VARIATION AND CHANGE IN NORTHERN ENGLISH VELAR NASALS: PRODUCTION AND PERCEPTION

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Humanities

2018

George Bailey

School of Arts, Languages and Cultures
# Contents

List of publications ............................... 4
List of figures ........................................ 5
List of tables .......................................... 8
Abstract .................................................. 10
Declaration .............................................. 11
Copyright statement ..................................... 12
Acknowledgements ....................................... 13

## Introduction ........................................ 16

1 Introduction .......................................... 17
   1.1 Questions that will be addressed .................. 18
   1.2 Organisation of this thesis ......................... 20
   1.3 Research outcomes ................................... 22

2 Velar nasals .......................................... 24
   2.1 Variation in (ng) .................................... 25
   2.2 Variation in (ing) .................................... 28
   2.3 Phonemic status ...................................... 30
   2.4 Typological status of the velar nasal .............. 33
      2.4.1 The German velar nasal ......................... 34
      2.4.2 Velar nasals in other Germanic languages ....... 37

3 The North of England ............................... 40
   3.1 The geographic north .............................. 40
   3.2 The linguistic north .............................. 42
      3.2.1 Dialectology in the North of England .......... 43
      3.2.2 The North West as a ‘speech community’ ....... 46
      3.2.3 Northern conservatism .......................... 47
   3.3 Perceptions of the North ........................... 49
List of publications

This is a journal format thesis, consisting of the following four original manuscripts, in order of their appearance in this thesis:


These will be referred to in the body of the thesis by their Roman numerals (e.g. Paper I, Paper II etc.). The pagination flows throughout the entire thesis, but Section, Figure, Table, and Footnote numbering resets at the start of each paper.
List of figures

Introduction

2.1 The regional distribution of [ŋɡ], based on (a) the 1950s Survey of English Dialects and (b) the 2010s UK Dialect Maps projects. 26
2.2 Velar nasal presence across a sample of the world’s languages. 34
2.3 Derivation of Dinge and Tango in German, illustrating the initial non-syllabification of schwa. 36

3.1 Map of the United Kingdom highlighting the three northern regions of England. 41
3.2 Map of the United Kingdom, indicating the different ways of delimiting the southern border of the North of England. 44
3.3 Average prestige ratings of 33 varieties of English on a 7-point Likert scale. 49

4.1 Modular feedforward architecture of grammar. 53
4.2 Stratal-cyclic version of the modular feedforward architecture of grammar. 54

5.1 Illustration of the forced alignment workflow. 66
5.2 Comparison between manual and automated identification of the dependent variable in (ing) and (ng). 68
5.3 Map of postcode regions of the North West, highlighting participant locations. 72
5.4 Distribution of participants by sex and date of birth. 73

Paper I

1 The life cycle of phonological processes in a modular architecture. 82
2 Example spectrograms showing [g]-presence and [g]-absence. 88
3 The rate of [g]-presence by morphophonological environment. 90
4 The rate of [g]-presence by morphophonological environment, for individual speakers. 91
Simulated grammars (1,000 iterations) with variable domain-specific rates of application of /ɡ/-deletion. .......................... 92
Apparent time change in the rate of pre-pausal [ɡ]-presence. ............... 93
Results of hierarchical cluster analysis. .................................... 95
Average rates of [ɡ]-presence by morphophonological environment (cluster #1). .................................................. 96
The phonological derivation of 1,000,000 (ng) tokens based on cycle-specific deletion rates. .......................... 99

Paper II

The impact of boundary strength on the sonorant (vowel+nasal) duration of (ng) tokens. .................................................. 126
Example spectrograms and waveforms of song with [ɡ]-absence (left) and thing with [ɡ]-presence (right). .............................. 128
Change in apparent time of pre-pausal [ɡ]-presence, based on a sample of 32 speakers. .................................................. 129
Rate of post-nasal [ɡ]-presence by boundary strength and the type of following segment. .................................................. 131
Relationship between sonorant duration, following pause duration, and [ɡ]-presence for pre-consonantal (ng) tokens. .............. 132
Pitch contours and ToBI annotation for two pre-pausal tokens in the IP-medial context: one genuinely IP-medial in (a) and one produced with phrase-final intonation in (b). .............................. 134
Visualisation of the distribution of pauses and IP boundaries in this study’s dataset, and the effect they have upon [ɡ]-presence. ............ 137

Paper III

Results from the norm-identification and self-report tasks in West Wirral, Merseyside (Newbrook 1999: 100–102). .............................. 155
Distribution of ratings by guise and environment for all four descriptive scales. .................................................. 161
Difference scores by environment, scale, and age group. .............. 163
Distribution of difference scores by scale and age group. .............. 163
A typology of velar nasal plus granularity, moving from more coarse-grained (top) to more fine-grained (bottom). .................. 172
<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proportional use of all three (ing) variants in conversation, on a speaker-by-speaker basis.</td>
<td>198</td>
</tr>
<tr>
<td>2</td>
<td>Rate of -in by speaker age and gender.</td>
<td>200</td>
</tr>
<tr>
<td>3</td>
<td>Rate of -in by phonological environment.</td>
<td>201</td>
</tr>
<tr>
<td>4</td>
<td>Rate of -in by speech rate.</td>
<td>202</td>
</tr>
<tr>
<td>5</td>
<td>Rate of [ɡ]-presence by phonological environment.</td>
<td>204</td>
</tr>
<tr>
<td>6</td>
<td>Rate of [ɡ]-presence by speech rate.</td>
<td>205</td>
</tr>
<tr>
<td>7</td>
<td>Rate of -in, -ing and -ingg by discourse style.</td>
<td>205</td>
</tr>
<tr>
<td>8</td>
<td>Distribution of nasal duration by (ing) variant.</td>
<td>206</td>
</tr>
<tr>
<td>9</td>
<td>Nasal duration by speech rate and (ing) variant.</td>
<td>207</td>
</tr>
<tr>
<td>10</td>
<td>Vowel plot of all conversational (ing) tokens.</td>
<td>209</td>
</tr>
<tr>
<td>11</td>
<td>F2 trajectories by vowel, following segment, and (ing) variant.</td>
<td>211</td>
</tr>
</tbody>
</table>
# List of tables

## Introduction

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Overview of the (ing) and (ng) variants.</td>
<td>25</td>
</tr>
<tr>
<td>2.2</td>
<td>The six possible variants for a word subject to both (ing) and (ng) variation, such as <em>ringing</em>.</td>
<td>25</td>
</tr>
<tr>
<td>5.1</td>
<td>Example of dictionary expansion, where (ing) words are represented by three distinct phonemic transcriptions and (ng) words by two.</td>
<td>68</td>
</tr>
<tr>
<td>5.2</td>
<td>Sociodemographic information for the 32 participants of the sociolinguistic interviews.</td>
<td>75</td>
</tr>
</tbody>
</table>

## Paper I

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The life cycle of post-nasal /ɡ/-deletion, highlighting the effects of domain narrowing on its application in different morphophonological environments.</td>
<td>83</td>
</tr>
<tr>
<td>2</td>
<td>Eligibility for deletion by morphophonological environment. Representations with onset [ɡ], where deletion cannot apply, are in grey.</td>
<td>85</td>
</tr>
<tr>
<td>3</td>
<td>Mixed-effects logistic regression model; [ɡ]-presence as the application value.</td>
<td>94</td>
</tr>
</tbody>
</table>

## Paper II

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The age and sex distribution of informants for the elicitation task (coded in a binary manner as <em>old/young</em> and <em>male/female</em>).</td>
<td>124</td>
</tr>
<tr>
<td>2</td>
<td>Mixed-effects logistic regression model for the interaction between following segment and date of birth.</td>
<td>130</td>
</tr>
<tr>
<td>3</td>
<td>AIC comparison between models, all of which include random intercepts of speaker and word.</td>
<td>135</td>
</tr>
<tr>
<td>4</td>
<td>Best-fitting logistic regression model; [ɡ]-presence as application value, with random intercepts of speaker and word.</td>
<td>135</td>
</tr>
<tr>
<td>5</td>
<td>List of elicited sentences by boundary category, vowel height, and following segment.</td>
<td>141</td>
</tr>
</tbody>
</table>
Paper III

1 Mixed-effects linear regression coefficients for (a) the northern scale, and (b) the professional scale. .......................... 165

Paper IV

1 Comparison between Zipf score and frequency per million words. . . . 197
2 Mixed-effects logistic regression model for (ing); -in as the application value. .................................................................. 199
3 Mixed-effects logistic regression model for velar (ing) tokens; [g]-presence as the application value. .......................... 203
4 Mixed-effects linear regression model for nasal duration. .............. 208
5 Mixed-effects linear regression model for front diagonal of /ɪ/ in (ing). 210

Conclusion

6.1 Example derivations illustrating how different rule ordering can lead to word-final [ŋ] in Standard German and [ŋk] in northern German varieties. ................................................................. 240


**Abstract**

This thesis is a sociophonetic and phonological study of velar nasals in the North West of England, concerned specifically with the variable presence of \[ŋɡ\] clusters in words such as *wrong* and *singer*, which have undergone obligatory coalescence to \([ŋ]\) in almost every other region of the English-speaking world.

The pathway of change followed by post-nasal /ɡ/-deletion in the Late Modern English period has been employed as evidence for the life cycle of phonological processes. However, this theory of sound change and the architecture of grammar that underpins it both make a number of predictions about how this process should behave synchronically that are as yet untested. In this thesis I draw upon complementary sources of data from informal sociolinguistic interviews, elicitation tasks, and a matched-guise perception survey to provide a unified account of northern English (ng) with respect to its observed pathway of change, its variation in contemporary dialects, and the wider indexicality and social meaning of its locally-restricted \[ŋɡ\] variant.

The results lend support to the life cycle in three ways. Synchronic variation in northern English (ng) is conditioned almost entirely by morphophonological factors in ways predicted by the life cycle, providing support for a cyclic model of phonology stratified into stem-, word-, and phrase-level strata. Differences between these domains with respect to the rate at which deletion applies also reflect the age of each avatar of this process, providing further empirical insight into the history of this change. Finally, a cross-linguistic comparison of the distribution of \([ŋ]\)~\[ŋɡ\] suggests that rule generalisation has given rise to dialectal microtypologies that represent different ordered stages of the same diachronic pathway of deletion.

This thesis also reveals that \([ɡ]\)-presence is increasing over time in pre-pausal position. Crucially, results of the perception task suggest that this is not a case of evaluation-driven change. Contrary to popular conceptualisations of the speech community, there is no evidence of a shared evaluative norm in the case of (ng). This is a case of incipient social meaning: the variable is increasing in salience as a dialectal feature, but in the early stages of this change there is not yet a shared consensus on the content of its evaluation. The results also highlight theoretical issues – specifically the granularity of social meaning – surrounding the incorporation of sociolinguistic evaluation into explanatory models of sound change. Instead, I propose an alternative account of this innovation, namely that it reflects a more general motivation for pre-boundary strengthening: alongside other cues to phrasal boundaries, \[ŋɡ]\ is used as a clear speech variant in pre-pausal position as a device to signal turn-taking and to mark a juncture in speech.

These results highlight the importance of combining synchronic and diachronic accounts in explanations of language variation and change: in the case of Northern English (ng), historical evidence can be used to inform and explain its pattern of synchronic variation, and conversely its synchronic status can lend insight into its historical trajectory of change.
DECLARATION

I declare that no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
COPYRIGHT STATEMENT

The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
During the second year of my doctoral studies I came across an amusing comparison between the PhD experience and Frodo’s journey in *The Lord of the Rings*. Needless to say, as a lifelong Tolkien fan, this really struck a chord with me. While the fate of the world was never really riding on the successful completion of this PhD, it sometimes felt like it. Much like the One Ring, at times I felt a curious attraction to this piece of work, and at other times I wanted to cast it into the depths of a fiery volcano. I only hope that readers will agree more with the former than the latter.

Of course, the heart of this analogy lies in the importance of fellowship. It’s often said that the PhD can be a lonely and isolating experience, but this was not the case for me: I was blessed to have had so many wonderful people accompany me on this adventure. First and foremost, I was lucky enough to enjoy doctoral life under three supervisors, to whom I will be eternally grateful: Maciej Baranowski, Ricardo Bermúdez-Otero, and Laurel MacKenzie.

Maciej has played a huge role in cultivating my curiosity in Northern English dialects. When enrolling on his third-year ‘Manchester speech community’ course, I didn’t realise at the time that I would discover a love for sociophonetic research that would define my postgraduate research career. In the following years Maciej would be ever-present, providing a wealth of advice and accompanying me to conferences all over the globe, which was a big help early on as a young and nervous first-year navigating my way through a scary new academic world.

Ricardo often went above and beyond the call of duty as a supervisor, providing incisive commentary on all aspects of my research, whether through 3am emails or meetings that regularly ran into the night. Ricardo is an endless source of knowledge and enthusiasm, and while the former proved to be an immensely valuable asset during this PhD, it is the latter that has had a real and lasting impact. On some level I think I always had a dormant interest in the scientific endeavour, but in certain cases you need somebody to provide the spark to ignite it. I hope I can offer the same when I myself move from supervisee to supervisor.

Laurel was instrumental in my development as an independent researcher, and quickly became the kind of mentor that everybody needs in the early stages of their academic career. Working with Laurel also gave me my first taste of the conference experience when we presented at *Methods in Dialectology XV* in Groningen. I was lucky to also share the floor with Danielle Turton, who deserves mention as a brilliant and friendly source of advice and guidance. During these conferences I’ve had the pleasure of meeting so many amazing people from every corner of the earth, who are too numerous to name so I’ll reserve my thanks for the next time we cross paths at a wine reception.

The department at Manchester is a truly special place to work and learn. I should know, having been there for seven years now. During this time I’ve had the pleasure of working with so many great scholars who (more importantly) are also great people. Thanks in particular go to Patrycja Strycharczuk for every pearl of wisdom, for overseeing the development of a phonetics lab that I will dearly miss, and for being...
more Manc than I could ever dream to be. Also to Wendell Kimper, who stepped in as my independent reviewer when Laurel’s move across the Atlantic prompted a reshuffle of my supervisory team. It’s a testament to everyone that the ship remained steady during this time. Special mention also goes to Andrea Nini, for reigniting my interest in using Twitter data for linguistic research, and for bringing the excitement of murder mystery to the department.

My day-to-day experience during the PhD was enriched by the presence of so many fellow students who are brilliant, clever, funny, and kind in equal measure. Thanks in particular go to Stephen Nichols, who provided so much help with various aspects of this write-up (not just linguistically but also with \LaTeX), and to Stefano Coretta for his stats wizardry. To them, and to the rest of my postgraduate family – Fernanda Barrientos, Massimiliano Canzi, Simone De Cia, Deepthi Gopal, Henri Kauhanen, Sarah Mahmood, Lorenzo Moretti, Donald Alasdair Morrison, Nicole Rajan-Brown, Jane Scanlon, and Kaiyue Xing – I say thank you: for the constant laughter, the custom Slack emojis, the yearly mPILs, the daily discussions about typography, and most importantly for the table tennis tournaments. Finally, special thanks also go to Hannah Booth and Chris Hicks, who both started this journey alongside me in the September of 2015 and who have become good friends. Thanks for the weekly lunches/dinners (delete as appropriate), and for making sure that I socialised outside of the phon/phon bubble.

At this point I’d also like to express my gratitude to the ESRC, who have funded me throughout four years of postgraduate research and allowed me to pursue a Masters and PhD in the first place. None of this would have been possible without their generous financial support. I’m also immensely happy that at the end of this journey I get to start yet another almost immediately. The road goes ever on and on, in this case leading me across the Pennines to the University of York. Thanks for offering me a job while a good chunk of this thesis remained unwritten, and for ensuring that I remain on the right side of the foot–strut isogloss for a while longer.

I’d also like to thank the unsung heroes of this research: the informants. Thanks for your time, for lending your voices to science, and for not being too mad when I couldn’t tell you what exactly I was listening out for. At least now you can read the subsequent seventy thousand words if you want to know that badly.

There are of course countless people outside of the academic world who also deserve recognition. To James and Carlos, for reminding me that there is a life outside of this thesis, even if it does largely revolve around pubs and football. To my family, for being a constant source of love and support, and my parents in particular for instilling in me a passion for learning from a young age. To my grandparents, who – through the powers of the sociolinguistic interview – granted me valuable insight into the local dialect of the 1930s, but more importantly gave me the opportunity to capture the countless tales and anecdotes that have given us all so much joy and laughter over the years.

And of course, to Jenny. I have so much to thank you for, not only your unwavering support but your encouragement to change degrees to linguistics in the first place! Neither of us thought I would still be studying it seven years later, but as you always remind me: everything happens for a reason. Writing this thesis pales in comparison to what I achieved in first year when I somehow convinced you that I was worthy of your time, and since then you’ve accompanied me every step of the way. After the best part of a decade I leave university with a handful of degrees, but most importantly I leave with you by my side.
No language is justly studied merely as an aid to other purposes. It will in fact better serve other purposes, philological or historical, when it is studied for love, for itself.

—J. R. R. Tolkien, 1955

Dark and true and tender is the North

—Alfred, Lord Tennyson, 1847
Introduction
Dialects of the north of England are renowned for their conservative status with respect to language change, with a number of the more rural varieties spoken in this region characterised by the use of traditional forms that have since been lost elsewhere in the country (Wells 1997). One such example of this resistance to linguistic change is the variable presence of post-nasal [ɡ] in words such as wrong and singer, which have undergone ng-coalescence in almost all of the world’s varieties of English but can still be realised as [ɹɒŋɡ] and [sɪŋɡə] in some areas of the North (Wells 1982b; Wakelin 1984; Hughes et al. 2012). This retention of [ɡ] also interacts with existing sociophonetic alternations: post-nasal [ɡ] can occur word-finally in unstressed (ing) clusters – e.g. waiting [weɪtɪŋɡ] – leading to a situation in which the widely-studied variation between [ɪn] and [ɪŋ] takes on a rather unique status in these regions as a case of ternary, rather than binary, variation. As will be shown in this dissertation, the presence of these [ŋɡ] clusters is not simply a relic of an earlier stage in the history of English, but rather a fruitful area of sociophonetic research that lends insight into a number of issues in phonological theory.

While we know a great deal about the historical aspect of post-nasal /ɡ/-deletion, particularly its diachronic pathway of change (see e.g. Bermúdez-Otero & Trousdale 2012), its variation in contemporary northern dialects is severely understudied. This dissertation seeks to address this gap, and does so by situating the synchronic status of (ng) variation in the context of its diachronic change throughout the Modern English period. This interplay between synchrony and diachrony forms a key element of the life cycle of phonological processes (Bermúdez-Otero 2013), but crucially the life cycle makes a number of predictions regarding how (ng) should pattern synchronically that are as yet untested. Exploring these predictions presents an opportunity to not only provide an explanatory account of (ng) variation in Present Day English, but to also gain insight into theories of language change and of the architecture of grammar more generally.

As a further consequence of the lack of research into (ng), it remains unclear
exactly how the local [ŋɡ] variant is perceived by listeners. It has been claimed that the non-coalesced form has local prestige and is seen as ‘posh’ or correct, possibly under the influence of the orthographic representation of these clusters (Foulkes & Docherty 2007: 64; Beal 2004: 127), but this has not been substantiated empirically by actual perception studies. This dissertation provides the first experimental evidence on how [ŋɡ] is evaluated and the indexical meanings with which it is associated; in doing so, this research further addresses the extent to which evaluation of [ŋɡ] can be used as an explanatory factor for ongoing changes in its production, an issue that also speaks to wider questions about the mechanisms that drive the incrementation of sound change.

In order to pursue these various strands of research, this dissertation draws upon three complementary sources of data: (i) a collection of sociolinguistic interviews conducted with a socially-stratified population sample from the North West of England, (ii) an elicitation task to investigate the realisation of (ng) in a more controlled set of phrasal environments, and (iii) a matched-guise perception survey to uncover listener attitudes towards [ŋɡ].

The remainder of this introduction is as follows: in Section 1.1 I make explicit the exact research questions that will be explored in this thesis, in Section 1.2 I outline how the remainder of this work is structured, and in Section 1.3 I summarise the outcome of these investigations.

1.1 Questions that will be addressed

There are two over-arching aims of this dissertation: to investigate synchronic patterns of (ng) variation in contemporary dialects of the North West of England, and to establish how the local [ŋɡ] variant is socially evaluated by members of these speech communities. As reflected by the title, there are in fact four aspects of these velar nasal clusters that will be covered: variation in production, variation in perception, change in production, and change in perception. The results of this study present evidence of all four.

However, while the general aim is to provide a thorough and empirically-grounded description of variation and change in Northern English [ŋɡ], and of its sociosymbolic meaning as a feature of local dialects, this body of research began specifically as an exploration of its morphosyntactic conditioning. This is formulated in the following research question (RQ):

RQ1: To what extent does synchronous variation in Northern English (ng) exhibit sensitivity to morphosyntactic structure, and does this sensitiv-
ity result in patterns of variation that are predicted by the life cycle of phonological processes and by an architecture of grammar with cyclic phonological computation?

This question, which is addressed in Paper I, provided the point of departure for the wider body of research. The study uncovered a hitherto-unreported change in progress, restricted to phrase-final contexts, from which a number of follow-up questions emerged. These, listed below, are addressed in the follow-up studies presented in Papers II, III and IV.

RQ2: What are the exact mechanisms that condition variable [ɡ]-presence in phrase-final environments? Specifically, is this variation directly influenced by prosodic boundary strength, or does it reflect sensitivity to some of the collinear phonetic properties that mark the right-edge of prosodic constituents, such as segmental duration or length of pause?

RQ3: What is the relative salience of post-nasal [ɡ] as a characterising feature of northern dialects, and how does this modulate the degree to which it is subject to overt sociolinguistic evaluation? Furthermore, how does its association with local dialects correlate with its perceived level of overt prestige, and to what extent can its sociolinguistic evaluation act as a driving force in this phrase-final change?

RQ4: What role does [ɪŋɡ] play in the three-way variation in Northern English (ing)? Is its presence conditioned by the same factors as in the binary (ng) variable? Additionally, do the more geographically-widespread [ɪn] and [ɪŋ] variants behave in the same way here as in other varieties of English?

