particular shared activity, hobby, profession, or habit, and not solely by broadly defined societal units specified by economical status, age, or gender. This is true even more so since some of these categories, like economical status or especially gender are difficult to define in broad terms. To take the example of gender, on one hand, binary categorisation of gender membership (female/male) is oversimplified and potentially masks finer patterns of variation present which depend on, for instance, the extent to which an individual assigns herself or himself to a gender group. On the other hand, even though gender is always investigated as a subordinate factor to social class, in cases it can account for more variation in sound change alone than in conjunction with class (cf. e.g. Cheshire 2002).
A network-based approach is also supported by the fact that, contrary to what the classic variationist model of urban language use would lead us to believe, many competing norms can emerge in a community. This is apparent, for instance, in the Stuart-Smith et al. (2007) study in Glasgow. In this study we find cases where the vernacular variants are played up in formal interview settings, and this can be seen as an open emphasis on the norm, but, in this case, the vernacular norm.

Network-based studies take it for granted that identity plays a role in the linguistic behaviour of any particular social network. This involves adherence to or rejection of a norm variable and the use of innovative vernacular variants to define the group with respect to other groups, and to express group solidarity.

In sum, speaker indexation is gaining significance in field studies on the social context of sound change. The interaction between speaker indexation and change is also apparent in simulations used to lay out and investigate the mechanics of language variation and change.

8.1.2 Simulations on the role of indexation

This section details two simulations of language change. The reason I talk about them at length is that the first one, by Roberts (2010), gives an elegant account of the emergence of markers in language use, while the second one, by Baxter et al. (2009), presents a framework in which the effect of speaker indexation — and, indirectly, salience — on sound change can be assessed directly.

Roberts (2008, 2010) sets out with an evolutionary approach to language change. He argues that, from an evolutionary point of view, cooperation — a typical human trait — is beneficial, since groups are better at utilising available resources than separate individuals. Group behaviour, however, brings forward the ‘free-rider problem’: individuals could potentially travel from group to group, exploiting the groups’ altruistic, coöperative behaviour, and exploiting each, only to move to the next group when the previous one
recognised their selfish behaviour and excluded them.

The best way to battle free riders is to create group markers which identify group members, and therefore also outsiders. This relates to the original assumption on the multi-facetedness of communicative utterances: an utterance can signal that an individual demands certain resources, but it also signals the individual’s membership in the group, with the potential consequence that, as a group member, the individual is obliged to return the favour later. Roberts notes that language is one of the best identity markers: it is robust, reliable, takes a long time to develop, and is, therefore, resistant to changes.

His main research question is whether linguistic group markers develop naturally, during the course of group interactions, or whether identity marking also has an effect on them. This is precisely the question, albeit formulated differently, that I talk about above. In order to answer it, he develops an experimental paradigm of group interactions, based on an artificial language and a competitive situation in which resources have to be shared.

During the experiment, participants have to learn an artificial language first. This is required because what he is interested in is whether they would develop group markers, so he has to exclude the participants’ native language, already riddled with these. After this first step, the experiment consists of a number of rounds. Each round starts with interaction: participants command a pool of resources, and they are paired with another participant to discuss trading these resources. Giving someone resources multiplies their value, so participants are better off when they trade with each other rather than sit on resources and not share with anyone else. The discussion phase is followed by actual trading with the previous discussion partner, where participants can decide whether they oblige to their requests or ignore them. Finally, participants get a summary of trade in that round.

The experiment has two independent variables. The first one is competitive versus cooperative: In the competitive condition, players are divided into two teams. Team scores are calculated at the end of each round, and the team
with a highest final score wins. In the coöperative condition, all participants belong to the same team, and the sole aim is the accumulation of resources (for details, see Roberts 2010). The competitive condition, in which participants cannot tell if they are talking to a team member or an ‘enemy’ in the discussion phrase, encourages the use of the artificial language to signal team membership beside its basic purpose of communication. This pressure is absent in the coöperative condition. In the competitive condition, participants are asked after each round whether they were able to recognise their team mates.

The second independent variable is frequency: in the low-frequency condition, the participants have 15 rounds, in the high-frequency one, they have 30 rounds.

Roberts has two predictions for the competitive condition, (i) that players would be able to identify their team-mates using only linguistic cues, and (ii) that the alien language, identical for all players at the beginning of the game, would diverge into dialects.

The main results confirm these predictions, but only in the high frequency condition. In this condition, players do significantly better than chance at recognising their team mates when asked after each round, and this also shows in their habits of resource exchange. Dialect divergence, that is, the increase in the number of variants used for each word, is clearly observable in the high frequency condition.

Roberts goes into detail about this dialect divergence, and the various communicative strategies used by the participants to recognise their team mates. What is relevant to us is that, in his experimental setting, competitive behaviour results in the rapid development of linguistic markers. This strongly suggests that individual recognition/individual self-assertion plays a relevant role in the emergence of linguistic markers, and, consequently, in language change in general.

Roberts’ findings underline the importance of speaker indexation in language change, supporting the view of sound change proposed in network-based sociolinguistic studies, such as Milroy (1980). They also provide evidence
on how speakers utilise variation that arises from independent circumstances (such as the noise in transmission and learner errors) to create group identity markers, a process acknowledged by, amongst others, Foulkes & Docherty (2006). Speaker indexation and sound change are then inseparable. There is, however, another important question that Roberts does not address, namely, the way variants that carry social indexation behave differently from those that do not – the consequences of the marker/indicator distinction on sound change.

This is a core problem in the simulation discussed in Baxter et al. (2009), which aims at modelling the emergence of New Zealand English. Before we are able to see how their model works, we have to briefly look into both their theoretical foundations and the particular example they set out to model.

The network-based model of language change that their simulation takes as a premise is proposed by Croft (2000). Croft takes parallels between biological evolution and language change as his starting point, positing interactions between individuals as the source of both the actuation (emergence) and propagation (spreading) of novel variants.

Starting on an evolutionary premise adopted partly from Dawkins (1976) and Hull (1988), Croft establishes two principal units of change. The replicator is a piece of structure replicating itself, while the interactor is the entity that interacts with its environment in such a way as to cause replication to be differential. In biological evolution, the replicator is the gene, and the interactor is the organism, which has a certain chance to produce offspring, partly affected by its adaptation to the environment, and partly affected by pure chance.

In language, Croft argues, the replicator is the utterance, a multi-structured linguistic unit used to express a certain type of convention. The interactor is the speaker, and replicators, that is, utterances, are replicated during face-to-face interactions: one speaker, based on her linguistic competence, formulates an utterance, which is then processed and stored by the listener. Crucially, this is indirect replication, affected by noise (in a very
Speaker indexation in sound change

concrete sense, here) and the listener’s linguistic competence. This extensively reduces the chance of successful identical replication, explaining why language change proceeds much faster than biological evolution. (Another obvious reason is that utterances replicate themselves – in Croft’s sense – much quicker than genes do.)

There are obvious differences between the behaviour of a gene and an organism as replicator and interactor *vis-à-vis* utterance (or set of structure) and speaker. Assuming an exemplar-based cognitive framework, relying on spreading activation (e.g. Rumelhart & McClelland 1987), it is reasonable to take the imperfect replication of utterances as the source of linguistic innovation – and actuation. The propagation of a linguistic variant is a different issue, because, though innovation, in its strict sense, is omnipresent (no utterance is repeated twice *verbatim*), this does not lead to unchecked variation in the language.

Croft talks in detail about the actuation, the origin, of a new linguistic variant. This work, however, largely ignores the actuation problem. This is true for this chapter as well. It is intuitive to think that salience comes into play when a variant is transmitted, since this is precisely the point in time when the listener can notice or ignore it. Therefore, below, I concentrate on propagation in Croft’s model.

In this model, two main factors affect propagation: On one hand, the *frequency* of an actuation will have an effect on the success and speed of its propagation. The frequency of listener exposure to a particular innovative variant depends on both its probability to emerge during innovation, and the listener’s position in the speech community – the listener might be a member of a social network that uses this variant prevalently. Only taking frequency into account, the success of a new variant is mostly due to chance. The same innovation can spread quickly if proposed in a well-connected speaker group and fade if proposed in a small one. Certain types of changes occur with a higher chance and are more likely to be transmitted successfully than others (cf. Keller 1994). This factor, however, is below the detail of the model
Salience and language change

discussed above.

On the other hand, innovative variants might have social value attached to them, and this promotes (or hinders) their transfer by listeners. Simply put, an innovation can propagate faster if it is frequent in the first place, and if it sounds cool.

Baxter and colleagues build on this premise and note that the spread of a novel variant can be modelled in two ways. Since language is used in a community, a speaker is more likely to interact with certain speakers than others, depending on the speaker’s social network. This, in turn, can depend on geographical and social factors. What follows is that even if we do not order utterances based on social value, the frequency of interactions can give an advantage to certain utterances over others. This is what they call neutral interactor selection. This stands in comparison with weighted interactor selection, where utterances can be preferred or avoided based on their social value. In the case of neutral interactor selection the only effect on the propagation of a variant is its frequency, in the case of weighted interactor selection, both frequency and prestige (social value) play a role. (Note that frequency here is a complex concept: if variants A and B have the same token frequency in the beginning, variant A can spread while variant B can die out if, for example, variant A originates in the largest and best connected speaker group, while variant B is present in more, peripheral, smaller speaker groups simultaneously. For details on the mechanics of the model, cf. Baxter et al. 2006, 2009.)

The reason they propose this distinction is that they want to test a claim made by Peter Trudgill (cf. e.g. Trudgill 2008), that the emergence of New Zealand English, like all other colonial English dialects, is purely deterministic, without any influence of what people at the time of dialect formation regarded as prestige forms. Though Trudgill mostly uses ‘prestige’ in the sense of the prestige carried by the Standard English variants of the time, his argument is clearly all-encompassing: simply put, the emergence of a colonial English dialect, such as New Zealand English, is only attributable to the background,
number and geographical distribution of its speakers – from a sociolinguistic point of view, purely to chance.