While the sociophonetic variables (ng) is clearly the direct focus of this research, the results have much wider implications for a number of other issues in both sociolinguistic and phonological theory. That is, (ng) is used as a testing site for a number of more general theoretical issues, and the results speak to a wider range of questions such as:

- How does the morphosyntactic conditioning of phonological processes vary and change over time, and what does this tell us about the nature of phonological representations and the architecture of grammar?
• How are processes of external sandhi influenced and modulated by phonological factors such as prosodic constituency, phonetic factors such as segmental lengthening, and psycholinguistic mechanisms of production planning?

• What is the structure of community-wide evaluation in the early stages of the development of social meaning?

• How should a speech community be defined if not in terms of uniform evaluation and adherence to a set of shared sociolinguistic norms?

• What is the granularity of social meaning, and to what extent can it be used as an explanatory force in language change?

• What happens in cases of ternary variation for variables that are binary in all other speech communities? Does the locally-restricted variant take on the sociosymbolic meaning of one of the existing forms or does it occupy its own social and linguistic role in the variation?

1.2 Organisation of this thesis

A major portion of this ‘journal format’ thesis is made up of four independent manuscripts, which are all united in studying the same sociophonetic phenomenon: variation in the realisation of /ŋɡ/ clusters in Northern Englishes. Each paper pursues a different set of research questions and addresses a single aspect of this variation, and as such they can be read independently of one another and in a different order from how they are presented. However, despite the self-standing nature of these manuscripts, there is a logical order in their order of appearance: in some cases one paper might lay the foundation for another, or might speak to questions raised in another, and to a large extent their presentation in this thesis reflects the order in which the research was conducted.

These papers are preceded by a number of introductory chapters, addressing a range of topics in order to contextualise the research and provide a detailed background to the variable in question. In some cases these chapters provide more detail than was possible to include in the papers themselves, and in some cases they also address wider topics that are not touched upon at all in the original publications.

In Chapter 2 I provide a background on the status of velar nasals in English and in particular their behaviour in varieties of Northern English. This chapter introduces the variation present in unstressed (ing) and stressed (ng) clusters, including how and when this variation originated. It also provides a discussion of the underlying representation of these clusters, specifically the manner in which velar nasals are derived
through nasal place assimilation and the implications that variable ng-coalescence has for the phonemic status of the velar nasal in Northern English. Finally, Chapter 2 provides a brief overview of the typological status of [ŋɡ] clusters, comparing their status in varieties of British English with their wider cross-linguistic behaviour.

The discussion in Chapter 3 is primarily concerned with the geographic and linguistic status of where exactly this variation takes place: the North West of England. It provides an overview of this region and of how England’s North-South divide has been defined in geographic, political, and linguistic terms. There is also a brief discussion of the status of (ng) with respect to the North’s tendency for linguistic conservatism, and of how the North of England is perceived with respect to both attitudes and the exact location of its southern border.

Chapter 4 outlines the theoretical issues that inform many strands of this research, most notably theories of the architecture of grammar, of language change, and of how morphosyntactic factors come to play a role in the course of sound change. It also presents a brief discussion of a crucial debate in sociolinguistic theory regarding the social meaning of variation and its role in the incrementation of change.

Although each paper includes its own methodology section, this is supplemented by the content of Chapter 5, which provides an overview of the sources of data drawn upon in each aspect of this research, including the manner of data collection, their complementary strengths and weaknesses, and the methods of analysis employed in each investigation.

After this background, the papers themselves are presented. Paper I establishes the morphosyntactic sensitivity of synchronic (ng) variation and how this reflects the diachronic pathway of change predicted by the life cycle of phonological processes; in doing so it also finds evidence of ongoing change within the community, with an increase in [ɡ]-presence in phrase-final position.

The discovery of this diachronic change motivates the investigation described in Paper II, which uses carefully-controlled lab speech data to probe a number of collinear factors that may be conditioning this pre-boundary innovation. Specifically, three potential factors are explored: the strength of the prosodic boundary following (ng), the duration of the velar nasal (which is modulated by boundary strength), and the presence and duration of a pause marking the prosodic juncture.

In Paper II the question is also raised regarding the extent to which this change in the pre-pausal environment is socially motivated, i.e. whether or not this move towards increased [ɡ]-presence in the highly-salient pre-pausal environment reflects a desire to be associated with the positive values indexed by use of this northern variant. This possibility in part provides the context and motivation for Paper III,
which employs a matched-guise technique to investigate the social meaning of (ng) and in turn provide direct perceptual evidence to determine whether or not attitudinal factors could play a role in this variation.

Paper IV is somewhat more independent than the other three in that it is the first and only study presented in this thesis that investigates unstressed -ing suffixes, e.g. walk-ing. However, although the analysis considers factors that influence the use of the alveolar [ɪn] variant, and situates these results in the context of previous work on (ing), this paper still addresses the variable presence of post-nasal [ɡ] and in particular whether or not its occurrence in this environment is constrained by the same factors as in the stressed (ng) environment.

These four papers are followed by Chapter 6, in which I provide a discussion of all the results from the four papers in light of each other. In this chapter I also present a reanalysis of the cross-linguistic status of /ŋɡ/ clusters in line with the life cycle of phonological processes, and finally offer thoughts for future avenues of research.

1.3 Research outcomes

The results of Paper I indicate that variation in (ng) is almost completely accounted for simply by assuming cyclic application of /ɡ/-deletion across stem-, word-, and phrase-level morphosyntactic domains. This lends empirical support to cyclic models of phonological computation and, by consequence, accounts of sound change such as the life cycle of phonological processes (Bermúdez-Otero 2013). Furthermore, the correlation between the depth of the stratum and the rate at which deletion applies in it suggests that post-nasal /ɡ/-deletion did indeed follow the systematic progression through the grammar consistent with a life cycle analysis that involves successive rounds of domain narrowing.

The analysis also reveals that the behaviour of (ng) in pre-pausal position is still involved in change: phrase-final [ɡ]-presence is increasing in apparent time, to such an extent that younger speakers now appear to have a categorical ban on pre-pausal [ŋ] without a following velar stop. I propose that this reflects an innovative process of [ɡ]-insertion that applies before pause, superimposed upon the original cyclic deletion rule. Results from a separate elicitation task described in Paper II also suggest that this innovation does not reflect a more gradient property of duration, i.e. how pre-boundary lengthening of the velar nasal itself might modulate the relative timing of velum raising and release of the lingual gesture, determining whether or not a post-nasal stop will be articulated. Instead, this pre-boundary innovation seems to be conditioned primarily by pause itself, independent of segmental duration and to
some extent prosodic phrasing.

Crucially, the incrementation of this change does not appear to be externally motivated by attitudinal factors. Contrary to existing claims in the literature (e.g. Beal 2004: 127), the results of Paper III find no evidence that [ŋɡ] has local prestige. In fact, there is no evidence of any shared norm across the speech community with respect to this variable. It does not appear to be a salient marker of northern dialects and, even though awareness of its dialectal status is increasing over time, this is still not leading to uniform evaluation. The magnitude of inter-speaker variation with regard to norm alignment makes an evaluation-driven account of this change unlikely.

Alternatively, I propose that this change in pre-pausal [ɡ]-presence reflects a more general velar fortition rule before pause, the outcome of which is also seen in phrase-final /ŋk/-ejectivisation in Glasgow English (McCarthy & Stuart-Smith 2013); functionally, this could be utilised by speakers as a turn-taking or boundary-marking device in the navigation of conversation, adding to existing cues such as pitch (Pierrehumbert & Hirschberg 1990; Swerts 1997), segmental lengthening (Lehiste et al. 1976; Gussenhoven & Rietveld 1992; Wightman et al. 1992), and non-modal voice quality (Cutler & Pearson 1985; Gordon & Ladefoged 2001; Ogden 2004; Garellek 2015).

Paper IV, which addresses the three-way variation between [ɪn]~[ɪŋ]~[ɪŋɡ] in the unstressed (ing) environment, presents one of the first detailed analyses of the phonetic realisation of these variants. Focusing on two properties that seem to differentiate the alveolar and velar forms – namely the vowel quality of [ɪ] and the duration of the following nasal segment – the results suggest that (ing) is in fact a case of lenition. The alveolar variant involves less articulatory effort than [ɪŋ(ɡ)] by virtue of its laxer vowel and shorter nasal. The data also reveals that this variable is not at all sensitive to part of speech, and by consequence does not reflect historical morphology as it does in other communities. Furthermore, (ing) appears to be much less of a ‘social’ variable in the North West of England, and is instead largely predicted by mechanical factors such as speech rate and assimilation with adjacent segments. Crucially, [ɪŋɡ] is rarely used in conversation but is more likely to occur before pause relative to all other phonetic environments; this indicates that the process(es) giving rise to /ŋɡ/ variation operate in the same manner in both environments, and serves to further highlight this exceptional pre-pausal behaviour.
Chapter 2

Velar nasals

Within the consonantal inventory of English, the velar nasal occupies something of a unique place: its positional restrictions – namely that it is forbidden in syllable onsets – set it apart from the other nasal segments, and its phonemic status is contested to a greater extent than that of any other sound in the language.

It is also subject to variation in two major ways. In unstressed syllables, the velar nasal alternates with [n] in what is arguably the most geographically-widespread and widely-studied sociolinguistic variable in contemporary varieties of English. This variation, described in more detail in Section 2.2, will hereafter be notated as (ing) following standard sociolinguistic convention.

In certain regions of England, namely parts of the North West and West Midlands, velar nasals are variable in another respect: they lack obligatory ng-coalescence, which results in variable post-nasal [ɡ]-presence in words such as young [jʌŋɡ]~[jʊŋ] and singer [ˈsɪŋɡə]~[ˈsɪŋə]. In these dialects, finger and singer are perfect rhymes. These two sources of variation have partially overlapping environments, as [ɡ] can also follow the velar nasal in unstressed (ing) clusters. The result is that, for these speakers, (ing) is in fact a case of ternary variation between [n], [ŋ], and [ŋɡ]. The variation in this environment will be notated as (ng), and the distinction between these two variables is made explicit in Table 2.1. Due to the productive nature of the -ing suffix in English, these two environments can and often do co-occur within the same word: e.g. singing, banging, ringing etc. This leads to a situation in which there are six possible variant pairs, as illustrated in Table 2.2.

In this chapter I provide an overview of the variation in these two environments – (ng) in Section 2.1 and (ing) in Section 2.2 – covering not only what has been said about their synchronic patterns of variation but also the proposals regarding their

---

1It should be noted that, among other features, Cockney English is characterised by a nasal+stop cluster in words like anything [ˈɛnɪθɪŋk] and something [ˈsʌmθɪŋk] (Cruttenden 2014: 216). However, this is lexically restricted to the quantifiers and should not be conflated with the nasal+stop clusters in northern varieties.
Table 2.1: Overview of the (ing) and (ng) variants.

<table>
<thead>
<tr>
<th>variable</th>
<th>word</th>
<th>[n]</th>
<th>[ŋ]</th>
<th>[ŋɡ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(ing)</td>
<td>waiting</td>
<td>[weɪtɪŋ]</td>
<td>[weɪtŋ]</td>
<td>[weɪtŋɡ]</td>
</tr>
<tr>
<td></td>
<td>ceiling</td>
<td>[siːlɪŋ]</td>
<td>[siːlɪŋ]</td>
<td>[siːlɪŋɡ]</td>
</tr>
<tr>
<td>(ng)</td>
<td>tongue</td>
<td>—</td>
<td>[tɒŋ]</td>
<td>[tɒŋɡ]</td>
</tr>
<tr>
<td></td>
<td>singer</td>
<td>—</td>
<td>[sɪŋə]</td>
<td>[sɪŋɡə]</td>
</tr>
</tbody>
</table>

Table 2.2: The six possible variants for a word subject to both (ing) and (ng) variation, such as *ringing*; columns correspond to (ing) variants, rows to (ng) variants.

<table>
<thead>
<tr>
<th>[n]</th>
<th>[ŋ]</th>
<th>[ŋɡ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ŋ]</td>
<td>[ɹɪŋɪŋ]</td>
<td>[ɹɪŋɪŋɡ]</td>
</tr>
<tr>
<td>[ŋɡ]</td>
<td>[ɹɪŋɡɪŋ]</td>
<td>[ɹɪŋɡɪŋɡ]</td>
</tr>
</tbody>
</table>

historical origins. In Section 2.3 I also discuss the implications of post-nasal [ɡ]-presence for the phonemic status of the velar nasal in English; this is supplemented with an overview of the cross-linguistic behaviour of [ŋɡ] clusters in other Germanic languages in Section 2.4, and in particular the status of the velar nasal in German.

### 2.1 Variation in (ng)

The 1950s *Survey of English Dialects*, published by Orton et al. (1978), attests post-nasal [ɡ]-presence in words like *tongue* across the southern part of the North West of England and the western part of the Midlands. The region covered by this isogloss, which is illustrated in Figure 2.1a, has been described as “most of the western half of the Midlands and Middle North, including Birmingham, Coventry, Stoke-on-Trent, Manchester and Liverpool” (Wells 1982b: 365). The regional distribution of this variant is corroborated with an even earlier dialectal survey from the 1870s (Ellis 1889; Maguire 2012) as well as more recent survey data collected between 2013–2017, which is mapped in Figure 2.1b (MacKenzie et al. 2017). The presence of post-nasal [ɡ] has also been reported in a wide range of descriptive dialectological literature (Wake- lin 1984; Trudgill 1999; Hughes et al. 2012) and specific sociophonetic studies both in the North West (Knowles 1973; Newbrook 1999; Watts 2005; Schleef et al. 2015) and in the West Midlands (Heath 1980; Mathisen 1999; Thorne 2003; Asprey 2015).


3While post-nasal [ɡ]-presence is largely restricted to the North West and West Midlands of England, there is also a small pocket of [ŋɡ] use in Maidstone in the South of England (illustrated in Figure 2.1a; Orton et al. 1978). Furthermore, non-coalesced forms appear sporadically in other regions outside of England, including Utah (Alzoubi et al. 2013), New York (Gordon 2004: 289; Newman 2014: 87), and Toronto (Walker 2013, 2017). It is entirely plausible that [ɡ]-presence in these varieties, particularly outside of England, reflects an independent innovation rather than retention of a historical form.
Figure 2.1: The regional distribution of [ŋɡ]: (a) the 1950s *Survey of English Dialects* isogloss for [ŋɡ] in *tongue* (adapted from Ph242 in Orton et al. 1978), and (b) the 2010s *UK Dialect Maps* isogloss for [ŋɡ] in *singer* (adapted from MacKenzie et al. 2017). In (b), the data has been smoothed using local spatial autocorrelation (Ord & Getis 1995), with higher values corresponding to more [ŋɡ].
This lack of ng-coalescence was not always so geographically restricted: Wells (1982a: 188) notes that it was only around 1600 when the post-nasal [ɡ] “ceased to be pronounced in educated London English”. Before this time, all varieties of English featured non-coalesced [ŋɡ] clusters. This change, according to Wells, involved the addition of a synchronic rule to speaker’s grammars that deletes word-final /ɡ/ – or word-medial /ɡ/ before a morpheme boundary – when preceded by a nasal. Later work has revealed that this change was not so abrupt, but that it progressed along a structured pathway of change. As discussed by Bermúdez-Otero & Trousdale (2012), there is suggestive evidence of intra-speaker variation around the time of the 18th century, with post-nasal [ɡ] still present in certain contexts. Along with Garrett & Blevins (2009), they cite suggestive evidence from the orthoepist James Elphinston, who retains word-final pre-vocalic [ŋɡ] in his formal, conservative register (e.g. *sing aloud* [sɪŋɡ əlaʊd]), but deletes it in this position in his more casual, innovative register, which is said to reflect a more advanced stage of the change. In both styles, Elphinston deletes /ɡ/ in word-final pre-consonantal position (e.g. *sing tunes* [sɪŋ tjuːnz]), suggesting that deletion in this environment represents an earlier stage of the change.

The importance of these distinctions will be addressed in Chapter 4 and in Paper I, which offer a more in-depth discussion of this diachronic change. For now, let us simply consider the end-point of this change as it stands now in Present Day English: for speakers outside of the North West and West Midlands, monomorphemic words such as *finger* [fɪŋɡə] invariably surface with [ɡ]-presence, and polymorphemic words such as *singer* [sɪŋə] invariably surface with [ɡ]-absence. The former category of words where /ŋɡ/ is not word- or stem-final – other examples being *tango*, *single*, *angle* etc. – are the only examples of non-coalesced clusters in these Present Day English varieties. Crucially, this process of post-nasal /ɡ/-deletion has remained variable for speakers in the North West and West Midlands of England, who exhibit variation between [ŋ] and [ŋɡ] to this day. How this variation patterns along sociodemographic dimensions, and in particular how it is constrained by language-internal factors, is relatively under-studied – with the exception of Mathisen (1999) in the West Midlands and Watts (2005) in Cheshire – although it has been proposed that the non-coalesced [ŋɡ] variant has local prestige and is considered ‘posh’ (Beal 2006: 48; Foulkes & Docherty 2007: 64). This dissertation will lend further insight into (ng)

---

4In addition to this category of words, there are also a number of listed exceptions: the comparative and superlative forms of the high-frequency adjectives *long*, *strong*, and *young* consistently surface with [ɡ]-presence contrary to what we would expect given the stem-final position of post-nasal /ɡ/ (Bermúdez-Otero 2012: 70–71). Additionally, [ɡ] is invariably present for all speakers in root-based items in which morpheme-final /ŋɡ/ occurs before a stem-level suffix, such as *elongate* (Bermúdez-Otero 2011: 2021–2022).
variation in contemporary northern varieties, and in addition will directly investigate the evaluation of [ŋɡ] to provide empirical evidence of local attitudes towards this form.

2.2 Variation in (ing)

Velar nasals are subject to one other major aspect of variation in English: they often surface as an alveolar variant in unstressed /ŋ/ clusters, regardless of whether this cluster corresponds to a synchronically segmentable -ing suffix transparently attached to a stem (e.g. *talking* [tɔːkɪŋ]-[tɔːkɪn]) or to an item that is monomorphic or at best nonproductively based on a bound root (e.g. *pudding* [pʊdɪŋ]-[pʊdɪn]). Of course, this environment also overlaps with the environment for post-nasal /ɡ/-deletion, resulting in three-way variation for speakers in the North West and West Midlands of England; that is, for these speakers, a third variant of [tɔːkɪŋɡ] or [pʊdɪŋɡ] is also possible.

Variation between [ɪn] and [ɪŋ], often notated as -in and -ing, is not subject to the same restricted regional distribution as [ŋɡ]; in fact, it has been attested all across the English-speaking world, including varieties spoken in the United Kingdom (Tagliamonte 2004; Watts 2005; Schleef et al. 2011; Levon & Fox 2014), the United States, (Huspek 1986; Labov 2001; Wagner 2012c; Tamminga 2016), Canada (Woods 1979; Walker 2016), and Australia (Shopen 1978; Shnukal 1982). It also exhibits very regular and consistent patterns of conditioning factors with respect to both sociodemographic and language-internal predictors. The velar [ɪŋ] variant is perceived to be the ‘prestige’ variant, favoured in more formal discourse styles and used more often by women, by middle-aged speakers, and by those of higher socioeconomic status (see Hazen 2006 for an overview). This variable is also highly salient, to such an extent that it is often reflected orthographically (e.g. <walkin’>) and subject to a great deal of metalinguistic commentary: a number of studies have investigated these perceptions of (ing) directly, revealing that use of [ɪn] – while stigmatised (Labov et al. 2006, 2011) – is associated with a wide range of indexical meanings (Campbell-Kibler 2006, 2011b).

The variation is also said to be subject to both regressive assimilation and progressive dissimilation, such that a following velar consonant favours use of [ɪŋ] but a preceding velar consonant favours use of [ɪn] instead (Shuy et al. 1968; Cofer 1972; Houston 1985; Forrest 2017). However, by far the strongest and most widely-reported linguistic predictor of (ing) is the part of speech of the token in which it occurs: [ɪn] is more frequent in verbal tokens, while [ɪŋ] is favoured in more nominal categories.
(Houston 1985; Huspek 1986; Labov 2001; Tagliamonte 2004).

Of particular note is how this grammatical conditioning of its synchronic variation has been connected to the supposed origin of the competition between these two forms. Visser (1966) tracks this development back to the Old English period, in which two distinct suffixes were used to mark participles and verbal nouns: -inde and -ynge, respectively. These suffixes later underwent a number of reductive phonological processes, including reduction and then deletion of the final vowel and coalescence of the nasal-stop cluster. In the case of -ynge, as we know from the discussion in the preceding section, /ŋɡ/-coalescence never ran to completion in some regions of the country. Nevertheless, the result of these lenition processes led to increased phonetic similarity, which could have been partly responsible for the functional conflation of these two suffixes. This conflation – which saw both the velar and alveolar forms begin to be used in both verbal and nominal constructions – is said to have occurred first in the South of the country, resulting in inter-dialectal differences in the magnitude of this synchronic grammatical conditioning effect (Houston 1985, 1991).

Given the historical origin of the synchronic (ing) variable, it is tempting to analyse the variation between [m] and [n] as an alternation between allomorphs – which just so happen to only differ with respect to the place of articulation of the nasal – rather than between allophones. This would clearly reflect the fact that this variation originates from two distinct grammatical morphemes in the Old English period, and how that distinction is still registered today in its grammatical conditioning. However, this account relies on a segmentable -ing suffix to act as the locus of variation, which requires us to treat items such as pudding and morning as morphologically complex despite the fact that they are synchronically non-compositional. On the other hand, the sensitivity of (ing) to its immediate phonological context might lend support to the idea that this is indeed a phonological variable, which could at least produce variation in both monomorphemic and polymorphemic contexts. However, this has been disputed by Labov (2001: 87), who reports no significant phonological conditioning in Philadelphia and claims that “(ing) has all the hallmarks of a morphological alternation rather than a phonological reduction”.

Recent evidence from Tamminga (2016) speaks to exactly this question, with reports of a significant priming effect such that the choice of one variant in one instance increases the probability of using that same variant next time. Crucially, the results suggest that priming effects are only present when the two (ing) tokens are matched for morphological composition. For example, use of [m] in a polymorphemic (ing)
token will increase the likelihood of using [ɪn] again for the next instance of poly-
morphemic (ing), but has no influence on the choice of variant if the following (ing)
token is monomorphemic instead. This leads Tamminga (2016) to claim that (ing) vari-
ation in contemporary dialects of English actually represents two distinct processes:
a phonological alternation in the case of monomorphemic items, but a morphological
alternation in cases of polymorphemic items\(^6\). This corroborates earlier claims, that
“as a sociolinguistic variable, (ING) is a coherent whole with two variants; however,
(ING) is not the result of a single linguistic process” (Hazen 2006: 583).

This question will be returned to in the discussion of Paper IV’s results, but for
now I remain theory-neutral with respect to the locus of this variation, particularly as
such priming effects may be dialect-specific and it has yet to be established whether
or not morphological composition plays any role in the variation of (ing) in Northern
Englishes.

2.3 Phonemic status

The velar nasal is arguably one of the most interesting segments of English in terms
of phonemic status, and is made even more interesting due to the regional differences
in its positional distribution.

Given the existence of minimal triplets such as *ran* [ɹan], *ram* [ɹam], and *rang*
[ɹaŋ] in almost all varieties of Present Day English, it may seem fairly uncontroversial
to claim that English has three nasal phonemes differing in their place of articulation
(Giegerich 1992: 36–37).

As noted by Hogg (2002: 9) and Cruttenden (2014: 48), this phonemic opposition
between the velar nasal and the other nasal segments of English was of course only
initiated when post-nasal [ɡ] was lost as a result of the historical process discussed
earlier in Section 2.1. What does this mean, then, for speakers of Northern English
varieties who still exhibit post-nasal [ɡ]-presence? An alternative account is possible:
in these cases, the velar nasal only ever occurs before a velar stop (e.g. *sink* [sɪŋk]
or *sing* [sɪŋɡ]), and can therefore be analysed as a positional allophone of /n/ rather
than a phoneme itself.

Of course, this allophonic analysis is still possible for all speakers of English, re-
gardless of the surface realisation of these (ng) words. It would simply require posit-
ing an underlying /ɡ/ and a combination of nasal place assimilation and post-nasal
/ɡ/-deletion, where the latter process applies across the board for most speakers and

\(^6\)A similar account has been proposed for (td)-deletion, whereby the surface alternation reflects not
only a phonological variable but also variation in whether or not the /t/ is morphologically present
(Fruehwald 2012).
only variably for speakers of North West and West Midlands English.

The requirement of a nasal place assimilation process is not problematic. There is enough evidence to suggest that this phenomenon is widespread in English, such that it can even apply across word boundaries\(^7\) – the preposition \(\text{in}\) is \([\text{ɪm}]\) before alveolars (e.g. \(\text{in Tooting}\)), but often \([\text{ɪm}]\) before bilabials (e.g. \(\text{in Blackburn}\)) and \([\text{ɪn}]\) before velars (e.g. \(\text{in Kent}\)) (Gussenhoven & Jacobs 2011: 70) – as well as to all coronal consonants more generally (e.g. \(\text{this shop} \rightarrow [\text{ðɪʃ ʃɒp}]\)) (Cruttenden 2014: 312). We also know a great deal about the exact environment in which nasal place assimilation takes place: an analysis by Kiparsky (1979) (see also Halle & Mohanan 1985: 62; Anderson & Ewen 1987: 65) suggests that obligatory categorical nasal place assimilation only takes place when the nasal and following consonant are both contained within the same foot, accounting for the difference between pairs such as \(\text{cόngress} [\text{kon̩ɡɹəs}]\) (invariably \([\text{n}]\)) and \(\text{congressional} [\text{kən'ɡɹɛʃn̩əl}]\) (invariably \([\text{n}]\)).

Given this independent evidence, it is fairly uncontroversial to assume a process of nasal place assimilation in English. However, do we want to also assume that these speakers of non-northern varieties have an underlying \(/\text{ɡ}/\) in words like \(\text{sing}\) and \(\text{longing}\) if it is never actually present in their realisation of these words? The earlier discussion in Section 2.1 highlighted how post-nasal \(/\text{ɡ}/\) was historically present in all varieties of British English prior to the initiation of ng-coalescence in the Late Modern English period, but the synchronic system does not necessarily have to reflect that. Wells (1984) does not commit to either account, simply acknowledging that the argument for a purely allophonic \([\text{n}]\) receives stronger support in the North West and West Midlands of England compared to other regions, because for these speakers at least the post-nasal \([\text{ɡ}]\) is actually present.

This account has proved to be popular in generative phonology (e.g. Chomsky & Halle 1968; Kahn 1976; Withgott 1982; Halle & Mohanan 1985; Borowsky 1986), but the idea actually predates the advent of this approach: in discussing the English velar nasal, Sapir (1925: 49) notes how “no naïve English-speaking person can be made to feel in his bones that it belongs to a single series with \(m\) and \(n\)”.

Nevertheless, there are a number of complications that come with an allophonic

\(^7\)It should be noted that assimilation across word boundaries is optional and can vary in its implementation, involving categorical delinking and spreading in some cases but only gradient gestural overlap in others: see Bermúdez-Otero & Trousdale’s (2012: 694–695) discussion of the articulatory evidence provided by Ellis & Hardcastle (2002).

\(^8\)It has been claimed, based on their vulnerability to phrase-level assimilation, that coronals are unspecified for place of articulation features (Avery & Rice 1989); under this account, the underlying representation does not specify a place of articulation, which instead is inherited from its surrounding segmental environment. Non-coronals, on the other hand, are specified for place of articulation and as a result do not exhibit the same vulnerability to assimilation (although see Coleman et al. 2016 for counter-evidence to this).
analysis of the velar nasal. A thorough discussion is provided by Wells (1982a), who cites the English velar nasal as a borderline case of phonemic status. As noted in Section 2.1, the comparative and superlative forms of strong, long, and young all retain post-nasal [g] despite it appearing in the environment for deletion. Wells highlights that these are simply exceptions, and that [g]-absence is the norm, based on the fact that forcing native speakers to produce novel words such as a comparative and superlative form of wrong (i.e. wronger and wrongest) usually results in a production without [g] (unless, of course, the speakers are from the North West or West Midlands of England). A similar line of argumentation has been put forward by Chomsky & Halle (1968: 370), who give the examples cunningest [ˈkʌnɪŋəst] and willingest [ˈwɪlɪŋəst], and by Bermúdez-Otero (2012: 70–71), citing the example winningest [ˈwɪnɪŋəst].

Other scholars working under a generative framework have attempted to account for these exceptions in a number of different ways: Anderson & Ewen (1987) claim that /ɡ/ is retained in linger (and other finger-type items) because it is ambisyllabic, but that ambisyllabcity is blocked by the morpheme boundary in singer, hence the application of /ɡ/-deletion here. The exceptional behaviour of longer and stronger stem from the fact that in these cases the morphological boundaries “are not phonologically relevant” (Anderson & Ewen 1987: 62). Alternatively, Halle & Mohanan (1985: 62) account for the exceptional behaviour of these comparative and superlative forms by assuming that for these words the suffixation of inflectional -er and -est occurs in an earlier stratum. This means that these words undergo bracket erasure – removing post-nasal /ɡ/ from its vulnerable constituent-final position – before /ɡ/-deletion can apply.

However, other words exist that are exceptional in a different way: the words gingham, Langham, and dinghy all unexpectedly surface without post-nasal [g], in spite of what we would predict given their monomorphemic composition (Wells 1982a). In accounting for this exceptional behaviour, Chomsky & Halle (1968: 234) actually propose an underlying /nx/ cluster in words like dinghy and gingham, with assimilation of /n/ to [ŋ] triggered by the following velar fricative, which itself undergoes deletion. A more plausible account is suggested by Gussmann (2000: 507), who notes that unexpected deletion in place names like Birmingham and Langham could reflect speakers analysing these words as morphologically complex, which they were historically.9 If this is the case – and we do have independent evidence that place

---

9In place names such as Altrincham, Rotherham, Buckingham etc. -ham is historically a morpheme meaning ‘homestead’; for example, Nottingham was historically Snot-inga-hām, literally ‘home of Snot’s people’ (an Anglo-Saxon chieftain who ruled the settlement around the seventh century) (Gurnham 2010). Likewise, Birmingham was historically Beorm-inga-hām, literally ‘home of Beorma’s
names in Dutch, at least, are synchronically complex (Köhnlein 2015) – then post-
nasal /ɡ/ would be syllable final in a phonological domain and therefore eligible for
deletion. Even if this is not the case and these words are synchronically monomor-
phemic, Gussmann suggests that their underlying representations could contain /h/,
blocking syllabification of the preceding /ɡ/ into the onset and consequently leaving
it vulnerable to deletion\textsuperscript{10}.\footnote{Gussmann (2000) also notes that in some varieties, such as American English, words like Birmingham
do actually surface with an audible [h] and, for some speakers at least, with an unreduced vowel, e.g. [hæm].}

Gussmann (2000) also suggests that the exceptional [ɡ]-absence in words like
dinghy and hangar could stem from analogy with words like thing-
y and hang-er, both of which are morphologically complex, leaving the /ɡ/ eligible for deletion. Yet, there remain a number of other exceptions such as [ɡ]-less Singapore ['sɪŋəpɔː], for which none of these explanations are applicable.

In sum, it is clear that there is no established consensus on the phonemic status
of the velar nasal in English. To a large extent the choice between these competing
theories is dependent on a researcher’s own theoretical standpoint regarding abstract
underlying representations and the acceptability of lexical exceptions to this rule-
based framework. However, it is clearly less controversial to reject a phonemic velar
nasal, and instead posit underlying /ɡ/ with processes of nasal assimilation and post-
nasal /ɡ/-deletion, in varieties of northern English where the [ɡ] is often present on
the surface. This account is strengthened by the investigation described in Paper I,
which presents evidence that synchronic (ng) variation closely reflects the diachronic
change discussed earlier in Section 2.1.

### 2.4 Typological status of the velar nasal

The velar nasal exhibits interesting distributional properties in two distinct ways.
Firstly, it is highly clustered with respect to its areal distribution across the world’s
languages. Of the 469 languages mapped in the World Atlas of Language Structures
(WALS), just over half (235) have no phonemic velar nasal; of those that do, approxi-
mately 37% of them do not allow it to appear word-initially (Anderson 2013). This
is mapped in Figure 2.2. Secondly, there is cross-linguistic variation with respect to
its distribution within languages in terms of the environments in which it appears
alongside a homorganic velar stop. While this might suggest cross-linguistic differ-
ences in how post-nasal /ɡ/-deletion operates, providing a detailed analysis of [ŋɡ]
clusters in these languages is beyond the scope of this thesis. In Chapter 6 I propose
2.4.1 The German velar nasal

In Section 2.3 I highlighted the contested phonemic status of the velar nasal in English, but the debate is even greater in the case of German. Like English, German has a three-way distinction of nasals in word-final position, but many scholars propose that the velar nasal is not phonemic, and is instead derived from an underlying sequence of /Nɡ/ with rules of nasal place assimilation and post-nasal /ɡ/-deletion (Vennemann 1968, 1970; Dressler 1981; Kloeke 1982; Hall 1989; Moltmann 1990; Wiese 2000; Ito & Mester 2003). This deletion rule, like in English, is said to delete underlying /ɡ/ in coda position when preceded by a nasal.

One line of evidence, discussed by Kloeke (1982), is that the final /n/ of German loanword Balkon ‘balcony’ can be realised as either alveolar or velar; crucially, if it surfaces as alveolar, the preceding vowel is long [oː], but if it surfaces as a velar nasal the preceding vowel is short [o]. According to Kloeke (1982), this is to be expected if the velar nasal is in fact underlyingly a sequence of two segments, /Nɡ/ and we assume a constraint against long vowels followed by complex codas.

The exact manner in which this deletion rule operates is also subject to debate. Boase-Beier & Lodge (2003) claim that the difference between words like evangelisch [evaŋˈɡeːlɪʃ] ‘Protestant’, which has [ɡ]-presence, and words like Enge [ˈɛŋə] ‘narrow-
ness’, which doesn’t, is because in the latter case the [ŋ] is ambisyllabic; the former example, on the other hand, features a syllable boundary intervening between the nasal+stop cluster. It should be noted, however, that Boase-Beier & Lodge (2003) do not agree with an abstract analysis of /g/-deletion in the first place. The authors suggest instead that German does have phonemic /ŋ/, and that surface velar nasals come from two sources: some are simply allophones of /n/, derived through nasal place assimilation with a following [ɡ] or [k] that is present on the surface, but some are also present in the underlying representation.

Wiese (2000: 224–226) does argue for an underlying /Nɡ/ sequence and a combination of nasal place assimilation and /g/-deletion, but draws attention to the fact that the alternation between [ŋ] and [ŋɡ] in German cannot be attributed to the presence of a morpheme boundary: there are examples of stem-final [ɡ]-presence (e.g. fing-ier-en [finˈɡiːrən] ‘to fake’) and stem-final [ɡ]-absence (e.g. läng-lich [ˈlɛŋlɪç] ‘long’).

The explanation proposed by Wiese (2000: 225) relies on the assumption of “non-syllabification of consonants which eventually surface in schwa syllables”, accounting for the application of deletion in words like Finger [ˈfɪŋə] ‘finger’ and the non-application in words like Tango [ˈtaŋɡo] ‘tango’. In Finger (and polymorphemic words like Ding-e [ˈdɪŋə] ‘thing.pl’), the schwa is represented as an empty prosodic position and remains detached from the skeletal tier in the first round of syllabification, leaving /ɡ/ final in the syllable and therefore vulnerable to deletion; the schwa attaches to the prosodic tree later in the derivation after a second pass of syllabification. On the other hand, in a word like Tango, or polymorphemic fing-ier-en, the vowel following underlying /Nɡ/ is not reduced; this means that it is syllabified in the first pass, and subsequently /ɡ/ can move into the onset and thus be saved from deletion. This proposal, which is also supported by Hall (1989), is illustrated in Figure 2.3.

However, Wiese (2000: 226) openly admits that this non-syllabification of schwa contradicts a requirement stipulated elsewhere in the same body of work, namely that schwa syllables are present in the lexicon: “one set of facts seems to require syllabification of final material, while another set of facts seems to argue, with equal force, for its extrasyllabic, or at least invisible, status”.

Whatever the reasons behind the deletion of post-nasal /ɡ/ in unstressed schwa syllables (see Chapter 6 for my own proposal), it is clear to see that the deletion process in German applies to a wider set of environments than it does in English, where the /ɡ/ is invariably present in monomorphemic words such as finger.
Figure 2.3: Derivation of Dinge and Tango in German, illustrating the initial non-syllabification of schwa. Adapted from Wiese (2000: 225).
2.4.2 Velar nasals in other Germanic languages

Other Germanic languages show the same positional restrictions on the velar nasal, namely that it only appears syllable-finally and after short vowels. However, this has not led to a unanimous treatment of the velar nasal as an underlying sequence of /ŋɡ/ with nasal place assimilation and post-nasal /ɡ/-deletion as discussed for German.

In Dutch, for example, Booij (1995: 35) notes that while its restriction from appearing in onsets has led some scholars to argue for an underlyingly /Ng/ sequence (see e.g. Kager 1989), such an analysis has been shown to be ‘untenable’ by others (Booij 1980; Trommelen 1983). Dutch is, however, similar to German with respect to the environments in which [ŋɡ] occurs: for example, the [ɡ] is absent in monomorphic words such as engel [ɛŋəl] ‘angel’ (cf. the more conservative deletion rule in English, which does not apply in a word like angle [æŋɡəl]). That is, Dutch mirrors German in forbidding post-nasal [ɡ]-presence in pre-schwa position, regardless of morphological composition.

In Danish, Basbøll (2005: 75) argues for an underlying /nɡ/ sequence, proposing that ‘at the morphophonemic level the bisegmental interpretation of the velar nasal is the one to be adopted’. This is based on the existence of morphological alternations as in (1). Crucially, Basbøll claims that [ɡ]-presence in these cases is determined by stress, with deletion blocked if the /ɡ/ can syllabify as the onset of a stressed syllable.

(1) a. diftong [diftʰʌŋ] ‘diphthong’
   b. diftongere [diftʰʌŋˈɡ̊eːɾə] ‘to diphthongize’
   c. Kolding [ˈkʰʌleŋ] ‘[name of town]’
   d. koldingenser [kʰʌleŋˈɡ̊ɛnˈsɐ] ‘inhabitant of Kolding’

A similar situation is found in Norwegian. Standwell (1975: 350–351), who also argues for an underlying /ŋɡ/ sequence, notes that /ɡ/-deletion is blocked before stressed suffixes, citing examples such as tangere [taŋˈɡɛɾə] ‘to touch’ and fungere [funˈɡɛɾə] ‘to function’. On the other hand, Kristoffersen (2000: 39–40) offers an alternative proposal, claiming that /ŋ/ is present underlingly in Norwegian and is therefore phonemic; this is based on the fact that such alternations between [ŋ] and [ŋɡ] – for example in diftong [diftɔŋ] ‘diphthong’ and diftongere [diftɔŋˈɡ̊ɛɾə] ‘diphthongise’ – are exceedingly rare. Kristoffersen instead analyses these alternations as cases of lexical allomorphy.

For Swedish, Riad (2013) follows the same argumentation as Kristoffersen (2000) in claiming that the velar nasal is phonemic. Interestingly, the same alternations from Danish and Norwegian are present again in Swedish, with post-nasal [ɡ] invariably present in foot-initial position. This is illustrated by the examples in (2).
(2)  a. harang [haˈranː] ‘harangue.sg’
   b. haranger [haˈranːɛ̠r] ‘harangue.pl’
   c. harangera [haraŋˈɡeːra] ‘to harangue’
   d. diftong [dɪfˈtɔŋː] ‘diphthong.sg’
   e. diftonger [dɪfˈtɔŋːɛ̠r] ‘diphthong.pl’
   f. diftongera [dɪftɔŋˈɡeːra] ‘to diphthongise’

However, while Basbøll (2005) claims that this reflects a blocking of /ɡ/-deletion in Danish, and Kristoffersen (2000) dismisses it as lexical allomorphy in Norwegian, Riad (2013) makes yet another proposal for Swedish: that this reflects a separate process of [ɡ]-excrescence, which inserts [ɡ] in foot-initial position

11. Put another way, the same distributional pattern has led to one account that assumes underlying /ŋ/ with [ɡ]-deletion, one that assumes underlying /ng/ with /ɡ/-deletion, and one that posits no phonological process at all.

The status of the velar nasal is not so contested in other languages. Yiddish, for example, serves as an extreme case in which ng-coalescence does not take place at all. Jacobs (2005: 113) claims that the velar nasal is completely allophonic, as it only ever occurs through assimilation with a velar consonant following the nasal, as in zing [ziŋɡ] ‘(I) sing’, or preceding it, e.g. zogn [zoɡŋ] ‘say’

12. Crucially, this assimilation is perfectly transparent since the velar stop is always present on the surface, even word-finally (recall that every other language discussed thus far forbids word-final [ŋ]). It should be noted, though, that there have been reports of word-final /ɡ/-deletion in some varieties of Yiddish, e.g. zing [ziŋ] (Prilutski 1940, cited in Jacobs 2005: 113), and that Herzog et al. (1992: 41) argue for phonemic /ŋ/ in Samogitia and Courland Yiddish. In standard varieties, at least, an allophonic analysis of [ŋ] is difficult to contest.

The velar nasal has a similar distribution in both Icelandic and Faroese. Árnason (2011: 109) notes that velar (and palatal) nasals only occur in coda position before a homorganic stop, e.g. Icelandic langur [laʊŋkʏr] ‘long.masc’, banka [pauŋka] ‘bank’, and Faroese longur [lɔŋkʊr] ‘long’. In Faroese, velar nasals can also occur without a homorganic stop under certain conditions. As the examples in (3) demonstrate, even in these environments [ŋ] is still amenable to an analysis under which it is derived from an underlying velar stop, but this is opaque on the surface due to metathesis and cluster reduction. That is, these velar nasals are derived through a sequence of

11. The same argument is proposed by Eliasson (1986: 276), who also applies it to the Norwegian case, contra Kristoffersen (2000).

12. It should also be noted that German exhibits a similar process of progressive nasal assimilation, such that the infinitive -en suffix surfaces with velar place of articulation in words like hängen [hɛŋŋ̍] ‘to hang’ (Wiese 2000: 223).
metathesis between /s/ and /k/, which moves the velar stop adjacent to the preceding nasal segment, followed by nasal place assimilation and finally deletion of the velar stop (Árnason 2011: 177–178). Metathesis does not take place in (3c) or (3d), and as such the nasal surfaces as alveolar in these cases.

(3)  
a. *danskt* [taw̃st] 'Danish.neut’
b. *enskt* [ɛaw̃st] 'English.neut’
c. *danskur* [taŋ̊skʊɹ] 'Danish.masc’
d. *enskur* [ɛŋ̊skʊɹ] 'English.masc’

This section has highlighted not only the magnitude of cross-linguistic variability with respect to the environments in which velar nasal+stop clusters are reduced, but also how the same (or similar) distributional patterns have led to completely different analyses and assumptions regarding the underlying representation of these clusters. The extent to which these various distributional patterns reflect different ordered stages of post-nasal /ɡ/-deletion will be addressed in Chapter 6, where I provide an analysis in the context of the life cycle of phonological processes.
Chapter 3

The North of England

Given that this thesis is concerned with velar nasals in ‘Northern English’ in particular, some definition is required of what constitutes ‘the North’. There is a great deal of linguistic interest in northern varieties of British English, highlighted most prominently in the collection of work in Hickey (2015a); despite this, there is no widely-agreed definition of what – or rather where – the North is and although the exact placement of regional boundaries is to a large extent quite arbitrary, this question can be approached from geographic, perceptual, and dialectological angles. In the following sub-sections, I discuss the geographic North with respect to administrative regions and county borders, and also provide an overview of the linguistic features that differentiate varieties within the North as well as those shared among them to form a ‘supra-regional’ Northern English.

3.1 The geographic north

Defined purely in terms of geography, the North uncontroversially begins at the official border between England and Scotland; however, it is relatively more difficult to pinpoint where the North ends (i.e. its southern border), and this remains a debated topic. The ‘official’ region of the North of England is made up of three sub-regions, defined by the highest level of the Nomenclature of Territorial Units for Statistics: the North East, the North West, and Yorkshire and the Humber. These divisions are illustrated in Figure 3.1. Historically, this part of the country was also referred to as the ‘seven county’ North, made up of Cumberland, Cheshire, Durham, Lancashire, Northumberland, Westmorland, and Yorkshire.