Baxter et al. (2009) take on this claim and model the emergence of New Zealand English using Croft’s terminology. Translated to this terminology, Trudgill’s prediction would be that this dialect emerged through neutral interactor selection only. They find that neutral interactor selection alone is insufficient for the dialect to reach its present state in the allotted time frame. What follows is that, at least to some extent, weighted interactor selection also has to play a role in its emergence.

To be precise, the actual results of this model – for a fairly specific situation, dialect emergence in isolation – is much less relevant to us here than the mechanics used for the simulation. Baxter and colleagues model the spread of indicator and marker variables in a speaker network, and this is an ideal basic set-up for a model of the role of salience in sound change.

The main consequence of their model is that variants that propagate via weighted interactor selection, that is, markers, spread differently from indicators, variants that spread through neutral interactor selection. The main difference is that, given proper tailwind, the former spread much quicker. In this model, the effect of prestige on propagation is taken for granted.

The theory of salience presented in this work is able to input a network-based model of sound change on which variants are salient, can possibly carry some form of prestige, and as a result spread differently. One prediction is that a prestige variant is adopted more readily by a larger number of speakers, spreading quickly out of the group in which it is actuated and reaching new networks. A neutral variant propagates at a slower rate, and is more likely to remain in the group where it is actuated.

That there is such a difference between salient marker and non-salient indicator behaviour in real life can be seen in the two case studies presented in this work.
8.2 Salience in the propagation of a change

Milroy (2007) proposes a distinction between two typical ways of the propagation of a sociolinguistic variant. *Off the shelf* changes are readily accessible to a large number of speakers, while *under the counter* changes spread in a socially and geographically more limited domain. Conflating her distinction with levels of indexicality through the Croft model of language evolution is a tempting one. One could argue that the propagation of markers is propelled (or hindered) by weighted interactor selection, so if it is propelled, it is more easily available to a larger number of speakers in a shorter time span. At the same time, an indicator is restricted to a smaller domain as its propagation is only propelled by frequency.

This simplifying view equates the distinct notions of indexicality (the social meanings listeners associate with a variable), selection (the readiness of listeners to adopt the variable) and salience (the extent to which listeners can be aware of the variable). The three concepts, however, are connected through salience: a salient variable is more likely both to carry social indexation and to be adopted by listeners (as long as listeners assess the attached social values favourably).

We expect that salient variation shows different patterns of use to non-salient variation. Language users attribute social meaning to it and adopt and transmit it more readily. This is the picture that emerges if we compare glottalisation in the South of England with derhoticisation in Scotland, two variation patterns discussed in detail in Chapters 5 and 7, respectively.

### 8.2.1 Glottalisation in England

Certain patterns of glottalisation, the replacement of a voiceless coronal stop with a glottal stop, show similarities with what Baxter and colleagues call propagation through weighted interactor selection. Glottalisation has been discussed in Chapter 5 in detail. To recapitulate the main findings, Glottalisation is originally a Southern English English feature, typical of the
working-class speech of London, that is rapidly spreading in use, reaching both more formal registers and other cities in England, being mostly associated with urban speech (see also Roach 1973; Wells 1982; Kortmann & Schneider 2004). (The birthplace of glottalisation in the British Isles is likely Scotland (Andrésen, 1968), but that is irrelevant here.)

It is interesting to note that the spread of glottalisation in London resembles a change from below in the sense that it originates in lower-class speech and that it is both phonetically gradient (with realisations ranging from pre-glottalised variants to glottal stop variants) and, in its more prevalent forms, ignored by the speaker community. It is only word-final pre-vocalic glottal replacement (the focus of Chapter 5) and word-medial intervocalic glottal replacement that are salient for speakers. We have good evidence that precisely these two patterns are the ones that behave differently in other UK cities – as far as glottalisation patterns go, these are the interesting ones.

Milroy et al. (1994), in a paper on glottalisation in Newcastle upon Tyne, survey the rapidity with which glottal replacement increased in in cities throughout the UK, such as Cardiff, Liverpool, or Belfast. They note that in Newcastle the glottal replacement pattern actually competes with a local pattern of glottal reinforcement, a kind of phonetic variation that is similar but not entirely identical to it. Glottal reinforcement, the co-articulation of a glottal constriction with the oral constriction in the case of the lenis stops [p],[t],[k], is a local, entrenched variety, which also shows a large extent of phonetic variation with a uni-modal distribution. It is a local prestige form used more predominantly by men. Glottal replacement, the replacement of the oral constriction with a glottal one in the case of the lenis stop [t], is the new variety, which has a bi-modal distribution: to a large extent, coronal stops are either replaced by glottal stops or not. It is more restricted than glottal reinforcement, only showing variation with [t]. It is more typically used by women.

Milroy and colleagues argue that, while the two show phonetic similarities, glottal reinforcement and glottal replacement are two distinct patterns of
variation in Newcastle and the Tyneside. The latter is a supra-local norm that was imported to this urban centre, supported by the fact the kind of gender bias shown by its use: women favour it more than men. In variationist studies of urban speech, women opt for the standard variant, which makes it surprising that they would take the lead in the use of glottalisation, an iconic English English sub-standard pattern. Milroy and colleagues reach the conclusion that this apparent contradiction is explained if we see this pattern, glottal replacement, as part of an emerging middle-class norm in the British Isles, a norm to which female speakers in Newcastle converge.