This thesis is primarily concerned with the North West in particular, given that the occurrence of post-nasal [g] is restricted to this western portion of Northern England. The North West can be further subdivided into the five counties of Cheshire, Cumbria, Greater Manchester, Lancashire, and Merseyside, and had a population of just over 7 million in the 2011 census (Office for National Statistics 2012), making it the most
populous of the three northern regions (cf. approximately 5.3 million for Yorkshire and the Humber and 2.6 million for the North East\footnote{Even when the size of the geographic area is taken into account, the North West ranks highest of the three regions with respect to population density: 500 people per km$^2$, cf. 340 people per km$^2$ for Yorkshire and the Humber and 300 people per km$^2$ for the North East}).

The North West is largely dominated by rural areas, although it also features two major urban centres around Liverpool and Manchester; the latter of these has a long and storied industrial history dating back to the turn of the 19th century and its role in the industrial revolution as a centre of textile – specifically cotton – manufacturing. A large proportion of the speakers recorded and analysed for this research were in fact born and raised in the satellite towns of Greater Manchester, a metropolitan county comprising the boroughs of Bolton, Bury, Oldham, Rochdale, Stockport, Tameside, Trafford, and Wigan, as well as the cities of Manchester and Salford. Greater Manchester is relatively new, having only been established in 1974, but is alone home to over 2.5 million people. Although the linguistic behaviour of these northern communities will be discussed in more detail in Section 3.2.1, it is important to establish at
this point that Greater Manchester is a dialectally-diverse region; despite the popularity of dialect levelling as an explanatory force in language change – the effects of which have been reported both here (Barra 2006) and elsewhere in England (Watt & Milroy 1999; Williams & Kerswill 1999; Watt 2002; Dyer 2002) – the accent of central Manchester (‘Mancunian’) is dramatically different from those spoken in the likes of Wigan and Bolton etc².

### 3.2 The linguistic north

As discussed by Hickey (2015b: 1), there exists a ‘holistic’ approach to Northern English that emphasises the features shared among all the inhabitants of this wide geographic area and highlights their absence among southern varieties of British English. Equally, however, there is a great deal of variation within the North itself; this will be addressed in more detail in Section 3.2.1.

The use of the term ‘Northern English’ implies that, in spite of these differences, there are enough shared features among them to posit a single identifiable pan-regional variety of English, and this is indeed the case. Arguably the two most prominent linguistic features of northern varieties – and certainly the most salient – are the lack of distinctions between the vowels of the FOOT–STRUT and TRAP–BATH pairs of lexical sets (Wells 1982b; Wales 2006), such that the former are both produced with a rounded /ʊ/ vowel (rather than the southern /ʌ/ in strut) and the latter as short /a/ (rather than the southern long /ɑː/ in bath). That is, throughout the North, words such as bath, laugh and grass are all realised with the same vowel that appears in words such as trap, cat and pack, and likewise words such as foot and cut are perfect rhymes. It should be noted, however, that these two features are not necessarily equal with respect to their salience and the extent to which they elicit metalinguistic commentary, as the quote below indicates:

> “There are many educated northerners who would not be caught dead doing something so vulgar as to pronounce strut words with [ʊ], but who would feel it to be a denial of their identity as northerners to say bath words with anything other than short [a].” (Wells 1982b: 354)

Furthermore, Trudgill (1986) notes how although Southerners comment on Northerners’ bath and strut vowels, Northerners supposedly only stereotype Southerners as

²The M60 ring road has been employed as a geographic ‘cut-off’ point for the Manchester dialect area and speech community (Baranowski & Turton 2015; Baranowski 2017).
having /ɑː/ in bath, with comparatively little overt comment on /ʌ/. This commentary is even reflected in media representations of northern dialects, with phonetically-motivated orthography such as grim oop north to indicate a rounded [ʊ] vowel in up (see Wales 2000: 10; Montgomery 2007: 118).

Based on these two vocalic features, how then would we delimit the North linguistically? Wells (1982b) proposes an isogloss running from the River Severn in the West to The Wash in the East, which would of course include much of the Midlands. Alternatively, Chambers & Trudgill (1998) propose different isoglosses for trap–bath and foot–strut: both run from The Wash in the East of the country but the latter runs lower than the former and they end at different latitudes on the West side of the country, suggesting that speakers in the southern Midlands might pattern with the North with respect to one of these features but with the South for the other. It has also been claimed that many speakers in this transitional zone have ‘fudged’ or ‘scrambled’ lecxts with intermediate variants for strut such as [ɤ] (Upton 1996; Chambers & Trudgill 1998: 114, Britain 2001: 230–231). These various ways of delimiting the southern border of the ‘linguistic North of England’ are illustrated in Figure 3.2.

Although northern varieties are perhaps most strongly characterised by the short /a/ vowel in bath and the rounded /ɔ/ in strut, there are a number of other consistent features that contribute to a holistic view of Northern English. The vowels in the goat and face lexical sets tend to be monophthongal /oː/ and /eː/ respectively, particularly in rural Lancashire and Yorkshire (Beal 2004: 123; Ferragne & Pellegrino 2010). Some of these more rural areas also exhibit pockets of rhoticity (Wells 1982b: 368; Barras 2018), and the positional allophony in /l/-velarisation that is present in most varieties of English is not as apparent in the North: in Tyneside /l/ is realised as a light variant in almost all environments and in large parts of Lancashire /l/ is consistently velarised and surfaces as the dark [l] variant even in syllable onsets (Kelly & Local 1986; Watt & Milroy 1999; Carter 2002; Cruttenden 2014; Turton 2014, 2017).

### 3.2.1 Dialectology in the North of England

There is a long history of dialectology in the north of England; not only has there been extensive coverage of northern varieties in general dialectological handbooks (e.g. Wakelin 1977; Wells 1982b; Beal 2004; Hughes et al. 2012), contemporary variationist linguistics has often placed the lens of inquiry on northern dialects such as those spoken in Manchester (e.g. Drummond 2012; Baranowski & Turton 2015; Baranowski 2017), Liverpool (e.g. Honeybone 2007; Watson 2007; Cardoso 2015), Tyneside (e.g. Milroy et al. 1994; Watt 2002; Beal et al. 2012), and Yorkshire (e.g. Petyt 1985; Tagliamonte 2004; Tagliamonte & Temple 2005).
Figure 3.2: Map of the United Kingdom, indicating the different ways of delimiting the southern border of the North of England: (a) foot–strut boundary, (b) Trudgill (1999), (c) bath–trap boundary, (d) Severn-Wash line (Wells 1982b). Adapted from Figure 1 in Montgomery (2015: 349).
Northern English can clearly be considered a supra-regional variety, given the afore-mentioned features consistent throughout this wide geographic area. However, at the same time it should be made clear that there is a great deal of diversity within the North itself, which has not gone unnoticed as the following claim indicates:

“Local differences in dialect and accent as one moves from valley to valley or from village to village are sharper in the north than in any other part of England, and become sharper the further north one goes.” (Wells 1982b: 351)

This is perhaps no more evident than in the differences between Liverpool and Manchester. Of the two major urban centres in the North West of England, Watson (2006) notes that Liverpool English is significantly more well-studied than Manchester English (or any other variety spoken in Lancashire), but this is no longer the case in light of subsequent work in the years following this (see for example Baranowski & Turton 2015; Baranowski 2017; Drummond 2018). Despite the close proximity of these two urban centres – which are approximately just 30 miles apart – these two northern varieties are markedly different.

Manchester English is characterised by a number of vocalic features such as super-lax [ɛ] in the HAPPY lexical set (Ramsammy & Turton 2012) – a feature so salient that it plays a central role in the enregisterment of Manchester English and is often reflected orthographically (e.g. citeh, in reference to Manchester City Football Club). Speakers of Manchester English also often have a fronted /aː/ vowel in the START and PALM lexical sets (Hughes et al. 2012: 117).

Liverpool English, on the other hand, is most often stereotyped by its spirantisation of /k/ to [x] and what is referred to as the ‘T-to-R’ rule that sees word-final /t/ replaced by [ɾ] when followed by vowel-initial words in certain highly-frequent collocations, such as get off [ɡɛɹɒf] and shut up [ʃʊɹʊp] (Clark & Watson 2011). Speakers of Liverpool English also merge the vowels of nurse and square, most often to [ɛː]; this departs from the way this merger operates in most of Lancashire, where the two are merged to [ɜː] instead (Watson 2006; Watson & Clark 2013).

Despite these differences, Liverpudlian and Mancunian varieties of English – possibly by virtue of their highly urbanised status – stand out from almost all other dialects spoken in the North West in not retaining the traditional monophthongal realisation of FACE [eː] and GOAT [ɔː]. This is characteristic of the more rural northern areas, of Lancashire towns such as Blackburn and Accrington, and of the satellite towns surrounding Manchester itself such as Bolton and Wigan (Beal 2004: 123; Ferragne & Pellegrino 2010: 17).
Despite this dialectal diversity within the North West, all speakers within this region are united in their use of post-nasal [g]. Although this lack of ng-coalescence has not been particularly well-studied, there is no evidence to suggest that its behaviour varies between communities within this region; the isogloss from the Survey of English Dialects covers almost all of this region as well as the northern part of the West Midlands, and this feature has already been attested specifically in the likes of Manchester (Schleef et al. 2015), Cheshire (Watts 2005), Liverpool (Knowles 1973), and West Wirral (Newbrook 1999).

3.2.2 The North West as a ‘speech community’

The participants involved in the studies that constitute this thesis, detailed later in Chapter 5, were all born and raised in the North West of England. Throughout this body of work, and particularly in Paper III, the North West is treated as a single ‘community’ with respect to this specific variable, which is common to all the dialects spoken in this region. However, the size of this broad geographic area raises the question as to whether or not the North West can be treated as a coherent speech community, although this in part depends on the definition of speech community which, as highlighted by Patrick (2008), is not widely-agreed upon.

It is clear that the North West, being a large and diverse region, cannot be viewed as a speech community in the same way as, for example, Manchester can. Nevertheless, it is the case that the notion of community is complex and refers to more than just geographic proximity. Indeed, this has been suggested on numerous occasions, with Eckert (2004: 108) highlighting how “place is as much ideological as it is physical” (see Hill & Williams 1996 for a similar view). I argue that different ‘layers’ of community co-exist alongside each other in a nested, hierarchical manner, and that this can be justified with reference not only to linguistic variation but also to political and/or cultural ideologies, and sociohistorical development (see Santa Ana & Parodi 1998 for a similarly nested approach to speech communities in Mexico).

There is undoubtedly a long-standing political and ideological division between the North and South of England, discussed further in this chapter in Section 3.3. However, these sociocultural divisions also exist within the North itself, e.g. the rivalries between Lancashire and Yorkshire (often referred to as the ‘Roses rivalry’), between Liverpool and Manchester (dating back to the construction of the Manchester ship canal), and likewise in the North East between Tyneside, Teesside and Wearside (Llamas 2001; Burbano-Elizondo 2008; see also Beal et al. 2012: §1.4.3 on local rivalries within the North East).

This complex, multi-layered approach to community and identity is also sup-
ported by linguistic evidence. Certain features may distinguish particular towns and cities within the North West, such as /k/-frication in Liverpool, HAPPY-laxing in Manchester, monophthongisation of goat and face in more rural areas of Lancashire etc. Other features may project broader divisions within the North. Take, for example, definite article reduction (e.g. going t’shops), which is characteristic of the dialects spoken around Lancashire and Yorkshire but is comparatively absent in the North East (Jones 2002). Furthermore, while there is a great deal of diversity within the dialects of the North East – Kerswill & Williams (2002a) present evidence from a dialect recognition task that suggests Middlesbrough English is closer to the variety spoken in Liverpool than in Newcastle – there are still a number of features consistent across this whole region, such as the glottalisation of voiceless stops, or the clear realisation of final /l/s (Beal et al. 2012). Moving up one level further, we can look to more pan-northern features that unite all of these speakers: most notably, they largely lack a distinction between the vowels of foot and strut, and of trap and bath.

What this suggests is that it is possible to justify various distinctions depending on the granularity at which you view the evidence, whether that be linguistic, sociocultural, political, or otherwise. As part of this approach to community, it is important to highlight how speakers can foreground different ‘layers’ of this nested community structure depending on the interactional situation and what ideologies it motivates. That is, speakers can invoke different identities depending on contextual factors: one particular speaker may identify strongly as a Northerner in one situation, but more specifically as a Mancunian in another interactional situation. This is summarised eloquently by Hill & Williams (1996):

“One person’s awareness of being a Northerner is not necessarily the same as another’s and, indeed, the same individual can subscribe to differing forms of Northern identity at different times.” (Hill & Williams 1996: 6)

This discussion becomes particularly pertinent in Paper III, in which the social meaning of post-nasal [ɡ] is related to wider conceptualisations of the speech community and the role of evaluative norms.

3.2.3 Northern conservatism

The North of England has a reputation for being relatively conservative with respect to language change. This is a topic that will be addressed in numerous points throughout this body of work, but at this point it is important to note that many of the features
discussed in the preceding sections – those said to characterise Northern English and the individual varieties spoken across the North – are not examples of innovations that have occurred only in the North of the country. Rather, they are examples of the North *resisting* innovations that have otherwise gained ground elsewhere in England, and instead maintaining an earlier, more traditional variant.

As pointed out by Wells (1997) in an exploration of this northern conservatism, the most salient characteristics of northern accents – the lack of vocalic differentiation between the TRAP–BATH and FOOT–STRUT pairs of lexical sets – are perfect examples of this. Speakers in the North did not merge these vowels; rather, the oppositions between /a/~/ɑː/ and /ɪ/~/ʌ/ represent innovations, i.e. vowel splits, that occurred in the South but were resisted in the North. The monophthongal FACE [eː] and GOAT [oː] vowels are further examples of this, reflecting more traditional realisations before these vowels underwent ‘Long Mid Diphthonging’ during the Great Vowel Shift around 1800 (Dobson 1968; Wells 1982a: 210–211; Prichard 2014: 94). On a related note, many northern varieties even retain long [uː] in words with <oo>, such as *book* [buːk] and *look* [luːk], which in all other varieties changed to short [ʊ] around the 17th century (Wells 1982a: 133; Beal 2004: 122).

The conservative nature of northern dialects is also seen in consonantal variation, such as the afore-mentioned pockets of rhoticity that remain across Lancashire, and is not just restricted to phonology either: the retention of archaic second person pronouns such as *thee* and *thou* in large parts of the North, particularly Yorkshire, points to a similar conservatism in the domain of morphosyntax (Hughes et al. 2012: 34; Buchstaller & Corrigan 2015: 85).

Of course, the most pertinent example of the North’s conservative status with respect to language change has already been discussed; post-nasal [g]-presence itself is a perfect illustration of this, as its presence in the North West of England reflects the failure of /g/-deletion to reach completion here as it did in almost all other varieties spoken throughout the country and indeed the English-speaking world (Bermúdez-Otero 2011; Bermúdez-Otero & Trousdale 2012).

---

3A further example of linguistic conservatism in the FACE lexical set can be seen in how some northern speakers still maintain the historical distinction between words like *eight* and *weight* with [eɪ], and words like *ate* and *wait* with [eː] (see Hughes et al. 2012: 105 and Petyt 1985: 119–124 on West Yorkshire varieties). Although Beal (2004: 123) claims that such a distinction is ‘recessive’ in these northern dialects, it is evident for a number of speakers from Blackburn recorded in this research, even as young as 20 years old.
Perceptions of the North

Thus far, the discussion of the North–South divide in England has foregrounded fairly objective measures based primarily on dialectological variation. However, the issue of determining where the North ends and the South begins is as much a question of perception and ideology as it is a question of physical geography or political and administrative boundaries.

Perceptions of the North are particularly interesting given the stereotypes and attitudes surrounding this salient North–South divide in England (see Wales 2000 for a thorough discussion of this historical divide). Montgomery (2015) draws attention to the perpetuation of the ‘depressed North’ and ‘prosperous South’ stereotypes in the decades following the Second World War, when dramatic decline of the North’s industry resulted in mass unemployment.

Attitudes towards the North are also rooted in linguistic variation and dialectology. The survey conducted by Coupland (2007) indicates that some of the most stigmatised regional accents – or at least those evaluated as having the lowest degree of social prestige – are those varieties spoken in the North of England (see Figure 3.3).

Based on the salient nature of this North-South divide in public discourse (Jewell...
1994; Billinge & Baker 2004), it is not surprising to find that non-linguists have strong opinions on where exactly ‘the North’ ends and ‘the South’ begins. Montgomery (2015) explores this quantitatively under a perceptual dialectology framework (Preston 1989), specifically with the use of the ‘draw-a-map’ task in which respondents are asked to draw lines on blank maps to denote what they believe to be the major dialectal divisions, and crucially to pinpoint where exactly those divisions are. Additionally, subjects are often asked to annotate with descriptive labels and provide some commentary to the regions they identify in order to elicit attitudinal data. These individual maps are subsequently aggregated, allowing for generalisations to be made based on common patterns.

Firstly, it is interesting to note that only 9% of respondents in Montgomery’s (2015) task opted for a tripartite division between the North, Midlands, and South; the fact that a majority of subjects (68%) adopted a two-way ‘North–South’ division highlights the salience of this binary opposition.

There is also a tendency for subjects to agree more on the eastern point of this North-South divide relative to the western point; this forms a neat parallel with the isoglosses and boundaries plotted earlier in Figure 3.2 based on regional patterns of linguistic variation, which all begin at the same easterly point (the Wash) but end at dramatically different latitudes at the western border of England.

However, arguably the most interesting result found by Montgomery (2015) is the role that proximity plays in determining where the North–South divide is placed. Subjects from more southerly locations are more likely to agree on the location of the dividing line, and crucially tend to place it at a more southerly latitude. Montgomery refers to this as ideologically-motivated ‘shifting’, suggesting that it reflects a desire to guarantee one’s ‘northerness’. This seems perfectly plausible given the kind of social commentary and attitudes elicited not only in this task but in similar perceptual studies (e.g. Montgomery 2007; Beal 2009), which reveal perceptions of northerners being ‘honest’, ‘friendly’, and ‘trustworthy’, and of southerners being ‘posh’, ‘up themselves’, and ‘aloof’ (Montgomery 2015: 363)⁴. It is also suggested that the stronger agreement among more southerly subjects stems from the divide’s increased cultural salience for these people, who live much closer to the contested transitional zone and therefore have ‘more to lose’ (or gain) from its exact position (Montgomery 2012, 2015).

⁴This highlights an inverse correlation between social prestige (as discussed earlier based on Coupland 2007) or linguistic capital, and the kind of personal qualities that are seen to be desirable. That is, although people living in this ‘transitional zone’ are undoubtedly aware that northern accents are still subject to linguistic prejudice, this is overridden by a desire to identify as northern and therefore be associated with the positive qualities indexed by this status.
Chapter 4

Theoretical issues

One of the major goals of this dissertation is to investigate the extent to which synchronic variation in (ng) aligns with predictions made by the life cycle of phonological processes; in doing so, the results should lend further empirical insight into a wider and ongoing theoretical debate regarding the nature of phonological representations.

In Section 4.1, I outline the tension between two major competing theories of grammatical architecture and phonological representation: on the one hand, modular architectures with abstract underlying representations, and on the other, episodic models in which the lexicon contains fine phonetic detail. In Section 4.1.1, I discuss how these theories make different predictions concerning the patterns of synchronic variation and pathways of change that should be possible.

Following on from this, the discussion in Section 4.2 concerns a different theoretical issue: namely, the role of social meaning in explanations of sound change. The degree to which social meaning play a role in the incrementation of change is still a contested topic in contemporary variationist linguistics (see e.g. Eckert & Labov 2017; Bermúdez-Otero forthcoming) and, as mentioned in Section 1.1, one of the goals of this dissertation is to investigate the indexicality of [ŋɡ] and how this might influence its diachronic and synchronic variation.

4.1 Theories of phonological representation

One of the major theoretical debates in phonetics and phonology concerns the nature of phonological representation, the level of abstraction in the phonological system, and more generally the architecture of grammar itself (see e.g. Hyman 1970; Kiparsky 1973; an overview is given in Pierrehumbert 2016). This debate is dominated by the tension between, on the one hand, advocates of generative approaches to phonology and, on the other, those working under usage-based frameworks; this in part reflects a more long-standing division between formal and functional approaches to linguis-
tic theory (see e.g. Croft 1995; Newmeyer 1998). This discussion provides a crucial background to the analysis of (ng) later in this thesis, as these competing accounts make sharply different predictions regarding the nature of phonological and phonetic variation and change, on which the findings of this thesis have a direct bearing.

Under a generative approach to phonology (e.g. Chomsky & Halle 1968), the grammar consists of abstract underlying representations in the lexicon, from which surface representations are derived through application of phonological and phonetic rules. However, in usage-based accounts of functionalist phonology (e.g. Bybee 1998, 2002), the lexicon plays a much more central role: rather than assume abstract, feature-based representations in the lexicon, under this phonological framework words are stored as memory traces with fine phonetic detail, and production takes place by aggregating over clouds of these stored exemplars, giving rise to possible word-specific phonetic behaviour.

Of particular relevance to this dissertation is what is commonly referred to as a ‘modular, feed-forward’ architecture of grammar (Pierrehumbert 2002; Bermúdez-Otero 2007), which at its most general level of representation can be illustrated diagrammatically as in Figure 4.1. Under this architecture, lexical representations are stored, and phonological and phonetic computation take place, in separate ‘modules’, with a single interface between adjacent modules. Lexical representations are stored with discrete, abstract features and are subject to categorical phonological rules to produce a surface phonological representation. This is then fed into rules of gradient phonetic implementation, which assign continuous values in phonetic space to abstract phonological categories in order to produce the acoustic and articulatory output. Bermúdez-Otero (2007) highlights two crucial properties of this architecture:

1. both lexical and phonological representations contain only discrete values.
2. due to the modular feedforward nature of this architecture – where information travels unidirectionally between modules that share an interface – rules of phonetic implementation cannot apply directly to lexical representations.

What this means is that gradient phonetic rules can apply only to surface phonological representations (by virtue of the phonetics–phonology interface), and should therefore be blind to lexical identity due to the lack of interface between the lexicon and the phonetic module.