The reader might observe that the supra-local glottalisation pattern, glottal replacement, behaves in a way that Baxter et al. (2009) would expect from a marker – a variant that propagates through weighted interactor selection. It spreads very rapidly, as echoed throughout the literature from Wells (1982) to Kortmann & Schneider (2004), and speaker preferences play a role in its adoption. It is no wonder, then, that Milroy (2007) also mentions it as classic off the shelf variant, one that is easily available for a large number of speakers.

Besides being a perfect off the shelf marker, glottal replacement also meets both the perceptual and distributional criteria for sociolinguistic salience. It is noticed by listeners and its use leads to clear distributional differences. This has been discussed in depth in Chapter 5. This means that, in this case, it is the salient variant that behaves like what we call for in the case of a marker.

The argumentation above might seem obtuse. I point at the study on glottalisation in London featured in this work and link it up with another study on glottalisation in another place, Newcastle, throwing in the Croftian model of language evolution to leave the reader confused forever. What is crucial to see, however, is that these studies only provide empirical evidence for a notion that is intuitive in itself.

Glottal replacement is a (near-)categorical pattern that creates strong, perceptible segmental differences in the distributions of a type of variation already present in most English dialects (including the Newcastle city vernacular), stop pre-glottalisation and stop glottalisation. Since these segmental
differences are clear and perceptible, the variation is readily available for speakers to utilise: it is a salient, off the shelf pattern. If it is salient, we expect it – borrowing the Crofian term – to propagate through weighted interactor selection, and, accordingly, spread faster and across more speaker groups.

8.2.2 Derhoticisation in Scotland

The categorical, city-hopping, norm-shaping salient variant, the glottal stop, should be contrasted with a non-salient variant that does everything differently: it remains local, shows slow spreading, is phonetically gradient, and speakers find it difficult to notice, not to mention having explicit attitudes toward it. Beside examples like the local variety of glottal reinforcement in Newcastle (or, indeed, in London), one pattern explored in this work is derhoticisation, the loss of coda /r/ in the Southern, Urban Belt of Scotland.

Chapter 7 saw an extensive discussion of this process. What we ought to recall here is that this is more of a working-class pattern, typical of men, and it shows a large extent of phonetic variation. This is both true for the realisation of /r/, ranging from tapped variants to approximants to fricative variants, and for the extent to which coda /r/ is deleted – here, gradience in acoustic variation is closely connected to the gradience in the articulatory gestures behind /r/ realisations. Unlike other recent working-class innovations in Scottish English, such as th-fronting, derhoticisation commands less attention, with no clear patterns of style shifting and no open stereotypes. In Chapter 7 I argued that this state of affairs is due to the vast, phonetically gradient variation of /r/, which makes it more difficult for clearly distinguishable variants to emerge, and for the speakers to pick up on them.

The behaviour of coda /r/ matches that of an indicator, that is, a variant which propagates through neutral interactor selection. It is mainly local, determined by frequency of interactions in the community, and, as it is entirely ignored by the speakers, very little attention is given to it. Here, we observe the inverse of the patterning of glottal replacement. In both cases, however,
sociolinguistic salience (or lack thereof) matches a certain pattern of sound change. Coda /r/ deletion would be classified by Milroy as an under the counter change, hard to notice and local in nature.

### 8.3 Concluding remarks

This chapter talked about the link between speaker indexation, and, indirectly, sociolinguistic salience in sound change. We saw that a network-based model of change, where the sum of the interactions between the individuals plays a central role in the propagation of a variant attribute large importance to speaker/variant prestige in sound change. A model similar to that of Baxter et al. (2009) is able to give an account of sound change in a way that incorporates the findings of Milroy (1980), Eckert (2000), and Milroy (2007), where different kinds of variants behave differently in the long run. The theory of salience advocated in this work lends additional support to such a model, by giving independent tools to ascertain the salience of a variable.

The Croft/Baxter model is insensitive to salience. It only makes a crude distinction between variants which spread through neutral/weighted interactor selection by contrasting a general model of sound change only relying on neutral interactor selection with one that relies on weighted selection. Croft’s theoretical framework also remains silent about whether it is the variant that carries social indexation or the speaker who uses it. A full model of salience and sound change should probably contain both types of interactor selection (with various strengths of weights) and associate prestige with both speaker and token. This is, however, beyond our present discussion. What this chapter has shown is simply how a theory of variable salience can augment a theory of sound change based on speaker interactions.
Chapter 9

Conclusions

The previous chapters proposed an empirical approach to assessing the salience of phonological variation, offered a methodology and a number of case studies for this approach, and highlighted its relevance for a theory of language change. This final chapter sums up the main findings of this work and tries to see whether salience in phonological variation is related to some general property of sound change.

9.1 The source of salience

In this work I suggest that one source of the ability of the variant to carry indexation (and, arguably, the main linguistic source) is salience, a cognitive-perceptual property that follows from the way the variant is realised in the language structure. Assuming a phonological segmental level of linguistic structure, certain variants change the distribution of segments in such a manner that is conspicuous for the listener. This applies when the listener has a different native dialect or a projection of a different dialect. The latter is the case in the experiment of MacFarlane & Stuart-Smith (2012), discussed in Chapter 7. Their Glaswegian participants have a strong judgement on what a middle-class young university student in Glasgow sounds like, down to the sub-phonemic level, even if the participants themselves belong to a
different social group and do not speak in this particular way. If conditions for salience are met, the speech community can utilise the variable to signal social indexation, which has its specific implications to language use and language change.

9.1.1 From cognitive properties to language use

The model of variable salience discussed in this work is built on cognitive linguistic and phonetic findings on perceptual salience. There is ample evidence from studies on visual cognition that a relevant source of salience in attention deployment is contrast with the surroundings, that is, surprisal (see references in Chapter 2). This can be carried over to linguistics: Saffran et al. (1996b), amongst others, demonstrate that listeners distinguish segmental transitions with a low probability from transitions with a high probability in listening tests. This evidence strongly suggests that transitional probabilities (TP-s) in the speech signal, first pointed out by Harris (1955), are at the disposal of the naïve language user. A theory of variable salience based on distributions and surprisal implies an information-theoretic (Shannon & Weaver, 1948) framework. The segmental realisation of a variable has a surprisal value depending on its distributions in a dialect. To refer back to Chapter 4, the reason why the reduced, glottal stop variant of the definite article is salient for the listener is that it entails a different distribution of glottal stops in the dialect with definite article reduction. This means that it is not the glottal quality of the variant that is surprising to the listener, rather, the relative frequency with which this variant is distributed.

The case studies in this work demonstrate how a frequency-based theory of salience can be applied to linguistic phenomena (cf. Kerswill & Williams 2002). In these studies, social indexation (evidenced by studies on listener attitudes and dialect data) tallies with sociolinguistic salience modelled as a notable relative difference in the target dialect. Salient variables are compared to the lack of salience in cases where the relative frequency difference is absent. The studies also show that the relative frequency difference in itself does not
necessarily lead to sociolinguistic salience. For instance, if the realisation of a variable is rampant and phonetically gradient, listeners have a hard time picking up on relative differences, even if these differences are very strong. The example to support this point is the loss of coda /r/ in Glasgow (cf. Chapter 7).

If we look at the results of Timmins et al. (2004) on coda /r/ realisation in Glaswegian, we find strong differences depending on age and social class. Older middle-class speakers are way more rhotic, that is, more likely to realise coda /r/, than their young working-class counterparts. The loss of coda /r/, despite its clear social stratification, seems to fail carrying social indexation, as evidenced by the absence of large scale code switching and strong listener reactions. The reason for this is likely that coda /r/, when realised in Glasgow, has a wide range of variants, which are phonetically difficult to distinguish from each other. Not realising a coda /r/ does not starkly contrast with a realised /r/ in the coda, rather, is an endpoint of a gradient phonetic continuum. Listeners find it difficult to tackle this continuum, which explains why variation in coda /r/ realisation does not carry the same social significance as in other, similar cases.

The salience of sociolinguistic variables has consequences for language change. Since variables that carry social indexation behave differently in language change from variables that do not do so (as corroborated by a model of Baxter et al. 2009), the source of this distinction, sociolinguistic salience, along with its fundamentals, becomes material to a theory of language change.

Chapter 8 notes that this theory of salience attributes to relative frequency a strong effect on language change. An incidental and complex state of affairs – segmental distributions in a dialect – can influence the manner in which an entire language system changes.

The effects of frequency on language structure have been a recurring topic in linguistics (cf. Pierrehumbert 2003; Mendoza-Denton et al. 2003). As a matter of fact, frequency is the main focus of my research unit (DFG GRK 1624) at the University of Freiburg. This work is relevant to this line of study,
Conclusions

because, along with the other researchers in this research unit, it highlights other ways in which frequency is relevant to linguistics, that is, beyond the acknowledged effect of raw frequency on analogy and phonetic reduction phenomena (Köpcke, 1988; Plug & Ogden, 2000).

Again, in the present case, the listener is assumed to compare relative frequency differences in various structural positions in order to apprehend a variable, and this is more or less independent of raw token or type frequency. Nonetheless, it provides evidence to support the arguments other, general frequency-based studies make, such as the importance of sub-phonemic detail (Myers & Li, 2009) or the prevalence of rich memory language models in linguistic theory (Wedel, 2007).

9.1.2 Consequences for phonological modelling

The theory of salience discussed here subsets the multitude of variation with respect to which variants are at the disposal of the speech community to carry social indexation. This becomes a key aspect when approached from the perspective of phonological modelling.

Thanks to the duality of patterning, we can assume a finite set of minimal units available for manipulation above the phonological level. For example, if we look at variation in word order, we can set up a typology of possible variation patterns by manipulating the order of word-sized units or part of speech classes. As a matter of fact, even this typology will over-generalise, as not all possible word order combinations are attested in dialects of the language. If we go below and look at morphology, we can posit a typology of meaningful units and thereby establish the limits of the array of variation we are going to find.