A notable version of this modular architecture is the stratal phonological system described in Bermúdez-Otero (2007), in which the phonological module is further separated into stem-, word-, and phrase-level processes and representations. This is
illustrated in Figure 4.2. These domains correspond to different stages of the phonological derivation, and as such are differentiated by the size of the constituent in the phonological representation: stem-level rules apply to stems (i.e. they take place before affixation), word-level rules apply to entire words (i.e. after affixation), and finally phrase-level rules are sensitive to phrasal content (i.e. they can ‘see’ across word boundaries).\footnote{This aspect of the modular architecture bears some resemblance to the earlier theory of Lexical Phonology, a framework that makes similar distinctions between phonological processes that apply lexically (interleaved with morphological processes) and post-lexically (after syntactic operations) (Kiparsky 1982a; Mohanan 1982, 1986; see Rubach 2008 for an overview).}

Phonological processes take place within these ‘levels’, and during the derivation of a word the phonological computation proceeds outwards from the more embedded domains. In other words, processes apply at the stem level, which produces the input to the word level, which in turn feeds the phrase level computation. A direct consequence of this architecture – and in particular this model of the morphosyntax-phonology interface – is that the phonological computation can lead to opaque surface forms due to morphologically-induced misapplication. Because of the manner in which computation of the part precedes computation of the whole, derived forms can exhibit apparent opacity on the surface resulting from purely transparent application or non-application in an earlier cyclic round. Bermúdez-Otero (2011) highlights two complementary examples, one of which bears directly on this thesis: post-nasal

**Figure 4.1:** Modular feedforward architecture of grammar. Adapted from Bermúdez-Otero (2007: 502).
/ɡ/-deletion, and Belfast dentalisation.

Post-nasal /ɡ/-deletion appears to overapply on the surface in words like singer, where the regular rules of English syllabification would place /ɡ/ in the onset of the second syllable and thus be ineligible for coda-targeting deletion. However, deletion has already taken place in an earlier cycle – specifically to the stem sing, before word-level suffixation of -er takes place – and as such we find a counter-bleeding interaction with opaque application of /ɡ/-deletion on the surface.

Conversely, dentalisation in Belfast English – a process whereby coronal /t, d, n, l/ take on a dental place of articulation when immediately followed by a rhotic segment – exhibits opacity of a different kind: while post-nasal /ɡ/-deletion exhibits morphologically-induced overapplication, Belfast dentalisation seemingly underapplies in certain conditions. When the rhotic segment is present in the stem-level representation, i.e. in monomorphic or root-based words such as Peter [piːtəɹ] or ladder [laːdəɹ], we find transparent dentalisation. However, when the rhotic segment belongs to a suffix attaching to a free stem, e.g. heat-er [hitəɹ], it does not trigger dentalisation. This suggests that pre-rhotic dentalisation is confined to the stem level, and that word-level suffixation of -er in this case counter-feeds the dentalisation process.

These cases of opacity fall out naturally from the architecture of grammar in the-
ories of stratal phonology, leading to the following two predictions:

• **Russian Doll Theorem**: “Let there be the nested cyclic domains \([γ \ldots β \ldots α \ldots]\). If a phonological process \(P\) is opaque in \(β\) because its domain is \(α\), then \(P\) is opaque in \(γ\)” (Bermúdez-Otero 2011: 2023).

• **Cyclic containment**: “In cases of morphosyntactically induced phonological opacity, a linguistic expression inherits its opaque phonological properties from a constituent defining an immediate cyclic subdomain” (Bermúdez-Otero 2018: 103).

Other approaches to the morphosyntax-phonology interface, such as those that rely not on cyclicity in the derivation of words but on surface-to-surface computation, are not bound by the same constraints. Output-output correspondence (hereafter OOC) is one such account, which assumes that transparent application in the ’surface base’ directly influences the surface representation of a morphosyntactically related form (Benua 1997; Kager 1999). For example, the over-application of /g/-deletion in *singer* might instead reflect a high-ranked OO-identity constraint that would otherwise be violated if there was a mismatch between the morphologically related forms *sing* and *singer*. Likewise, the explanation for Belfast dentalisation would be that transparent non-application in base form *heat* is transmitted over to the derived form *heater* through transderivational correspondence, even though this creates an opaque surface representation where the [t] remains alveolar before [əɹ]. Such theories find empirical support from apparent cases in which the afore-mentioned cyclic containment prediction is violated, i.e. phonological paradigmatic dependencies between pairs of words in which neither is cyclically contained within the other (Steriade 2008, 2012; Steriade & Yanovich 2015).

### 4.1.1 Pathways of change

The competition between theories of phonological architecture outlined in the preceding section has implications not only for synchronic linguistic behaviour, but also for predicted pathways of diachronic change. That is, the theories impose constraints on what kinds of changes are possible in terms of their lexical and phonetic implementation as well as their interaction with morphosyntactic factors.

One major way in which their predictions diverge concerns the lexical and phonetic implementation of change. Within a modular architecture, this change can be phonetically abrupt if it involves an innovation in the phonological module, affecting

---

2For a recent summary of the competition between these two theories, see Bermúdez-Otero (2018).
discrete features\(^3\). This change would also be lexically abrupt, meaning that it affects all relevant words simultaneously and in equal measure. Alternatively, a change can also be phonetically abrupt if it involves feature substitution in the underlying representation of individual words, in which case it would be lexically diffusing, spreading through the lexicon one word at a time (Wang 1969).

Under this theoretical framework, a change can only be phonetically gradual if it involves an innovation in the phonetic module, i.e. a rule of phonetic implementation, which assigns values in continuous phonetic space. Crucially, the lack of interface between the phonetics and the lexicon prevents rules of phonetic implementation from making direct reference to lexical identity, and as such innovations of this kind surface as regular neogrammarian change (see Bermúdez-Otero 2007 for a discussion of the inventory of possible sound changes).

A direct consequence of this modularity is that there can be no change that is simultaneously gradual with respect to both phonetic and lexical implementation. Diachronic change of this kind is instead possible under Exemplar Theoretic accounts of phonological representation, in which phonological features are absent (at least in purely episodic approaches) and word-specific effects can be registered through changes to fine phonetic detail of individual words (Goldinger 1998; Hawkins 2003; Wade & Möbius 2010). This is purely mechanical in nature, resulting from the production–perception feedback loop: the phonetically-rich lexicon is constantly updated with new exemplars that a listener experiences, which then influences the next production of that word as speakers average over clouds of stored tokens in speech production (Sóskuthy & Hay 2017). A widely-cited line of evidence for this is schwa reduction in English, which is sensitive to lexical frequency such that high frequency words like *memory* are much more likely to surface with a reduced/deleted vowel relative to low frequency words like *mammary* (Bybee 1998). As Bermúdez-Otero (2007) explains, the presence of such synchronic frequency effects is still permissible in an expanded version of the modular architecture, in which surface phonological representations are encoded with information about lexical accessibility in order to facilitate hypo- and hyper-articulation (Lindblom 1990). However, episodic accounts of phonological representation do find further support from a wider range of sources (e.g. Walker & Hay 2011; Hay et al. 2015; Raymond et al. 2016; Hay & Foulkes 2016; Sóskuthy & Hay 2017; Sóskuthy et al. 2018).

It should be noted that theories of phonological representation that rely solely on phonetically-detailed exemplars fail to account for many phenomena that pose no

\(^3\)Phonetic abruptness, or ‘categoricity’, typically manifests itself as a discontinuity in quantitative space; evidence of bimodality along some phonetic dimension is typically viewed as a robust indicator of phonological – rather than phonetic – change (Scobbie 2005; Bermúdez-Otero & Trousdale 2012).
such problem to traditional generative approaches, such as the existence of regular neogrammian change, or the central role that type frequency (rather than token frequency) plays in phonotactic learning (see Pierrehumbert 2016: §2 for an overview of these issues). This has resulted in a move towards hybrid models of grammatical architecture, in which phonetically-rich cognitive representations exist alongside abstract categories (e.g. Pierrehumbert 2002, 2016; Nguyen et al. 2009; German et al. 2013; Docherty & Foulkes 2014; Ernestus 2014). In other words, these hybrid models are characterised by the association of phonetic detail with both words and abstract phonemic units.

Given that both theories are seemingly supported with evidence from a number of quantitative studies, this remains an open debate in phonological theory. This dissertation will lend further empirical insight into the debate by testing predictions made by the modular architecture – and in particular the theory of the life cycle of phonological processes (Bermúdez-Otero & Trousdale 2012; Bermúdez-Otero 2013) – about how (ng) variation should pattern synchronically.

The theory of the life cycle lays down possible trajectories of diachronic change, and provides an explanatory account for how these could mould synchronic patterns of variation. In addition to imposing constraints on the phonetic and lexical implementation of sound change, the life cycle of phonological processes and the modular architecture of grammar make a further crucial prediction, which is particularly pertinent to this dissertation: namely, how and when morphosyntactic factors come to affect the course of sound change.

As just discussed, the modular architecture of grammar stipulates that the phonetics does not share an interface with the lexicon, thus forbidding phonetically gradual change from registering word-specific effects. On a related note, the phonetics can also not interface with the morphology, meaning that there should be no evidence of morphologically-conditioned phonetic change. If we recall the stratal phonological architecture discussed earlier (and illustrated in Figure 4.2), the life cycle stipulates that changes entering the grammar from below – that is, as purely epiphenomenal effects that later become under cognitive control through phonologisation (Hyman 1976; Ohala 1981) – should begin as gradient phonetic effects before undergoing stabilisation as a categorical phonological rule. It is only at this point that a process can become sensitive to morphological structure.

Upon stabilisation, the new phonological avatar of this process begins applying in the widest morphosyntactic domain, i.e. as phrase-level, post-lexical rule. Over time, it undergoes progressive rounds of domain narrowing, progressing into the more en-
bedded word-level and stem-level strata. In other words, the directionality is the reverse of what was proposed earlier in the discussion of phonological cyclicity: diachronically, sound change progresses from the phonetics, into the phonology, and through an ordered sequence of phrase-, word-, and stem-level domains, whereas in synchronic speech production underlying representations pass through stem-, word-, then phrase-level phonology before finally undergoing rules of phonetic implementation.

The prediction that morphosyntactic factors only play a role later on in the life cycle of sound change, and that by consequence the initial stages of change should be blind to morphology, has been contested by apparent evidence of morphologically-conditioned gradient phonetic change (e.g. Bybee 2017; Strycharczuk & Scobbie 2016, 2017), and phonetic paradigm uniformity effects (e.g. Steriade 2000). However, it has been shown in other lines of work that this data is quite often amenable to an analysis in line with a modular architecture and regular patterns of sound change (Hill 2014; Morrison 2018).

The choice of (ng) as a testing site for the life cycle’s synchronic predictions is particularly pertinent: this variable has been employed as a prototypical example of the diachronic aspect of the life cycle, namely the process of domain narrowing, but the synchronic behaviour of post-nasal /ɡ/-deletion remains under-described.

Drawing on historical evidence from eighteenth-century orthoepist James Elphinston, Bermúdez-Otero & Trousdale (2012) reconstruct these different stages of domain narrowing for the process of post-nasal /ɡ/-deletion, highlighting its progression along the exact diachronic trajectory predicted by the life cycle during the Late Modern English period. However, we know comparatively little about the status of (ng) variation in contemporary dialects of Northern English, and consequently to what extent this synchronic variation reflects the proposed pathway of change. In exploring this relationship between the synchrony and diachrony of Northern English (ng), this dissertation is in a position to add to an existing range of sources in providing empirical support for the life cycle (see e.g. Turton 2014, 2017 on English /l/-darkening; Iosad 2016 on vowel length distinctions in Northern Romance varieties; Ramsammy 2015 on glide hardening and continuancy dissimilation in Cypriot Greek, /o/-lowering and umlaut in Swiss German, and word-final nasal velarisation in dialects of Spanish).

---

4The exact mechanisms that drive this theory of language change are described in more detail in Paper I.
4.2 Variation, sound change, and social meaning

It is of crucial importance to highlight the variable nature of these phonological and phonetic rules. Variation is a probabilistic component of the grammar that captures the fact that – within a single speaker – some of these rules may not necessarily apply in every possible instance, but rather with some variable rate of application. In many cases this probability can be partly predicted by macro-social properties of the speaker, and may be further modulated by a number of contextual and language-internal factors such that a variant appropriate for the linguistic/social/stylistic context is produced.

There have been a number of approaches to the investigation of this variation inherent in natural language, which scholars often categorise into three ‘waves’ of the variationist paradigm dating back to foundational work by Fischer (1958) in New England, Labov’s (1966) Social stratification of English in New York City study, and later research in the same vein such as Trudgill’s (1974) study of Norwich. These empirically-driven investigations are characterised by the emphasis they place on macro-social categories and how they correlate with the use of standard and non-standard forms; that is, they established orderly heterogeneity in the speech community (Weinreich et al. 1968) and uncovered broad generalisations such as the tendency of women using fewer non-standard forms, and a move away from the local vernacular as one climbs up the socioeconomic hierarchy.

However, later approaches – which are often referred to as ‘third wave’ studies – move away from these macro-social categories and instead investigate how patterns of variation correlate with categories and groups that have local importance (see e.g. Eckert 2000 on the ‘jocks’ and ‘burnouts’ of a Detroit highschool). In this approach to variationist linguistics, the social meaning of individual forms plays a crucial role: variation is not viewed as a direct reflection of social categories but rather a dynamic system of indexically-rich forms that are used in identity construction and stance-taking. Invoking Silverstein’s (2003) concept of indexical orders, the third wave of variationist study investigates how linguistic variables can become associated with particular groups, and with characteristics that may define those groups, and how speakers then use these forms agentively in acts of identity construction. Eckert (2012) provides a fitting summary:

“The emphasis on stylistic practice in the third wave places speakers not

---

5 It should be noted that while these ‘waves’ loosely reflect the chronological order in which the field has developed, there is no direct correlation: some early studies in the variationist paradigm invoke social meaning in explanations of synchronic variation, most notably by Labov (1963) on Martha’s Vineyard.
as passive and stable carriers of dialect, but as stylistic agents, tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation.” (Eckert 2012: 92–98)

The concept of ‘social meaning’, which forms a crucial component of Paper III, refers to the meanings that speakers and listeners associate with the use of particular linguistic forms. A central part of this approach to socially-meaningful variation is that these values are not static. A static approach to social meaning might predict, for example, that the alveolar -in variant (i.e. walkin’, talkin’) is evaluated as a stigmatised, non-standard form associated with ‘lazy’ or ‘uneeducated’ styles of speech, and that this has always been the case and always will be. Rather, these forms take on local social meaning, dependent on a range of contextual and interpersonal factors and are therefore fluid, open to re-interpretation and re-evaluation. This has been shown empirically in the case of the afore-mentioned (ing) variable. Campbell-Kibler (2008) finds that the same passages of speech prompts listener-specific evaluations that contradict one another: the same speaker’s use of -in is seen as more intelligent by some but less intelligent by others, compassionate by some, condescending by others. The range of these possible attitudes towards a single linguistic variant illustrates perfectly how the indexical field is, as described by Eckert (2008: 454), “a constellation of ideologically related meanings” and how it can be shaped by interpersonal factors.

For many years the emphasis on large-scale studies of the speech community overlooked the role played by social meaning in language variation and change. Although this balance has been redressed in more contemporary sociolinguistic research (see e.g. Zhang 2005; Podesva 2007; Eckert 2008; Campbell-Kibler 2009; Moore & Podesva 2009; Kiesling 2009; Pharao et al. 2014), this new focus on the local should not be at the expense of the global. It is important to understand both the correlation between macro-social categories and rates of use of certain features within the speech community, but also the role played by local instantiations of social meaning within this larger network.

Indeed, the two are closely linked, so much so that the speech community has been described as the “major social unit for the construction of [social] meaning” (Eckert 2000: 32). Indexical meaning emerges through everyday interaction both within and outwith the speech community, and while the speech community provides the context for social meaning to develop, it does not wholly determine the content of that meaning. That is, speakers may agree on what sociodemographic group a speaker belongs to, but they may hold very different views regarding that particular social group (Eckert 2000). This is at the heart of the dynamic approach to indexicality that, as discussed earlier, sees individuals settle upon very different evaluations of the same lin-
guistic phenomenon. Crucially, social meaning mediates between linguistic variants and broader social categories at the community-level, allowing speakers to employ linguistic variation as a tool to make ideological moves, take stances, and construct and project identity in stylistic practices.

This dissertation – specifically the investigation described in Paper III – pursues a number of questions relating to the local construction of this social meaning and the extent to which it can be employed as an explanatory force in language variation and change.

The mutable and complex nature of the indexical field necessitates careful design of perceptual studies seeking to uncover listener evaluations of variable forms. Simply asking listeners which variant sounds most 'correct', for example, could overlook the more fine-grained values attributed to the forms partaking in this variation. Acquiring a complete picture of the indexical field of (ng) ideally requires qualitative, ethnographic research into how it is used ‘on the ground’ in everyday interaction. The matched-guise technique used in Paper III, details of which are provided in Chapter 5, is a starting point: by selecting a number of possible indexical values – chosen based on the results of earlier, related work (e.g. Labov et al. 2006, 2011; Levon & Fox 2014; Schleef et al. 2015) – we can investigate how listeners evaluate [ɡ]-presence along these different social dimensions.

As the results of that paper will highlight, the evaluative patterns that emerge in the North West of England in relation to [ŋɡ] are not necessarily expected if [ɡ]-absence is taken to be a high-prestige variant associated with a standard RP variety. Equally, the magnitude of community-level variation uncovered in Paper III suggests that claims of the dialectal form having 'local prestige', possibly reflecting its closer alignment with the orthography (Beal 2004), are an over-simplification. Furthermore, the results call into question the role that social meaning plays in the intergenerational incrementation of sound change, adding further insight into this tension between the role of sociolinguistic evaluation against more mechanical factors such as density of communication (Bloomfield 1933; Labov 2001; Trudgill 2014).
CHAPTER 5

SOURCES OF DATA

RESEARCHING the questions outlined in Section 1.1 requires a range of different
types of linguistic data, covering the domains of both speech production and
perception. In this thesis I take a three-pronged approach to data collection,
with the four papers that constitute its results drawing from the following types of
data:

1. conversational speech from sociolinguistic interviews

2. carefully-controlled lab speech from an elicitation task

3. attitudinal data from an online perception survey

Papers I and IV utilise the naturally occurring language data in (1), while Paper II
uses the lab speech in (2) and the perception study presented in Paper III is based
on the attitudinal data in (3). These three sources of data have various strengths and
weaknesses, and complement each other in different ways. Consequently, the use of
all three provides a more well-rounded approach to the study of this sociolinguistic
phenomenon.

For example, data from sociolinguistic interviews arguably provides the most di-
rect access to the vernacular and is therefore the most reliable in identifying actual
patterns of variation. Any process of recording a speaker inevitably encounters the
‘observer’s paradox’, i.e. the need to use recording equipment to capture the acoustic
signal, but equally the need for the subject to speak naturally as though they weren’t
being recorded. However, for a number of reasons this problem is minimised in the
sociolinguistic interview relative to other forms of data collection in more experimen-
tal contexts.

The format of the sociolinguistic interview is typically described as a series of
hierarchically-structured ‘topical’ modules, covering a range of topics such as life
in the neighbourhood, childhood, family, career, travel, and hobbies and interests
(Labov 1984; Tagliamonte 2006). They are also designed to elicit ‘narratives of per-
sonal experience’, in which the speaker is said to be paying the least attention to
their speech (Labov 1984: 34; Labov 2010a). That is, one of the primary goals of these interviews is to make the participant forget they are being recorded in order to minimise the possibility of self-monitoring. This is also achieved in part simply due to the length of these interviews, which are typically recommended to last for at least an hour; as the interview progresses, the participant should become even more relaxed with the process, more accustomed to the situation, and as such any self-monitoring is reduced to the point where its effects are barely observable. In this way, the interview approximates the kind of social interaction that makes up a majority of everyday language use.

However, the downside to this method is that in sociolinguistic interviews it is of course impossible to control what the interlocutor actually says, which in turns makes this method of data collection much less efficient in terms of eliciting the words and segments of interest. In the case of (ng), this is particularly problematic as these clusters do not occur regularly in speech. Further to this, it is impossible to control the environments in which the target segment occurs; the analysis in Paper II, for example, which explores the behaviour of (ng) before prosodic boundaries of varying strengths, is only made possible because of the ability to have fine control over the stimuli in elicitation tasks. A further limitation of the sociolinguistic interview stems from the afore-mentioned need for the participant to relax and forget about the recording set-up; this makes it necessary to use a small, unobtrusive microphone, ideally out of the subject’s line of sight. In this study, a lavalier microphone was used, attached to the participant’s clothing just below and to the side of their chin.

The attitudinal survey departs from the other two methods of data collection in that it addresses language evaluation rather than production; however, it still complements interview data in that it provides access to the content of sociolinguistic evaluation, which can only be approximated indirectly through studying patterns of synchronic variation in production. Differentiation along sociodemographic lines and stylistic stratification between informal and formal discourse both provide suggestive evidence of overt sociolinguistic prestige, but these methods of determining language attitudes suffer from a number of limitations. While style-shifting between spontaneous conversation and word list tasks may sometimes reflect differences in formality and therefore pressure to avoid non-standard variants, it is confounded by prosodic differences in intonation and speech rate, and the fact that word list elicitations can be subject to performative acts of identity construction that may contradict overt norms (Coupland 1985; Schilling 2013). Furthermore, while it may be possible to infer sociolinguistic norms from differences in production between social groups, such as men and women or lower and higher social classes, this is very much rooted in
the ‘first wave’ of variationist study (see discussion in Eckert 2012) and may overlook more fine-grained indexical values associated with different variants. Consequently, a more reliable insight into the social meaning of a linguistic variable can only be attained through either ethnographic access or experimental methodologies such as the matched-guise technique (see for example Campbell-Kibler 2009, 2011b; Levon & Fox 2014).

In the following sections I discuss these sources of data in more detail, and in Section 5.4 I also provide information on the subjects who took part in these various tasks. Section 5.5 details the data analysis and statistical methods employed in these studies.

## 5.1 Sociolinguistic interviews

In total, the corpus of naturally-occurring speech used in this analysis consists of 32 dyadic sociolinguistic interviews largely conducted between 2015 and 2017. Two of these interviews were conducted in 1971 in Manchester, and are included in this corpus to provide extra time depth to the diachronic analyses described in Paper I.