Looking at German, the presence or absence of verb-final relative clauses after the complementiser weil is an example to word order variation. Having the participle wo stand in place of wer, was, wann in relative clauses is an instance of morphological variation. This is a simplified example, but similar cases will easily come to mind.
When we look at the phonological level, minimal units of representation are much harder to find. Phonological descriptions of a language can make do with the assumption of a minimal unit of representation, such as the phoneme (Bloomfield, 1933), the underlying contrastive segment (Chomsky & Halle, 1968), or the phonological constraint (Prince & Smolensky, 1993). It is debated whether adequate phonological descriptions can manage at all with word-level representations built up from minimal absolute units (Silverman, 2006), but there is no room to go into that debate here.

Sociolinguistic descriptions of phonological variation need to go way below phonemic detail. As Foulkes & Docherty (2006) note, it is only possible to completely cover the array of variation, both on the social side (the contexts and speaker groups associated with certain variants) and on the phonological side (the extent of detail that is sensitive to social stratification and geographical setting). This is why they advocate the use of a modelling architecture like Exemplar Theory (Nosofsky, 1988), which is, in theory, able to register phonetic detail and context without any lower boundary.

An Exemplar Theory representation is able to record all the variation of all the environments attested during field work. This main strength, however, is its main weakness as well, as such a framework will heavily over-generate. As Bermudez-Otero (2006) notes, assuming an exemplar-theoretic framework predicts language change creating a much larger array of variation than attested.

This problem is duplicated on the sociolinguistic side. If speakers are able to keep track of minute phonetic variation along with the relevant contexts, we would not expect to have the indicator/marker distinction. In theory, then, all phonological/phonetic variation should be subject to social awareness, because it is part of the speakers’ competence.

The logical answer to this is, of course, that some variation is too minute to reliably index speaker differences. The theory of salience presented in this work is an attempt at giving an explanation on how and why variation becomes highlighted for the language users – it is argued that variation
does not become salient at random, but rather through meeting a set of cognitive-perceptual properties. This is an important step, both because it gives predictions on how variation is structured and because it tells us something about the way social evaluation affects the behaviour of linguistic variants and, consequently, sound change.

If we accept that sound change is heavily affected by the social evaluation of linguistic variants – and the evidence compels us to do so – we have to also admit the relevance of frequency, and, in connection, the validity of a linguistic theory relying on sub-phonemic detail and relative frequency counts.

9.2 The predictability of salience

So far we have seen that salience, this probabilistic, low level process, is able to effect larger and more complex processes, such as the way language use changes in a society. It would be even better if we could accept that the cognitive pre-conditions of salience, transitional probability differences, do not arise in a haphazard manner, but, rather, are connected to general internal properties of language change. (Internal, that is, independent of the social factors talked about at length in Chapter 8.) If this were the case, we would not have to look at variation on a case-by-case basis to know whether it could be salient for listeners. We could say that the change we are observing is likely to be salient, because sociolinguistic salience is concomitant with this particular kind of change.

This, of course, only holds when variation follows from an ongoing change. Salient variation can be stable, with no change in progress, as in the case of the realisation of the suffix -ing in American English (Labov et al., 2006). Salient variation can also result from dialect contact or a general koineization process (Trudgill, 1986). Having said that, since we talked about salience and language change extensively already, the topic certainly merits some more attention. If one wants to find the property or properties that correlate with salience in phonological change, it is useful to look at some properties people
9.2.1 Types of phonological change

One classic distinction, touched upon in Chapter 8, is made by the Neogrammarians, between analogical and regular, phonetic change. The latter, regarded as ‘genuine’ language change, would subsume almost all processes discussed in this work: reduction phenomena, assimilations, vowel mergers and shifts, and so on. This summary does not aim at a thorough critique of the Neogrammarians method. It is important to mention again, however, that the basic tenet that phonetic change is not directly observable, and is, at the very least, completely beyond the awareness of the language users, has been challenged by variationist sociolinguistics.

The distinction between a type of sound change that is gradient, deterministic, and independent of the word level (affecting all word forms matching its structural description at once) and another type that is categorical and proceeds by lexical diffusion (affecting word forms one at a time, basically) survives into the Generative view of diachronic variation (Bermudez-Otero, 2006). The recognised strict difference between phonetic change and lexical diffusion is partly due to the examined evidence, and partly due to the usual model architecture, which only allows either a gradient change affecting all respective environments or a categorical change affecting individual lexical representations.

This view is consonant with Labov’s take on vowel change. Labov (1994) gives two examples of regular and diffusing change in Philadelphia. The raising of the variable (aeh) entails the raising of /æ/ in all environments, without respect to word-level differences. The tensing of the variable (ae) entails the tensing of /æ/ in a given lexical set. It involves a higher order categorical shift (from lax to tense) and shows lexical diffusion. Labov notes that the regular low level phonetic change is subject to social awareness and is easily learnable, whereas the high level case of lexical diffusion escapes...
Conclusions

listener attention as is hardly ever learned by people moving into Philadelphia
and taking up a local accent.