Participants were recruited through existing social networks within the region. Prior to each interview, I described the process to the participant and provided them with the project information sheet and consent form to sign. Participants were told that this was a study of ‘English dialects of the North West’, which aimed to investigate certain features that are characteristic of the varieties spoken in this region. To minimise the possibility of self-monitoring during the production of these (ng) clusters, speakers were never told of the specific variable of interest and were instructed to speak in a conversational manner that feels natural for them. The task was described as an informal interview where they would be asked about their interests, their life growing up in the North of England, and about any other topic they would like to discuss, followed by two short reading tasks.

On average, the interviews lasted for 52 minutes (with the shortest being 21 minutes and the longest being 85 minutes), and in all cases the conversation was followed up with two elicitation tasks: a word list containing 8 tokens of (ing) and 18 tokens of (ng) randomly dispersed between a further 83 distractor items, and a 239-word reading passage containing a further 21 tokens of (ing) and 17 tokens of (ng). The word list and reading passage stimuli can be found in Appendix A and Appendix B of Paper IV.

In total, these interviews contain approximately 28 hours and 251,000 words of spontaneous speech, all of which was manually transcribed using ELAN (Wittenburg
et al. 2006). The output of this transcription process was then fed into the Force Alignment and Vowel Extraction (FAVE) suite (Rosenfelder et al. 2011) to produce a time-aligned TextGrid.

### 5.1.1 Forced alignment

Forced alignment is a widely-used tool in contemporary sociophonetics whereby an audio file and a matching orthographic transcription are aligned together such that the transcribed words – and even the individual segments that make up each word – line up precisely where they occur in the speech signal. This allows for efficient and large-scale sociophonetic analysis that would have otherwise been impossible with traditional methods. Labov et al. (2013: 35), for example, conduct analysis on a dataset of 3000–9000 vowel measurements per sociolinguistic interview, which they contrast with manual vowel extraction and the forty hours it previously took to extract just 300–350 measurements using manual techniques.

The FAVE suite (Rosenfelder et al. 2011) was used to align the 32 sociolinguistic interviews collected for this project. The aligner takes as input the audio file and corresponding word-level transcription from ELAN, and makes reference to a further two external objects: a pronouncing dictionary containing phonemic transcriptions of a language’s words, and a set of pre-trained acoustic models for all of the phonemes of that language. These models take the form of Hidden Markov Models (HMMs) (Ghahramani 2001; Yuan et al. 2013) trained on a corpus of oral arguments from the Supreme Court of the US (SCOTUS). Despite being trained on varieties of American English, previous work has revealed that FAVE still works extremely well on British English varieties (MacKenzie & Turton 2013). Broadly speaking, the alignment process then consists of looking up the phonemic transcriptions of each word and comparing each phoneme’s acoustic model to the speech signal from the interview in order to find the most likely segmentation. This workflow is illustrated in Figure 5.1.

As discussed earlier, the FAVE suite also makes it possible to perform automatic vowel extraction. Although this thesis is primarily concerned with the consonantal variation in (ing) and (ng), Paper IV also addresses variation in the vowel of (ing) suffixes; as such, the sociolinguistic interviews were also processed with FAVE-extract.

---

1In this case, the Carnegie Mellon University dictionary (CMUdict, v0.7) was used, which at the time of writing is available here: [http://www.speech.cs.cmu.edu/cgi-bin/cmudict](http://www.speech.cs.cmu.edu/cgi-bin/cmudict). The CMU dictionary includes over 134,000 entries, and uses the ARPAbet phonemic transcription system with additional markings for primary and secondary lexical stress.

2Other aligners are available, such as PLA (ProsodyLab-Aligner; Gorman et al. 2011), SPPAS (SPeech Phonetization Alignment and Syllabification; Bigi 2015), and MAUS (Schiel 1999; Kisler et al. 2017), which largely work in similar ways to the methods described here.
...that first concert I went to was Bowie as Ziggy...

<table>
<thead>
<tr>
<th>word</th>
<th>transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOVINE</td>
<td>B OW1 V AY2 N</td>
</tr>
<tr>
<td>BOWIE</td>
<td>B OW1 IY0</td>
</tr>
<tr>
<td>BOWING</td>
<td>B AW1 IH0 NG</td>
</tr>
<tr>
<td>BOWING</td>
<td>B OW1 IH0 NG</td>
</tr>
<tr>
<td>BOWL</td>
<td>B OW1 L</td>
</tr>
</tbody>
</table>

Figure 5.1: Illustration of the forced alignment workflow, highlighting (a) the orthographic transcription input, (b) the audio file input, (c) the pronunciation dictionary, (d) the acoustic model, and (e) the aligned TextGrid output. Acoustic model diagram in (d) adapted from Fruehwald (2014).
returning static and dynamic measurements for over 280,000 vowels (with the dynamic measurements being sampled at 5 time points throughout the vowel: 20%, 35%, 50%, 65%, and 80%). These measurements were normalised using the Lobanov (1971) method prior to the analysis, to ensure that comparisons between speakers were not confounded by physiological differences in vocal tract size and shape.

It is also possible to use FAVE-align to automatically code for the dependent variable itself, such that the aligned TextGrid it returns contains a phone-level tier with a phonetic transcription that actually reflects the variant used for each (ing) and (ng) word, rather than the word’s canonical transcription as represented in the pronouncing dictionary. All that is required is to change the dictionary to include multiple entries with the same lexical transcription but different segmental transcriptions, each reflecting a possible pronunciation variant (see Table 5.1 for an example). This dictionary expansion was carried out using a Python script written to identify words falling under the envelope of variation of either (ing) or (ng) and to produce alternative segment-level transcriptions. If the aligner encounters multiple possible dictionary entries at the look-up stage, it will select the one with the phone-level transcription that best fits the acoustic signal. Coding the dependent variable using this automated procedure may be less accurate than manual annotation from a trained phonetician, but by its very nature this method is objective, reproducible, and consistent; for example, it is not prone to fatigue-related effects that may influence a human coder. It is also, of course, faster by an order of magnitude. The accuracy of this method has already been explored for a number of variable processes such as /h/-dropping, /th/-fronting, and /td/-deletion with promising results (see Bailey 2016); it has also been previously employed for (ing) itself, albeit only for the variation between [ɪn] and [ɪŋ] (Yuan & Liberman 2011).

Although the dependent variable was coded manually for this project, based on categorical presence/absence of post-nasal [ɡ], the aligner’s annotation was also recorded in order to at least assess the accuracy for future work. Across the whole dataset, the aligner reached an accuracy rate of 85% for (ing) and 84% for (ng); a more detailed breakdown of the accuracy rates and distribution of errors is given in Figure 5.2. It is interesting to note that the errors are not equally distributed across the two variants for (ng): the aligner is better at identifying tokens realised as [ŋɡ] than those surfacing as [ŋ]. In other words, when [ɡ] is present on the surface the aligner is particularly reliable in identifying it, but the aligner is also prone to finding one when it is in fact absent; in many of these cases it appears that the aligner is misidentifying

---

3This script, along with other versions for different variable phenomena, is available at https://github.com/grbails/.
<table>
<thead>
<tr>
<th>word</th>
<th>transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>SINGER</td>
<td>S IH1 NG ER0</td>
</tr>
<tr>
<td>SINGER</td>
<td>S IH1 NG G ER0</td>
</tr>
<tr>
<td>THING</td>
<td>TH IH1 NG</td>
</tr>
<tr>
<td>THING</td>
<td>TH IH1 NG G</td>
</tr>
<tr>
<td>SORTING</td>
<td>S AO1 R T IH0 NG</td>
</tr>
<tr>
<td>SORTING</td>
<td>S AO1 R T IH0 NG G</td>
</tr>
</tbody>
</table>

Table 5.1: Example of dictionary expansion, where (ing) words are represented by three distinct phonemic transcriptions and (ng) words by two.

![Figure 5.2](image_url)

**Figure 5.2:** Comparison between manual and automated identification of the dependent variable in (ing) and (ng).

the initial consonant of the following word – or in some cases weak energy in the acoustic signal – as a post-nasal [ɡ].

## 5.2 Lab elicitations

The elicitation task was designed for the research questions in Paper II, which is primarily concerned with the behaviour of (ng) clusters before prosodic and syntactic boundaries of various strengths. While the conversational data naturally includes tokens in different environments, the use of a sentence list designed specifically for this purpose provides a more balanced sample and allows for finer-grained distinctions to be made with respect to the environments in which (ng) occurs. In total, the elicitation task contained 48 sentences with a single token of (ng) in each, equally balanced.
across six different types of prosodic and syntactic environments (in part mirroring the stimuli of Sproat & Fujimura’s 1993 study of /l/-darkening). However, the final analysis focuses only on the word-final tokens, excluding the pre-suffix condition (e.g. singer, youngster, wrongful).

The sentences appeared on a laptop screen one at a time, and subjects were asked to read the sentence aloud as naturally as possible. This consequently led to some subjects not producing the desired prosodic phrasing that the stimuli had been designed to elicit. For example, sentences in the IP-final condition – in which the /ŋɡ/ cluster should occur before an intonational phrase boundary – followed the structure of “[quoted speech],’ [person] said”, and contrary to prior expectations speakers did not always pause or produce a boundary tone directly after the quoted speech. However, this is still preferable to explicitly instructing participants to pause, which would likely lead to an even more unnatural style.

Like with the sociolinguistic interviews, these recordings were also processed using FAVE-align. However, as the duration of [ŋ] is one of the crucial factors explored in this analysis, the time-aligned TextGrids were then subject to a round of manual adjustment in Praat to ensure that the segment boundaries were maximally accurate for the velar nasal in each (ng) word. Of the 2880 segment boundaries (the start and end of [ŋ] in 1440 tokens), 1796 (62.4%) did not need any manual adjustment. Of the 1080 that did, the onset of the nasal was, on average, adjusted by 30ms (σ = 29.9), and the nasal offset was adjusted by 27.8ms (σ = 31.4), and in only one case did a segment boundary placed by the forced alignment process require a major shift (501ms). The dependent variable was again coded manually as a binary alternation between presence and absence of post-nasal [ɡ].

5.3 Attitudinal survey data

The attitudinal data was acquired using an online questionnaire hosted on LimeSurvey. This survey was designed as a ‘matched guise task’, with embedded sound files that subjects were asked to listen to and subsequently rate on a number of descriptive Likert scales ranging from (1) ‘strongly disagree’ to (7) ‘strongly agree’, with (4) ‘neither agree nor disagree’ in the middle. Originally conceived by Lambert et al. (1960), the matched guise task is widely used to uncover attitudes towards language and linguistic variation. Listeners are exposed to stretches of speech that typically differ with respect to just one particular linguistic attribute – in this case, the variable presence of post-nasal [ɡ] – and are asked to evaluate them along some attitudinal dimension. Any differences in how the guises are rated can then be attributed to the
variable in question, providing evidence of how its variants are evaluated and the extent to which they index certain sociosymbolic meanings.

It is important to note that different methods of attitude elicitation can, to some extent, result in different responses (see Garrett 2010 for an overview). Specifically, there is a sizeable body of research to suggest that attitudes can operate at different levels of consciousness, with a primary distinction between implicit and explicit attitudes (e.g. Rydell & Mcconnell 2006; Greenwald & Nosek 2009), and that these contrasting types of attitude can be investigated using complementary methodologies. The most widely-used technique to investigate implicit attitudes is the Implicit Association Test (Greenwald et al. 1998), which measures reaction times in order to ascertain the strength of association between an object of study and a particular evaluative dimension. In a recent study, McKenzie & Carrie (2018) employ this methodology for the first time in the context of British English language attitudes, finding evidence of a shift towards greater ‘tolerance’ of northern English accents.

Explicit attitudes, on the other hand, are often investigated through self-report measures – in which subjects are explicitly asked to comment on some aspect of language – or through verbal-guise or matched-guise techniques; the latter, which is also employed in Paper III of this thesis, has been widely used for exploring the evaluation of specific sociophonetic variables, such as (ing) (Labov et al. 2006, 2011) and (th)-fronting (Levon & Fox 2014) in English, or /s/-fronting in Danish (Pharao et al. 2014).

Crucially, a meta-analysis conducted by Hofmann et al. (2005) suggests that there is only a weak correlation between the results obtained from these different measures, with this inconsistency emerging through differences in the acquisition and development of implicit and explicit attitudes. The former are said to be slower to acquire, accruing over long periods of time through socialisation, and as a result are fairly stable and resistant to change (Gawronski & Bodenhausen 2006). Explicit attitudes, on the other hand, are comparatively volatile and are more susceptible to social biases. Due to these differences, it has been said that any discrepancy between implicit and explicit attitudinal results could plausibly reflect ongoing evaluative change (Gawronski et al. 2017; McKenzie & Carrie 2018).

The evaluation experiment described in Paper III employs the matched-guise technique, following the tradition of the afore-mentioned perceptual studies that also explore attitudes towards specific sociophonetic features. However, when interpreting the results it is important to remember that the chosen methodology may only offer partial insight into the evaluative landscape of Northern English (ng); future work should call upon complementary techniques in order to further explore these poten-
tial nuances.

The matched-guise survey was followed up with a number of demographic questions as well as a short self-report task in which subjects were played two isolated realisations of an (ng) word differing only in the presence of post-nasal [ɡ] and asked to identify (a) the variant they think sounds most ‘correct’ and (b) the variant they think best matches their own pronunciation. As another technique for probing explicit attitudes, it is well established that in such tasks subjects’ answers tend to align with social norms rather than their own speech habits (Labov 2001); that is, subjects typically over-report their use of standard forms and under-report the use of stigmatised or non-standard variants (see Trudgill 1972), and it therefore provides further insight into how this variable is evaluated.

5.4 Participants

In total, there were 32 participants for the sociolinguistic interviews, 30 for the controlled elicitation task, and 35 subjects for the attitudinal survey. The former two have partially overlapping population samples; that is, in some cases the subjects that took part in the sociolinguistic interview were later recruited for the elicitation task.

All participants were asked where they were born and where they had lived between the age of 4 and 13 to ensure they were truly representative of a 'North West' speaker; a more detailed history of residence was also obtained for the speakers that participated in the sociolinguistic interviews, as during the course of these conversations discussion usually turned to childhood and the local area. In almost all cases, the participants were still living in the same town/city in which they grew up. These locations are plotted in Figure 5.3, which highlights the geographic distribution of informants within the North West and how they cluster in particular around northern Greater Manchester and Blackburn. The participants of the online questionnaire used for Paper III all self-identified as coming from the North West, but more detailed information regarding their exact town/city was not acquired.

For all three tasks; the population sample covers a range of ages, with participants’ dates of birth ranging from 1907 to 1998 for the sociolinguistic interviews, 1932 to

---

4 Although the analysis in Paper IV includes all 32 sociolinguistic interviews, the results in Paper I are based on only 30 speakers of this corpus; two speakers are excluded because they produce no tokens of (ng) in one of the morphophonological environments that play a central role in the research question.

5 Subjects who were informed of the variable of interest following the original interview were not asked back for the elicitation task, to avoid the risk that their knowledge of this might influence their behaviour in any subsequent recordings.
Figure 5.3: Map of the United Kingdom, highlighting a subsection of the North West region and the locations where participants of the production tasks were born and raised. Borders and labels indicate postcode regions (BB: Blackburn; BL: Bolton; CH: Chester; CW: Crewe; FY: Blackpool; L: Liverpool; M: Manchester; OL: Oldham; PR: Preston; SK: Stockport; WA: Warrington; WN: Wigan).
1999 for the lab elicitation task, and 1944 to 1998 for the online survey. The former two sets of participants are also equally balanced with respect to speaker sex (17F 15M for the interviews, and 16F 14M for the lab elicitations). See Figure 5.4 for a full breakdown of these three population samples.

More detailed sociodemographic information was also collected for the participants of the interviews, both through direct demographic questions following the conversational part of the interview but also discussions of career history and upbringing during the interview itself. The population sample of the interviews was controlled for socioeconomic status by ensuring that all participants were upper working class, where class is operationalised using a composite measure similar to that employed by Labov (2001) in his study of Philadelphia neighbourhoods and Trudgill (1974) in his study of the Norwich speech community. Although this measure is based primarily on occupational history, reflecting traditional distinctions between blue-collar and white-collar professions (which are further subcategorised into skilled/unskilled labour and managerial positions), it also takes into account education and upbringing. For younger participants, almost all of whom are students, this classification was based primarily on their parents’ occupations.

Table 5.2 presents sociodemographic information about the participants of the interviews, listing the pseudonyms used throughout this thesis when referring to individual speakers, and detailing their gender, where they were born and raised, }

---

6See Ash (2003) for a thorough discussion of how social class has been operationalised in sociolinguistic research, and Baranowski & Turton (2018) for a comparison of different indices of social class in the Manchester speech community itself.
and three time-related components: the year in which the interview took place, the speaker’s age at this point in time, and their date of birth. Date of birth plays a crucial role in the analysis and interpretation of language change: under the apparent time construct, a speaker’s vernacular is said to reflect the state of the dialect at the time they were acquiring language, which means that correlations between date of birth and the rate of use of a linguistic feature can be used as a proxy for intergenerational language change (see Sankoff 2006 for empirical support). However, this is confounded by a speaker’s age at the time of interview, and how chronological age interacts with life stage. In other words, a twenty year-old speaker interviewed in 1995 and a forty year-old speaker interviewed in 2015 share the same date of birth of 1975, and so should reflect the same stage of intergenerational change. However, they are at different life stages at the time of interview, and given widespread reports of age-graded variation and lifespan change – changes within individual speakers after the critical period (see e.g. Bowie 2005; Sankoff & Blondeau 2007; Van Hofwegen & Wolfram 2010; Wagner 2012b; MacKenzie 2017) – it may be the case that these two hypothetical speakers behave quite differently from one another.

Teasing apart these interacting effects is particularly problematic in large multi-stage corpora consisting of interviews conducted across a long period of time, which require complex statistical techniques to overcome what Fruehwald (2017b: 4) describes as an “inherent and worrying unquantifiable uncertainty in identifying language change in progress”. Although this thesis draws upon interviews largely conducted in a narrow time frame (all but two took place between 2015–2017), the inability to tease apart date of birth and speaker age necessitates some degree of caution in the interpretation of apparent time community-level change.

5.5 Data analysis

In all three cases, the data was analysed using R (R Core Team 2017); this was further supplemented with various scripts written in Python and Praat to extract phonetic measurements and prepare the data for statistical analysis.

The number of packages used in R to analyse this data is far too large to provide an exhaustive list, but a majority of the analysis relied heavily on the tidyverse (v1.2.1) collection of packages (Wickham 2017), while lme4 (v1.1-17) (Bates et al. 2015) was used for fitting mixed-effects regression models. All of the figures in this thesis were produced in ggplot2 (v3.0.0) (Wickham 2016).

---

7 Fruehwald (2017b) also draws attention to a further confound presented by the year of interview itself, reporting on a ‘zeitgeist effect’ whereby a particular case of intragenerational lifespan change is restricted to one particular period of time.
<table>
<thead>
<tr>
<th>Speaker</th>
<th>Gender</th>
<th>Date of birth</th>
<th>Age</th>
<th>Year of recording</th>
<th>Location</th>
<th>Region*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BegleyJ</td>
<td>M</td>
<td>1954</td>
<td>17</td>
<td>1971</td>
<td>Manchester</td>
<td>M</td>
</tr>
<tr>
<td>BethS</td>
<td>F</td>
<td>1961</td>
<td>54</td>
<td>2015</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>BruceG</td>
<td>M</td>
<td>1950</td>
<td>55</td>
<td>2015</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>ChrisT</td>
<td>M</td>
<td>1988</td>
<td>27</td>
<td>2015</td>
<td>Didsbury</td>
<td>M</td>
</tr>
<tr>
<td>ConnorL</td>
<td>M</td>
<td>1979</td>
<td>36</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>DaveJ</td>
<td>M</td>
<td>1997</td>
<td>18</td>
<td>2015</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>DotV</td>
<td>F</td>
<td>1951</td>
<td>64</td>
<td>2015</td>
<td>Mill Hill</td>
<td>L</td>
</tr>
<tr>
<td>EllenB</td>
<td>F</td>
<td>1966</td>
<td>49</td>
<td>2015</td>
<td>Accrington</td>
<td>L</td>
</tr>
<tr>
<td>FeliciaD</td>
<td>F</td>
<td>1965</td>
<td>50</td>
<td>2015</td>
<td>Feniscowles</td>
<td>L</td>
</tr>
<tr>
<td>FionaB</td>
<td>F</td>
<td>1992</td>
<td>25</td>
<td>2017</td>
<td>Feniscowles</td>
<td>L</td>
</tr>
<tr>
<td>FrankE</td>
<td>M</td>
<td>1998</td>
<td>17</td>
<td>2015</td>
<td>Feniscowles</td>
<td>L</td>
</tr>
<tr>
<td>FredN</td>
<td>M</td>
<td>1955</td>
<td>62</td>
<td>2017</td>
<td>Salford</td>
<td>GM</td>
</tr>
<tr>
<td>GaryP</td>
<td>M</td>
<td>1957</td>
<td>60</td>
<td>2017</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>GloriaJ</td>
<td>F</td>
<td>1933</td>
<td>82</td>
<td>2015</td>
<td>Radcliffe</td>
<td>GM</td>
</tr>
<tr>
<td>GraceG</td>
<td>F</td>
<td>1994</td>
<td>21</td>
<td>2015</td>
<td>Moston</td>
<td>M</td>
</tr>
<tr>
<td>GrahamR</td>
<td>M</td>
<td>1932</td>
<td>83</td>
<td>2015</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>GwynethP</td>
<td>F</td>
<td>1963</td>
<td>52</td>
<td>2015</td>
<td>Accrington</td>
<td>L</td>
</tr>
<tr>
<td>HarryG</td>
<td>M</td>
<td>1976</td>
<td>39</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>HenryM</td>
<td>M</td>
<td>1954</td>
<td>61</td>
<td>2015</td>
<td>Accrington</td>
<td>L</td>
</tr>
<tr>
<td>JimmyC</td>
<td>M</td>
<td>1962</td>
<td>54</td>
<td>2016</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>JoeM</td>
<td>M</td>
<td>1938</td>
<td>77</td>
<td>2015</td>
<td>Accrington</td>
<td>L</td>
</tr>
<tr>
<td>LillyR</td>
<td>F</td>
<td>1907</td>
<td>64</td>
<td>1971</td>
<td>Manchester</td>
<td>M</td>
</tr>
<tr>
<td>MaryB</td>
<td>F</td>
<td>1959</td>
<td>57</td>
<td>2016</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>MikeM</td>
<td>M</td>
<td>1981</td>
<td>34</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>MollyF</td>
<td>F</td>
<td>1987</td>
<td>28</td>
<td>2015</td>
<td>Whitefield</td>
<td>GM</td>
</tr>
<tr>
<td>TanyaC</td>
<td>F</td>
<td>1979</td>
<td>36</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>TheaS</td>
<td>F</td>
<td>1979</td>
<td>36</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>TrudyC</td>
<td>F</td>
<td>1935</td>
<td>80</td>
<td>2015</td>
<td>Rishton</td>
<td>L</td>
</tr>
<tr>
<td>WadeT</td>
<td>M</td>
<td>1991</td>
<td>24</td>
<td>2015</td>
<td>Tockholes</td>
<td>L</td>
</tr>
<tr>
<td>WandaJ</td>
<td>F</td>
<td>1979</td>
<td>36</td>
<td>2015</td>
<td>Bury</td>
<td>GM</td>
</tr>
<tr>
<td>WendyJ</td>
<td>F</td>
<td>1993</td>
<td>22</td>
<td>2015</td>
<td>Darwen</td>
<td>L</td>
</tr>
<tr>
<td>WillowA</td>
<td>F</td>
<td>1995</td>
<td>20</td>
<td>2015</td>
<td>Darwen</td>
<td>L</td>
</tr>
</tbody>
</table>

*M = Manchester; GM = Greater Manchester; L = Lancashire.

**Table 5.2:** Sociodemographic information for the 32 participants of the sociolinguistic interviews.
ORIGINAL PUBLICATIONS
Paper I


N.B. This paper is currently undergoing peer review. When citing, please refer to the final published version once it becomes available.
Insertion and deletion in Northern English (ng): Interacting innovations in the life cycle of phonological processes

George Bailey

University of Manchester

Abstract

In north western varieties of British English the historical process of ng-coalescence that simplified nasal+stop clusters in words like wrong and singer never ran to completion, with surface variation between [ŋ] and [ŋɡ] remaining to this day. This paper presents an empirical study of this synchronic variation, specifically to test predictions made by the life cycle of phonological processes; a diachronic account of /ɡ/-deletion has been proposed under this framework, but crucially the life cycle makes hitherto-untested predictions regarding the synchronic behaviour of (ng) in North West England. Data from 30 sociolinguistic interviews indicate that these predictions are largely met: internal constraints on the variable are almost entirely accounted for by assuming cyclic application of /ɡ/-deletion across a stratified phonology. There is also evidence of apparent-time change in the pre-pausal environment, which is becoming increasingly [ɡ]-favouring contrary to the life cycle’s predictions. It is argued that this reflects a separate innovation in the life cycle of (ng), with synchronic variation reflecting two processes: the original deletion, overlaid with a prosodically-conditioned insertion process. These results have implications for theories of language change and the architecture of grammar, and add to a growing body of evidence suggesting that the effect of pause on probabilistic phenomena can be synchronically variable and diachronically unstable.

Keywords: life cycle of phonological processes; variation; sound change; phonological theory; velar nasal; dialects of English
1 Introduction

Explanation in phonology has become increasingly concerned with the integration of sound change and synchronic variation, leading to the development of theories such as Evolutionary Phonology, which seeks to explain cross-linguistic generalisations by reference to patterns of language change (Blevins 2004, 2006), and ‘amphichronic’ models such as the Life Cycle of Phonological Processes (Kiparsky 2006; Bermúdez-Otero 2013), which predicts an ordered set of synchronic grammars resulting from pathways of change.

A key component of amphichronicity is the way in which synchronic and diachronic accounts can mutually inform one another: specifically how contemporary speakers’ grammars can reflect diachronic trajectories of change and, conversely, how theories of language change make predictions regarding synchronic variation which can then be tested empirically (see e.g. Turton 2017 on /l/-darkening).

This paper is concerned with a variable process of post-nasal /ɡ/-deletion affecting words such as young [jʌŋɡ]-[jʊŋ] and wrong [ɹɔŋɡ]-[ɹɒŋ] in dialects of British English spoken in the North West and West Midlands of England, the last vestige of a process of ng-coalescence that has run to completion in all other varieties (Wells 1982b). The diachronic trajectory of this change has been discussed in great detail in expositions of the life cycle of phonological processes (e.g. Bermúdez-Otero 2011; Bermúdez-Otero & Trousdale 2012), but the predictions made by the life cycle regarding present-day patterns of variation have yet to be scrutinised empirically. Drawing upon data from a corpus of sociolinguistic interviews conducted with speakers of Northern Englishes, the goal of this paper is to investigate variation in the realisation of /ŋɡ/ clusters in the North of England with a specific focus on testing these predictions.

Results suggest that variation in [ɡ]-presence is predicted most strongly by its morphophonological sensitivity, which patterns in ways predicted by the life cycle of phonological processes. However, the pattern is somewhat distorted by what appears to be a recent innovation entering the grammar from below: not only has the effect of a following pause changed from being deletion-favouring to deletion-inhibiting, the magnitude of this change points to the presence of an entirely new process of pre-pausal [ɡ]-insertion. These results provide further empirical evidence to support the life cycle of phonological processes, and in doing so speak to theories of language change and of the underlying architecture of grammar.

The structure of this paper is as follows. Section 2.1 outlines existing work on the diachronic development of /ɡ/-deletion across centuries of linguistic change, and
how this fits in with the theory of the life cycle; Section 2.2 makes explicit the predictions about how /ɡ/-deletion should behave synchronically, which emerge naturally from assumptions about the architecture of grammar and the stratified nature of the phonological module. The methodology of this study is described in Section 3. The results are split into three parts. Section 4.1 presents evidence of apparent-time change, which sees a dramatic increase in the rate of pre-pausal [ɡ]-presence in recent decades. Section 4.2 shows that the results of this innovation are synchronically reflected in patterns of inter-speaker variation: clustering analysis identifies a group of relatively older speakers with very low rates of [ɡ]-presence prepausally, reflecting the outcome of the life cycle of the original deletion process, alongside a group of relatively younger speakers with high rates of pre-pausal [ɡ]-presence brought about by an innovative pre-pausal insertion process. Finally, Section 4.3 presents evidence for domain-specific rates of application, testing another of the life cycle’s predictions. In Section 5.1 a new amphichronic account is proposed for (ng) variation in light of this new synchronic data; in Section 5.2 a revised formulation of the life cycle’s predictions is proposed, based on evidence of this new innovation; finally, in Section 5.4 these results are discussed in the wider context of homorganic nasal+stop cluster reduction in the history of English.

2 Amphichronicity and the life cycle of phonological processes

The life cycle of phonological processes (Bermúdez-Otero 2007, 2011; Ramsammy 2015; Sen 2016) is an ‘amphichronic’ approach to variation and change in that it seeks to combine diachronic and synchronic accounts that mutually inform one another (see also Kiparsky 2006): it specifies an architecture of grammar, which in turn makes predictions regarding synchronic variation, and pathways of change defined by that architecture. Bermúdez-Otero (2013) provides a detailed exploration of amphichronicity and the life cycle itself, but the central points will be summarised here to provide background for the current study.

The life cycle assumes a classical modular feedforward architecture of grammar, built of separate components that sees information travel serially through the grammar along interfaces that connect only adjacent modules, e.g. a phonology–phonetics interface, and a morphology–phonology interface, but no morphology–phonetics interface. Building upon ideas from Lexical Phonology (Kiparsky 1982a,b), the life cy-

---

1This isn’t to say that evidence of what appears to be morphologically-conditioned phonetics is incompatible with a modular architecture; in a life cycle framework it is possible for phonological and
cle also assumes stratification of the phonological module itself, into three domains:

- **STEM LEVEL**, where phonological rules apply to the stem

- **WORD LEVEL**, where rules apply to the word (i.e. after suffixation), but do not see across word boundaries

- **PHRASE LEVEL**, where rules can see across word boundaries

The life cycle predicts that an innovation starting out as purely extragrammatical should eventually undergo phonologisation through Ohalian hypocorrection (Ohala 1981), in which a learner misinterprets a form arising through physiological means as an actual phonetic target. At this point the process begins applying instead as a gradient rule of phonetic implementation. After this gradient effect has increased in magnitude, the rule can at a later date undergo stabilisation and be reanalysed as a categorical phonological rule. When this change takes place, the new phonological process begins applying in the widest morphosyntactic domain, i.e. the phrase level. Over time, the process may undergo successive rounds of domain narrowing into more embedded strata, applying later at the word level and eventually at the stem level. These modules, and the diachronic processes that govern progression between them, are illustrated in Figure 1.

While in historical innovations the direction is from lower to higher modules, in synchronic derivation information travels downwards; that is, the phonological computation of an underlying form first involves stem-level phonology, before being subject to word-level and then phrase-level processes, and finally rules of phonetic implementation.

The cyclic stratification of the phonological module and the process of domain narrowing that sees phonological rules progress through these strata both form a crucial component of this study, and as such they will be discussed in more detail in specific relation to /ɡ/-deletion.

### 2.1 Diachrony of (ng)

The presence of post-nasal [g] in words such as *sing* and *young* – occasionally referred to as ‘velar nasal plus’ in the sociolinguistic literature – is a dialectal feature exclusive to varieties spoken in the North West and West Midlands of England (Wells 1982b).
Variation in these clusters will hereafter be referred to as (ng), following standard sociolinguistic convention.

The presence of post-nasal [ɡ] was not always so regionally restricted, however; up until the Late Modern English period [ŋɡ] was present in all regional varieties, before it began to undergo deletion. There is historical evidence to suggest that this coda-targeting deletion rule, which ran to completion in all other varieties of British English, progressed through the grammar in ways predicted by the life cycle of phonological processes.

Based on reports from 18th century orthoepist James Elphinston, the behaviour of post-nasal /g/-deletion at various stages of its life cycle has been reconstructed to provide historical evidence of domain narrowing during the Late Modern English period (Garrett & Blevins 2009; Bermúdez-Otero 2011; Bermúdez-Otero & Trousdale 2012). In Elphinston’s speech, /ŋɡ/ clusters exhibit stylistic stratification such that in formal speech (Elphinston’s conservative register) [ɡ] is always present before a vowel, but in casual speech (his more innovative register) it is only present before a vowel within the same word. This is said to reflect the direction of change, specifically the effects of domain narrowing from phrase level to word level.

This is illustrated in Table 1. When deletion is a phrase-level process (Stage 1), it only applies when /ŋɡ/ is phrase-final or followed by a consonant-initial word; it is bled by phrase-level resyllabification in contexts such as sing it when a vowel-initial word follows, because the /g/ can resyllabify as an onset at the phrase level and thus save itself from deletion. Only when deletion progresses to the word level (Stage 2),
Stage Realisation of underlying /ŋɡ/ Rule Period or variety

<table>
<thead>
<tr>
<th>Stage</th>
<th>/ŋɡ/</th>
<th>/ŋɡ/</th>
<th>/ŋɡ/</th>
<th>/ŋɡ/</th>
<th>/ŋɡ/</th>
<th>—</th>
<th>EModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>—</td>
<td>EModE</td>
</tr>
<tr>
<td>1</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>phrase</td>
<td>Elphinston (formal)</td>
</tr>
<tr>
<td>2</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>word</td>
<td>Elphinston (casual)</td>
</tr>
<tr>
<td>3</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>/ŋɡ/</td>
<td>stem</td>
<td>Present-day RP</td>
</tr>
</tbody>
</table>

Table 1: The life cycle of post-nasal /ɡ/-deletion, highlighting the effects of domain narrowing on its application in different morphophonological environments. Adapted from Bermúdez-Otero (2011: 2024).

and is therefore blind to phrasal content, can it apply to word-final /ŋɡ/ regardless of what follows. It later progressed to the stem level (Stage 3) where it applies before suffixation, evidenced by [ɡ]-absence in words such as *singer* in almost all varieties of Present Day English; in such cases, deletion can apply to the stem *sing* where [ɡ] is in the syllable coda, before it has chance to become an onset through word-level suffixation of *-er*.

Simulations of grammar acquisition by Lignos (2012) provide further evidence for domain narrowing in the evolution of post-nasal /ɡ/-deletion: drawing upon a corpus of child-directed speech (CHILDES, MacWhinney 2000) and appealing to Yang’s TOLERANCE PRINCIPLE (Yang 2005), Lignos shows that the probability of input restructuring is dependent on the level of ambiguity between positing rules in different domains. This ambiguity can be quantified by calculating the number of exceptions to each stage using actual corpus frequency data; crucially, if this level of ambiguity is high enough and the number of exceptions does not exceed tolerance, reanalysis will occur and the rule will move up into a higher domain.

Take the example of phrase-level to word-level domain narrowing. For input restructuring to take place between these domains, learners must posit a [ɡ]-less form as the input to the phrase level, rather than [ɡ]-less forms being derived from a phrase-level deletion process. When /ɡ/-deletion takes place at the phrase level, the alternation is transparent: [ɡ] is present before a vowel, and absent before a consonant or pause. For a learner to produce this alternation, they must model their input to the phrase level as *[sɪŋɡ]*, which then undergoes phrase-level deletion if pre-consonantal or pre-pausal but not in pre-vocalic position. However, given that (ng) occurs in

Although not included in Table 1, this life cycle of /ɡ/-deletion can be extended to include varieties such as Scots, which exhibits deletion even in *finger*-type items (Johnston 1997). This simply reflects a case of rule generalisation (Kiparsky 1988; Bermúdez-Otero 2007), such that in Scots, deletion no longer targets weak position in the syllable (i.e. the coda) but weak (i.e. non-initial) position in the prosodic foot instead; that is, in addition to a change of morphosyntactic domain through domain narrowing, processes can also begin to apply in higher prosodic environments (see Turton 2014, 2016 on both coda- and foot-based /l/-darkening in English).
pre-consonantal and pre-pausal position three times more often than in pre-vocalic position, the chances of the learner incorrectly positing a [g]-less form as input to the phrase level, and thus deletion climbing up into the higher word-level domain, is relatively high.

Despite this uneven distribution of phonological environments, the Lignos (2012) simulations suggest that the first round of domain narrowing encountered more resistance than the second; input restructuring from phrase- to word-level only takes place under certain conditions on post-lexical resyllabification, e.g. that it is not onset-maximising and only takes place before vowel-initial words, not words with initial sonorous consonants such as [l] or [ɹ]. In contrast to this, the rule’s development from the word level to the stem level progressed with relative ease. As explained by Bermúdez-Otero (2013: 386), the vulnerability of word-level processes to this kind of analogical change in English is not surprising given its ‘impoverished’ inflectional system and the consequence that stem-final consonants rarely surface as onsets; this is further reflected by processes that remain stuck at the word level in languages where word-level suffixes beginning with vowels are used more frequently, such as coda-devoicing in Dutch (Booij 1995).

2.2 Synchronic predictions

In addition to laying out a diachronic trajectory of change, the life cycle also makes falsifiable predictions with respect to how /g/-deletion should behave synchronically. These predictions naturally fall out from the architecture of grammar and in particular the stratified nature of the phonological module. When a rule progresses into a higher domain, it often leaves behind an existing avatar in the original stratum; consequently, for speakers in the North West and West Midlands of England who still have a synchronic grammar that produces variation in (ng) clusters, there are in fact three phonological deletion rules: one that applies to stems, one that applies to words, and one that applies post-lexically. All are probabilistic in nature, i.e. they are processed with some variable rate of application.

If we consider the four morphophonological environments from Table 1 from a synchronic standpoint instead, as in Table 2, it is clear to see that certain tokens of (ng) will meet the criteria for deletion in more of the morphosyntactically-defined phonological strata than others.

Specifically, the post-nasal [g] is in onset position throughout the derivation of finger-type words, as there is no stem fing in which [g] is in the coda. Consequently, these tokens should exhibit only trivial rates of deletion, if any at all. singer-type tokens are only subject to deletion once – at the stem level – because upon reaching
word-level computation the [ɡ] will move to the onset of the second syllable. Tokens such as sing it are exposed to two rounds of deletion, with [ɡ] in the coda at the stem and word levels before undergoing resyllabification across word boundaries at the phrase level to become an onset of the following vowel-initial word. Pre-consonantal tokens (e.g. sing tunes) meet the criteria for deletion in all three cycles, as do pre-pausal tokens, because in neither case can the [ɡ] resyllabify as an onset to bleed the coda-targeting deletion rule. For convenience, these five sets of morphophonological environments will hereafter be referred to using the following labels:

- **FINGER**: pre-vocalic /ŋɡ/, word-medial (stem-medial)
- **SINGER**: pre-vocalic /ŋɡ/, word-medial (stem-final)
- **SING IT**: pre-vocalic /ŋɡ/, word-final
- **SING TUNES**: pre-consonantal /ŋɡ/, word-final
- **SING‖**: pre-pausal /ŋɡ/, word-final

The prediction is clear: if /ɡ/-deletion applies cyclically, tokens that meet the criteria in more cycles should be exposed to more rounds of probabilistic deletion during the derivation, and so on the surface should be less likely to exhibit [ɡ]-presence. That is, we expect a cline of [ɡ]-presence, decreasing from singer to sing it to sing tunes and pre-pausal sing‖.

Although synchronic variation in (ng) has yet to be analysed within a cyclic framework, there have been claims that deletion is more frequent pre-consonantally than pre-vocally (Knowles 1973; Upton et al. 1987; Watts 2005). Given these reports, and the diachronic evidence of the life cycle of /ɡ/-deletion as discussed in Section 2.1, synchronic variation in (ng) may well be amenable to such an analysis.

This kind of empirical prediction is also not new to the field of variationist linguistics: it actually builds on earlier work in the Lexical Phonology framework, such as
Guy’s (1991a; 1991b) study of /t,d/-deletion in American English. One of the most robust predictors of variation in (td) is the morphological category of the word (Wolfram 1969; Guy 1980; Neu 1980; Santa Ana 1992; Fruehwald 2012), such that regular past-tense items (e.g. missed) exhibit less deletion than irregular ‘semi-weak’ items (e.g. kept), which in turn exhibit less deletion than monomorphemic items (e.g. mist). In this pair of influential papers, Guy argues that this effect stems from repeated exposures to a deletion process: the word-final [t]/[d] attaches later in the derivation when it belongs to a past-tense morpheme compared to a monomorphemic item where the targeted segment is part of the stem and present throughout the derivation. The intermediate status of semi-weak items is said to result from how inflectional endings attach at different levels depending on their regularity: irregular inflection operates at level 1 and is thus exposed to fewer rounds of deletion than items with regular past tense inflection (which attaches at level 2).

However, Guy (1991b: 8) assumes the same rate of application in each domain. To exemplify, if we adopt a hypothetical application rate of 50%, a word that meets its structural description once in the derivation should have 50% segment presence. This should decrease to 25% for words exposed to two rounds of deletion, and decrease further still to 12.5% for words exposed to three rounds. However, this is not necessarily a safe assumption to make. As Turton (2016) explains in relation to /l/-darkening, under a life cycle framework we would expect a correlation between the rate of application and the depth of the cyclic domain in which it applies, at least while the change is active and in progress; that is, because a rule will have been active for longer at the phrase level, where it began, it should apply at higher rates than in a more embedded domain, such as the stem level, where it is much younger. This is referred to as the Variation Corollary of the Russian Doll Theorem, and is formulated as follows:

If a phonological process $\pi$ shows a rate of application $x$ in a small embedded domain $\alpha$, then $\pi$ will apply at a rate equal to or greater than $x$ in a wider cyclic domain $\beta$. (Turton 2016: 139)

Testing these predictions regarding the synchronic behaviour of post-nasal /ɡ/-

---

3Despite the robustness of this effect in varieties of American English, it has been claimed to be absent in British English (see Tagliamonte & Temple 2005; Temple 2009 on York English). There is recent evidence to suggest that this effect is present in Manchester English, but that it plays a much smaller role compared to the variation in American English and as such is only detectable in large-scale corpora (Baranowski & Turton forthcoming).

4Alternatively, the intermediate deletion rates in semi-weak items could reflect inter-speaker variation with respect to their representation: Guy & Boyd (1990) argue that age-grading within this class of items suggests that younger speakers are less likely to parse the /t/ of kept as a separate morph.
deletion, both of which are rooted in its diachronic pathway of change, forms the basis of this paper.

3 Methodology

In order to answer these research questions, data is drawn from a collection of sociolinguistic interviews conducted in the North West of England. Although using a prepared sentence list would make it possible to elicit tokens in different morphophonological environments, thus providing more data as well as controlling for possible confounds of speech rate, it would also have introduced a stylistic confound. The use of conversational data circumvents the 'observer’s paradox' (Labov 1972b) and provides a more reliable insight into the variation of these /ŋɡ/ clusters, which may well be levelled in more formal discourse styles.

3.1 Sociolinguistic interviews

In total, the corpus contains 30 sociolinguistic interviews, largely conducted between 2015 and 2017. On average, these lasted for about an hour, and were structured to follow typical conventions as described by Tagliamonte (2006: 37–49) and as used by Labov (1984) in Philadelphia; the interviews consisted of open-ended questions about a number of topics such as childhood, school life, the neighbourhood, and travel. Many of the questions were designed specifically to elicit narratives of personal experience, which provide the most direct access to a speaker’s vernacular (Labov 2010b). The interviews were followed up with two elicitation tasks, a word list and a reading passage, which contain tokens of (ng) but are not subject to analysis in this paper for the afore-mentioned reasons.

These interviews were recorded using a Sony PCM-M10 recorder and a lavalier microphone, saved in uncompressed WAV format at a sampling rate of 44.1KHz. They were later transcribed orthographically using ELAN, and force-aligned using the FAVE suite (Rosenfelder et al. 2011) to produce a time-aligned phone- and word-level TextGrid allowing for more efficient analysis.

3.2 Participants

The 30 participants of these interviews were all born and raised in the North West of England, specifically from the urban centres of Greater Manchester and Blackburn and their surrounding regions.

The population sample is stratified by speaker sex (14M; 16F) and is evenly distributed with respect to date of birth. Socioeconomic status was controlled for by
restricting the population sample to upper working class speakers, with this classification based primarily on occupation (see Baranowski 2017: 303 for a similar operationalisation of social class in Manchester).