It would then seem that, at least from the point of view of salience and
social indexation, Labov manages to turn the Neogrammarian view upside-
down. Regular phonetic change is subject to speaker awareness, whereas
lexically diffusing change is not. This tallies with his observation (Labov,
2001) that vowel shifts are salient for the listeners, whereas splits and mergers
are not.

He argues that the main difference is between the level of the change.
The low level change, affecting vowel height, is salient for the speakers of the
Philadelphia dialect, whereas the high level change, affecting the tense-lax
distinction, and lexical subset, is not.

There are two problems with interpreting this distinction on a general
level. First, it is orthogonal to social factors in sound change. A change from
below would involve a gradient shift and remain below speaker awareness,
whereas a change from above would involve the adoption of a categorical shift,
with social significance.

9.2.2 Consonants and vowels

The state of affairs is even worse if we have a look at consonantal changes,
because, apparently, consonants work the other way around. A gradient
consonantal change, such as the pre-glottalisation of stops in Newcastle,
would escape the attention of the speakers, while a categorical one, such as
glottalisation also in Newcastle, would carry social indexation. This example
has been discussed in Chapter 8. Other examples can be found. The fronting
of the dental fricative in English dialects, a categorical shift from [θ] to [f], is
salient, whereas the loss of coda /r/, a gradient change, is not.

The apparent contradiction in the social evaluation of gradient versus
categorical changes in the case of consonants and vowels (again, ignoring
the social dimension) has a lot to do with the role of consonants and vowels
in transmitting information (Fant, 1970). Vowels involve subtle changes in
The predictability of salience

The predictability of salience

the carrier signal, and inherently communicate most of the extra-linguistic cues. Pitch and intonation, the main cues of speaker age and gender, are all vocalic properties or properties that come across through vowel articulation. Consonants are sharper modulations of the carrier signal, and cue most of the information content. This is apparent when we consider that (i) a written text remains intelligible with the omission of vowels and that (ii) most sociolinguistic variables involve non-contrastive differences in vowel articulation – no wonder vowel differences have always been the main scope of variationist sociolinguistic studies.

At the same time, it has to be borne in mind that most of these concepts are not inherently categorical either. The existence of near-mergers blurs the boundary between gradient/categorical or non-neutralising and neutralising variation, while liquids and glides also fall in the middle of the consonant-vowel continuum.

9.2.3 Overview

In an ideal world, we would only have to tackle phonologically gradient, low level, non-neutralising changes versus categorical, high level, neutralising ones. We could then hopefully find out that either of these is typically salient for the listeners (again, the social aspect notwithstanding), while the other is not. This is, however, not the case.

The brief and non-exhaustive overview of these concepts should only tell us that there is no easy correlate of sociolinguistic salience in language change. Mainly, it has two main affecting factors, internal (linguistic) and external (social) ones. If we put the social aspects aside, and concentrate on the phonological/perceptual properties of the variation under scrutiny, we can make predictions on whether this variation is perceived as salient by the listeners or not. Indeed, this is what this book is about.

Finding one unmistakeable phonological property of change that renders the ensuing variation as perceptually salient is a whole other matter. At this point, it seems that none of the properties used to categorise sound change
Conclusions

in one way or another match up with salience in a trivial manner. Rather, the sociolinguistic salience of variation is always the result of a complex interaction of the type of change, the phonological properties of the units involved, and the social context. And all this, of course, only in sound change, the main subject of sociolinguists both in the study of salience and elsewhere. What follows is that whenever the salience of a variable becomes an issue, one should refrain from sweeping generalisations and proceed with the attention and caution the subject deserves.

9.3 Concluding remarks

This chapter provided a walkthrough of the development of the theory of salience in this work. This theory is, to a large extent, only applicable to consonantal variables. As much as the lack of a grand unified theory of salience is regrettable, this is partly understandable as it follows from the principal distinction between consonants and vowels in the transmission of linguistic information. Beside highlighting and emphasising this difference, this work also gave due consideration to a number of consonantal variables and the way consonantal variables generally behave during sound change.

Apart from its focus on the peculiarities on consonants, this work also stressed the importance of disjoining the concept of salience from any other general phonological or social property, by way of illustration of how it can be independent of and interact with many of these. In general, this work, while not aspiring for completeness, is an attempt at laying the ground rules of studying sociolinguistic salience, and hopes to be a useful heuristic for further hypotheses to be tested and further models to be run.
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Lo-Co-Ca: Glottal replacement of t in London (interview style) – from Altendorf (2003)</td>
</tr>
<tr>
<td>2.1</td>
<td>Difference of a single feature</td>
</tr>
<tr>
<td>2.2</td>
<td>Difference of a single feature</td>
</tr>
<tr>
<td>2.3</td>
<td>Difference in two features</td>
</tr>
<tr>
<td>2.4</td>
<td>Probability of Y following X</td>
</tr>
<tr>
<td>2.5</td>
<td>Possible following segment for a voiceless coronal stop in English</td>
</tr>
<tr>
<td>3.1</td>
<td>Probability of Y following X</td>
</tr>
<tr>
<td>5.3</td>
<td>Glottal replacement of /t/ in interview style – from Fabricius (2000)</td>
</tr>
<tr>
<td>5.4</td>
<td>Glottal replacement of /t/ in reading style – from Fabricius (2000)</td>
</tr>
<tr>
<td>5.5</td>
<td>Glottal replacement according to regions – from Fabricius (2000)</td>
</tr>
<tr>
<td>6.1</td>
<td>Weighted scores for condition (i)</td>
</tr>
<tr>
<td>6.2</td>
<td>Weighted scores for condition (ii)</td>
</tr>
<tr>
<td>6.3</td>
<td>Transitional Probabilities for the different patterns involving [j]</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

7.1 Derhoticisation versus social class and environment: reading (n=1103) ........................................ 144
7.2 Derhoticisation versus social class and environment: spontaneous speech (n=8170) .................. 146
7.3 (rV) versus social class: reading and spontaneous speech ... 151
7.4 Rhoticity in FRED ELN: Tokens ......................... 165
# List of Tables

1.1 Reduced definite articles (examples from Fred) ............... 21  
1.2 (r) versus social class: spontaneous speech (n=1474) .......... 28  
3.1 The IPA and the Disc phonetic alphabet .................... 70  
4.1 Reduced definite articles (examples from the Fred corpus) .... 77  
4.2 Grassington data on DAR use .................................. 78  
4.3 Conditions in the Jones 2007 study ........................... 79  
4.4 Glottal stops word-initially ................................. 86  
4.5 Word boundary probabilities ................................. 88  
4.6 DAR and phonetic environment ............................... 89  
4.7 DAR & connected speech sequences 1 ....................... 90  
4.8 DAR & connected speech sequences 2 ....................... 90  
4.9 stop clusters ............................................. 92  
5.1 Transcription t-counts in the LLC .......................... 108  
5.2 Coronal and glottal stop counts based on the LLC ........... 109  
5.3 Transcription t-counts in the SCALL ....................... 111  
5.4 Coronal and glottal stop counts based on the SCALL ........ 111  
5.5 Word-final [ʔ] TP-s (formal) ................................. 112  
5.6 Word-final [ʔ] TP-s (formal & informal) ..................... 115  
6.1 Lexical hiatus resolution in Hungarian ..................... 120  
6.2 Post-lexical hiatus resolution in Hungarian .................. 121
LIST OF TABLES

6.3 Definite article allomorphy ........................................ 122
6.4 [j] in Hungarian ......................................................... 130
6.5 Conservative ECH ....................................................... 131
6.6 String frequencies in the corpus sample ......................... 131
6.7 TP-s in the corpus sample ........................................... 132

7.1 (rV) versus social class: reading (n=1103) ...................... 142
7.2 (rV) versus environment: reading (n=1103) .................... 143
7.3 (rV) versus social class: spontaneous speech (n=8170) ..... 145
7.4 (rV) versus environment: spontaneous speech (n=8170) ... 145
7.5 Style shifting examples in the 1997 corpus ((rV) RS n=1103, SS n=8170; (l) RS n=656, SS n=3420; (t) RS n=1212, SS n=2385) ............................................. 147
7.6 (r) versus social class: reading (n=1834) ....................... 150
7.7 (r) versus social class: spontaneous speech (n=1474) ...... 150
7.8 Style shifting examples in the 1997 corpus: (r) (RS n=1834, SS n=1474) .................................................. 152
7.9 Mean auditory rhoticity scores in Lawson et al. (2011) ... 156
7.10 Rhoticity in FRED ELN: Tokens ................................. 164
7.11 Rhoticity in FRED ELN: Ratios ................................ 164
7.12 Coda [r] transitional probabilities in FRED East Lothian . 167
7.13 Coda [r] transitional probabilities in FRED East Lothian and Glasgow working-class youth conversations data .... 167
7.14 Variation in coda /r/ vocalisation ................................ 169
Bibliography


BIBLIOGRAPHY


Blevins, J. (2004). Evolutionary Phonology: The Emergence of Sound Patterns. CUP.


BIBLIOGRAPHY


Jones, D. & M. V. Trofimov (1923). *The Pronunciation of Russian*. CUP.


BIBLIOGRAPHY


BIBLIOGRAPHY


Wright, J. (1905). The English Dialect Grammar. OUP.
Die vorliegende Dissertation untersucht eine quantitative Hypothese zum soziolinguistischen Salienzbegriff. Sie stellt die Hypothese auf, dass der Unterschied zwischen nicht-markierten und markierten (i.e. salient) dialektalen Merkmalen in ihrer negativen Frequenz liegt. Um es einfacher auszudrücken, je unwahrscheinlicher es ist, dass die Realisation einer Dialektvariante im Sprachsignal vorkommt, desto auffälliger wird die Variante an sich. Die Hypothese wird an den Beispielen der Reduktion des bestimmten Artikels in Nordengland, der Glottalisierung des Phonems /t/ im Süden Englands, der Auflösung des ungarischen Hiatus, und der Rhotizität im schottischen Englisch überprüft.