The distribution of speaker ages is particularly important as it allows for an apparent time investigation of possible diachronic change, where the vernacular of older speakers is said to be representative of an earlier stage of the development of this dialect. This is made further possible by the inclusion of two interviews conducted in 1971, providing extra time depth with date of births spanning almost a century – from 1907 to 1998.

### 3.3 Data annotation

The envelope of variation is defined as any underlying /ŋɡ/ cluster that appears in stem-final position with primary stress, such as young, sing-er, wrong, hang-ing, i.e. words that are invariably realised with the plain velar nasal [ŋ] in most varieties of the English-speaking world. Tokens of pre-vocalic /ŋɡ/ that appear in monomorphemic or root-based items, such as finger, bungalow, elongate, tango etc., were excluded from the analysis after confirming that they do indeed invariably surface with [g]-presence.

The dependent variable was manually coded in a binary fashion, based on categorical presence/absence of a post-nasal stop. There was very little phonetic variability in the realisation of tokens coded as having [g]-presence, with the stop often surfacing as voiceless and occasionally with an ejective release when in phrase-final position. Prototypical examples of tokens with and without post-nasal [g] are given in Figure 2.

Among a number of other independent variables (e.g. speech rate, word frequency, and grammatical category), each token was also annotated for the immediate phonological environment (whether the following segment is a vowel, obstruent,
glide, liquid, nasal or pause); this, along with the morphological composition of the word, defines the relevant environments outlined earlier in Section 2.2.

Although in Garrett & Blevins (2009) it was argued that Elphinston’s post-nasal [ɡ] undergoes resyllabification before liquids as well as vowels, in this data there is no evidence that the synchronic system works in this way; there is no significant difference between the rates of [ɡ]-presence before liquids and nasals/obstruents/glides, suggesting that only pre-vocalic tokens should be included in the sing IT environment. This indicates that phrase-level resyllabification of post-nasal [ɡ] is not onset-maximising; that is, it occurs if the following word is vowel-initial but not if the word begins with a sonorous consonant, even if resyllabification would form a phonotactically-valid complex onset. This is particularly interesting in light of the simulations reported by Lignos (2012), discussed earlier in Section 2.1, which suggest that such a restriction was also in place during the diachronic development of this process.

In total, the interviews contain 1,444 tokens of (ng) from spontaneous, conversational speech; the 446 elicited tokens from the word list and reading passage are not discussed in this paper.

4 Results

The results of this analysis reveal that synchronic variation in (ng) is strongly predicted by morphophonological factors in ways predicted by the life cycle of phonological processes: there is an almost perfectly-linear relationship between the rate of [ɡ]-presence and the number of times it appears in a deletion-targeting environment (i.e. the coda) throughout the cyclic derivation. When deletion has one chance to apply (i.e. SINGER, before a tautolexical vowel) post-nasal [ɡ] is present 82% of the time; this decreases to 57% when deletion has two chances to apply (i.e. SING IT, before a heterolexical vowel), and decreases further still to 22% when deletion can apply in all three cyclic domains (i.e. SING TUNES and SING], before a consonant or pause).

However, the existing discussion of cyclic coda-targeting rules, at least in the case of post-nasal /ɡ/-deletion, makes an implicit assumption that should not be overlooked: cases in which an underlying /ɡ/ is subject to three rounds of deletion – the category of tokens that shows the highest overall rate of deletion – actually encompass two distinct phonological environments. The deletion rule can apply at the stem, word, and phrase levels if the underlying /ɡ/ occurs pre-consonantly or phrase-finally, because in neither case is it possible for the /ɡ/ to resyllabify as an onset and save itself from deletion in any of the cyclic domains. Although this makes logical
sense from a purely cyclic position, the two environments do differ with respect to prosody, and when they are considered separately it becomes clear that they condition (ng) variation in drastically different ways.

As indicated by Figure 3, the pre-pausal environment is strongly [ɡ]-favouring; in fact when we aggregate over the whole population sample in this way, a following pause is second only to a following tautological vowel in the degree to which it favours [ɡ]-presence. When we look at this on a speaker-by-speaker basis, as in Figure 4, it is clear that this distinctive V-shaped pattern is evident for many speakers in this sample; there are in fact a number of speakers who never show [ɡ]-absence in pre-pausal position (see for example GraceG, MollyF, WendyJ etc.).

The high rate of pre-pausal [ɡ]-presence is particularly problematic if it is assumed that surface variability in (ng) is derived purely from cyclic application of a probabilistic /ɡ/-deletion rule, hereafter referred to as an Elphinstonian grammar in light of the historical evidence discussed in Section 2.1. To be more specific, an Elphinstonian grammar refers to a system in which the distribution of [ŋ] and [ŋɡ] involves cyclic application of deletion alone; empirically, this in turn leads to a falsifiable prediction: that there cannot be more [ɡ]-presence pre-pausally than before a hetero-lexical vowel, since the former environment is subject to three rounds of deletion and the latter only two.

Consider the simulations in Figure 5, all of which are consistent with a cyclic account of (ng). A hypothetical dataset of 1,200 tokens, equally split between the four morphophonological environments under consideration, is input to a simulated grammar. Each token in this dataset is exposed to either one, two, or three proba-
Figure 4: The rate of [g]-presence by morphophonological environment, for individual speakers.
bilistic deletion rules depending on its morphophonological properties and the syllabic status of [ɡ] in each cyclic domain, with all three deletion rules having their own domain-specific rate of application as discussed in Section 2.2.

In (a), the phrase-level deletion rule is not sensitive to pause, which likely approximates Elphinston’s own system given that he does not report any such effect; in (b), the phrase-level deletion rule applies at a lower rate pre-pausally than pre-consonantally; in (c), it is blocked completely before pause. Crucially, as evidenced by these simulations, in an account driven by cyclic deletion alone, the rate of [ɡ]-presence pre-pausally cannot be greater than before a hetero-lexical vowel (i.e. SING IT); at most they can be equal, which would reflect the scenario in (c) in which the phrase-level deletion rule is blocked in this environment and as such the tokens are only exposed to two probabilistic deletion processes just like the SING IT tokens.

However, as shown in Figure 4, there are clearly a number of speakers for whom the rate of [ɡ]-presence pre-pausally is higher than in SING IT; this is incompatible with an account that solely involves cyclic deletion in non-onset positions.

In light of this, there are two likely explanations: either synchronic variation in post-nasal [ɡ]-presence is not derived through cyclic application of a deletion rule, or there has been a separate innovation in the development of this dialect, distorting what would otherwise be a perfectly Elphinston-compliant pattern of variation as predicted by the life cycle of phonological processes. In the following section, I provide strong empirical evidence for the latter explanation.

**Figure 5:** Simulated grammars (1,000 iterations) with the following domain-specific rates of application of /ɡ/-deletion: $R_{SL} = 0.2$; $R_{WL} = 0.3$; pre-consonantal $R_{PL} = 0.7$; pre-pausal $R_{PL} = 0.7$ in (a), 0.3 in (b), and 0 in (c).
Figure 6: Apparent time change in the rate of pre-pausal [ɡ]-presence; pre-consonantal environment included as a baseline for comparison. Individual speaker means are plotted as points; lines reflect linear models with 95% confidence intervals.

4.1 Change in progress

Importantly, the inter-speaker variation with respect to the behaviour of pre-pausal (ng) is not unconstrained but rather shows a strong correlation with date of birth. As illustrated in Figure 6, we have apparent-time evidence of change in progress towards increasing [ɡ]-presence in pre-pausal position, suggesting that this is actually a relatively recent innovation in this community.

This change finds statistical support from mixed effects logistic regression, with a model that includes [ɡ]-PRESENCE as the dependent variable and independent variables of MORPHOPHONOLOGICAL ENVIRONMENT, DATE OF BIRTH, and their interaction. Crucially, as reported in the model summary in Table 3, the interaction only reaches significance in the pre-pausal environment. Further evidence for this effect comes from ANOVA comparisons between nested models with and without the ENVIRONMENT-DATE OF BIRTH interaction; the inclusion of this interaction leads to a statistically-significant decrease in AIC (1046, cf. 1063; \( p < 0.001 \)) and therefore a better-fitting model.

This change in progress suggests that younger speakers actually have a new system: normal cyclic application of /ɡ/-deletion, giving rise to the monotonic patterning in the SINGER, SING IT, and SING TUNES environments, overlaid with this pre-pausal innovation at the phrase level. Crucially, this innovation cannot simply be a blocking of the phrase-level deletion process, but instead must be a new process of [ɡ]-insertion superposed upon the existing phrase-level deletion rule. That is, although some of the
### Table 3: Mixed-effects logistic regression model; [ɡ]-presence as the application value; random intercepts of speaker and word; sing tunes as the environment reference level.

| Fixed effects       | Estimate | Std. Error | z-value | Pr(>|z|) |
|---------------------|----------|------------|---------|----------|
| (Intercept)         | -2.1164  | 0.2900     | -7.2986 | <0.001 *** |
| Environment         |          |            |         |          |
| singer              | 3.3910   | 0.4131     | 8.2080  | <0.001 *** |
| sing it             | 2.8805   | 0.2597     | 11.0924 | <0.001 *** |
| sing                | 4.1011   | 0.3273     | 12.5321 | <0.001 *** |
| Environment × Date of birth |      |            |         |          |
| singer : dob        | -0.3175  | 0.3072     | -1.0332 | 0.3015   |
| sing it : dob       | -0.3533  | 0.2113     | -1.6722 | 0.0945   |
| sing : dob          | 0.7801   | 0.2598     | 3.0030  | 0.0027 ** |
| Date of birth       |          |            |         |          |
| date of birth (scaled) | 0.2400  | 0.2171     | 1.1054  | 0.2690   |

Pre-pausal tokens of [ɡ] are ‘survivors’ of deletion, many of them are in fact likely to have been inserted.

Regardless of the nature of this pre-pausal innovation, the evidence of recent diachronic change suggests that there was once a point in the history of this dialect at which there was more [ɡ]-presence before a hetero-lexical vowel than before a pause, i.e. a system that develops through /ɡ/-deletion progressing along its life cycle, with no separate innovation.

Importantly, we do not have to rely on Elphinston’s testimony for this, as there are in fact a number of speakers in this data set who show such a pattern. Although there is not enough statistical power to diagnose significant differences between environments at the level of the individual, it is possible to aggregate over groups of speakers who show similar patterns of variability. This grouping can also be carried out in an objective manner using cluster analysis rather than by subjectively hand-picking speakers.

### 4.2 Clustering analysis

Given the degree of heterogeneity within this community, it is clear to see how aggregating over the whole population sample would be problematic in obscuring this inter-speaker variation. Having presented evidence of this apparent time change, highlighting that the probability of a speaker having a particular system of (ng) variation is correlated with date of birth, one option would be to discretise the time dimension and group speakers based on age. However, it is unlikely that all speakers within a single date of birth cohort are equally advanced in this change, which means such aggregation may conflate speakers with theoretically (and empirically) different
Figure 7: Results of hierarchical cluster analysis, based on relative rates of [g]-presence in each of the four morphophonological environments under study.

systems. It also necessitates the use of arbitrary boundaries if date of birth is to be split into discrete groups.

By instead performing cluster analysis, it is possible to aggregate one level above the individual and avoid all of the afore-mentioned problems. Clustering is a method widely used in statistical data analysis in which sets of objects are grouped together based on some quantifiable similarity (Everitt 1974). In this case, it is possible to perform hierarchical clustering based on the pattern of (ng) variation across all four morphophonological environments, such that speakers with similar ‘systems’ are clustered together. Using this method, implemented in R using the hclust function, speakers are grouped together based on their actual behaviour and we can then test the hypothesis of change in the opposite direction; that is, we can test whether or not these objectively-different systems, as determined by cluster analysis, are correlated with date of birth.

The results of hierarchical classification are given in Figure 7, where the three main clusters are characterised by the behaviour of word-final (ng) in pre-vocalic and pre-pausal environments.

Interestingly, the type of pattern also correlates with date of birth; the speakers in cluster #1 have a median date of birth of 1955, which increases to 1966 for cluster #2, and 1980 for cluster #3. Focusing on cluster #1, Figure 8 aggregates over its five
speakers to plot the cluster-wide pattern of variation, revealing a perfectly monotonic pattern; these speakers in cluster #1 are clearly those with a pure Elphinstonian system of /g/-deletion, with no evidence of phrase-level innovation and a significantly lower rate of word-final [ɡ]-presence pre-pausally than pre-vocally ($\chi^2 = 4.09$, df = 1, $p = 0.043$). For these five speakers, no innovation has taken place: their grammars reflect a system in which the alternation between [ŋ] and [ŋɡ] is derived solely through cyclic deletion.

4.3 Domain-specific deletion rates

The life cycle makes predictions not only about the relative proportion of [ɡ]-presence in different morphophonological environments, but also regarding the rates of deletion in each of the three cyclic domains. This concept, discussed earlier in Section 2.2, has already been tested by Turton (2016) for /l/-darkening and the same methods can be applied here to calculate domain-specific rates of deletion. Recall that the prediction, made explicit by Turton (2016) as the Variation Corollary of the Russian Doll Theorem, is that as long as the change is active and in progress, stem-level deletion should apply at lower rates than word-level deletion, which in turn should

---

Footnote: It is important to note that while $p$ is close to the alpha level of 0.05, even a non-significant $p$-value would not be counter-evidence of an Elphinstonian system; this would simply indicate that there is no significant difference between the rates of pre-pausal and pre-vocalic [ɡ]-presence, which is still compatible with a system in which the surface alternation of [ŋ]–[ŋɡ] stems only from a cyclic deletion rule with no separate innovation.
be lower than phrase-level deletion (contra Guy 1991a,b, who assumes the same rate of application across all strata). This follows naturally from the fact that deletion began at the phrase level, and has therefore been active longest here, before climbing up into more embedded domains.

Calculating domain-specific rates of deletion is fairly trivial from a mathematical standpoint; SINGER-type tokens are only subject to deletion in one cycle – at the stem level – which means the rate of [g]-presence in this environment tells us what the rate of retention is at the stem level. This is shown in (1), where \( R \) is the retention rate and \( D \) is the deletion rate (simply \( 1 - R \)):

\[
\begin{align*}
R_{SL} &= R_{singer} \\
R_{SL} &= 0.8156 \\
\therefore D_{SL} &= 0.1844
\end{align*}
\]

Tokens in the SING IT environment are subject to two rounds of deletion – at the stem and word levels – and since we have already calculated the rate of application at the stem level, we can isolate the WL process as in (2):

\[
\begin{align*}
R_{s\text{ing\, IT}} &= R_{SL} \times R_{WL} \\
&= 0.8156 \times R_{WL} \\
R_{WL} &= \frac{R_{s\text{ing\, IT}}}{0.8156} \\
&= \frac{0.5735}{0.8156} \\
R_{WL} &= 0.7031 \\
\therefore D_{WL} &= 0.2969
\end{align*}
\]

It would be unwise to aggregate over the whole population for the pre-pausal environment given the vigorous change taking place here, so the phrase-level calculations are based solely on pre-consonantal (ng). Tokens in this environment are subject to three rounds of deletion – at the stem, word, and phrase levels – so we can simply apply the same methods as before to isolate the phrase-level deletion rate; this is done in (3):
\begin{align*}
R_{\text{sing tunes}} &= R_{SL} \times R_{WL} \times R_{PL} \\
R_{\text{sing tunes}} &= 0.8156 \times 0.7032 \times R_{PL} \\
R_{PL} &= \frac{R_{\text{sing tunes}}}{(0.8156 \times 0.7031)} \\
R_{PL} &= \frac{0.0995}{0.5735} \\
R_{PL} &= 0.1734 \\
\therefore D_{PL} &= 0.8266
\end{align*}

These cyclic-specific deletion rates are illustrated in Figure 9 for a hypothetical input of 1,000,000 tokens of (ng), indicating the number that are predicted to undergo deletion in each cycle. The domain-specific rates fall in line with predictions, decreasing from phrase (83%) to word (30%) to stem (18%) levels, but perhaps most interesting of all is how the word-level deletion is much closer to the stem-level rather than the phrase-level process. Recall that the simulations carried out by Lignos (2012), discussed in Section 2.1, suggest that the domain narrowing from word to stem level progressed fairly rapidly for (ng), most likely due to the reduced inflectional system of English and as a consequence of this the vulnerability of stem-final consonants to coda-targeting rules (Bermúdez-Otero 2013). This is arguably also reflected here: taking the domain-specific rate of application as a proxy for the age of each deletion rule, it suggests that domain narrowing from phrase level to word level took much longer than the narrowing from word level to stem level.

In summary, not only do these domain-specific deletion rates fall in line with the predictions made by the life cycle and the Variation Corollary of the Russian Doll Theorem, they also corroborate the results from independent simulations and as a result shed light on the speed at which /g/-deletion underwent successive rounds of domain narrowing.

5 Discussion

The results presented in Section 4 cast new light on the synchronic variation in (ng); at this point it is important to situate these results in the existing knowledge of the diachrony of (ng), and as such complete the amphichronic picture as was set out at the beginning of this paper.
5.1 A diachronic and synchronic account of (ng)

From a number of historical sources we can attest the presence of [ŋɡ] (alongside other homorganic nasal+stop clusters) both word-medially and word-finally in Proto-Germanic (Ringe 2006) and Old English (McCalla 1984; Voyles 1992; Hogg 2002), e.g. OE hringan ‘to ring’ and hring ‘ring’. From Elphinston’s testimony, we also know that this remained the case up until Late Modern English, at which point a process of /ɡ/-deletion had begun progressing through the grammar along a trajectory predicted by the life cycle of phonological processes, i.e. beginning in the phrase-level domain, before undergoing successive rounds of domain narrowing into more embedded morphosyntactic domains (Garrett & Blevins 2009; Bermúdez-Otero 2011; Bermúdez-Otero & Trousdale 2012).

The fact that this rule has run to completion in all other dialects outside of the North West and West Midlands of England, such that [ŋɡ] only occurs in a restricted set of environments⁶, lends support to this diachronic account. In this paper, it has been shown that the synchronic system of (ng) variation in the North West of England reflects this diachronic trajectory of /ɡ/-deletion, but also that it shows evidence of a more recent innovation: there appears to be a separate process of [ɡ]-insertion taking place in pre-pausal contexts, with the surface effects of this process increasing in magnitude in apparent time.

---

⁶These being monomorphemic or root-based items such as finger or elongate, as well as the exceptional comparative and superlative forms of long, strong and young.
An alternative explanation is that synchronic variation in (ng) is detached from the historical facts and the surface alternation between [ŋ]-[ŋɡ] in these northern varieties does not stem from cyclic phonological deletion at all. However, treating the morphosyntactic environments independently from one another overlooks the fact that for three of the four contexts, the behaviour of (ng) falls perfectly in line with the predictions made by the life cycle, outlined in Section 2.2. Furthermore, given the strength of the correlation between pre-pausal [ɡ]-presence and date of birth, it is highly likely that an earlier stage of the dialect did behave in a purely Elphinstonian manner. Indeed, results from the clustering analysis reported in this paper suggest that for some of the older speakers in this corpus this is exactly the case.

Although this paper provides evidence of innovation at the phrase level, it is difficult to pinpoint exactly when this change actually took place. It is not possible to conclusively identify [ɡ]-insertion at work until it applies at such a rate that, under this framework at least, it would be impossible for it not to be active, i.e. until the rate of pre-pausal [ɡ]-presence exceeds the rate of pre-vocalic [ɡ]-presence in sinɡ ɪt. At this point, we know that insertion must be active at the phrase level, but of course the innovation could have taken place even in a more conservative grammar that is still apparently life cycle compliant.

When the rate of pre-pausal [ɡ]-presence begins to increase, it remains apparently compliant with a cyclic analysis as long as it remains lower than the rate of [ɡ]-presence before a hetero-lexical vowel, although the actual likelihood of there being no separate innovation decreases. In other words, there is a point at which the dialect starts to show conclusive evidence of insertion, but that isn’t necessarily the point at which the innovation was genuinely actuated; it may have already begun at an earlier stage of the dialect.

It is also important to note that even in a purely Elphinstonian system of cyclic /ɡ/-deletion, it is permissable to have slightly more [ɡ]-presence in pre-pausal position relative to the sinɡ ɪt environment, simply because of the anti-conservative nature of how this environment has been defined and coded. All cases of word-final (ng) before a vowel-initial word were coded as belonging to the sinɡ ɪt environment, but in reality not all of these post-nasal stops will have undergone phrase-level resyllabification due to its sensitivity to prosodic factors such as speech rate and the temporal distance between the two words.

It is highly likely that some tokens coded in the sinɡ ɪt category were not resyllabified as onsets at the phrase level and were therefore subject to three rounds of deletion rather than two; as such, the surface rate of [ɡ]-presence in this category
may have been underestimated.

5.2 Life cycle predictions

In light of these results, I suggest a reformulation of the life cycle’s predictions that has not thus far been made explicit.

If there truly has been an innovation taking place at the phrase level, the predictions made by the life cycle do not necessarily have to hold; in fact they are expected not to hold. The theory of the life cycle states that changes enter the grammar from below and over time undergo progressive rounds of domain narrowing, climbing up into higher, more embedded strata; as such, its effects are first seen at the phrase level, then at the word level, and even later at the stem level. Given this trajectory of change, it is also predicted that the rate of application of a process should be inversely correlated with the narrowness of the morphosyntactic domain in which it applies; that is, because a process travelling along this life cycle will have been active longest in lower levels, it should apply at higher rates in those domains (see Turton 2016 on the Variation Corollary of the Russian Doll Theorem). However, this prediction only holds if there have been no further developments at the phrase level. If such a development has taken place, the surface pattern of variation will show traces of the life cycle of the original process along with the superimposed effect of the new development at the phrase level, where it has entered the grammar from below. This is exactly what has been described here for (ng) in the North West of England, i.e. a pattern of variation compatible with a life cycle of /ɡ/-deletion, overlaid with a new phrase-level innovation, evidence of which is provided from clear apparent time change over the past century.

5.3 Why has this innovation taken place?

Any discussion of why this innovation has taken place encounters issues relating to the actuation problem (Weinreich et al. 1968), but we can speculate as to the possible motivations behind this change. It is possible that social evaluation plays a role in this change; one might expect such an effect to be registered most strongly in this environment given the salience of phrase-final position (see Sundara et al. 2011; Dube et al. 2016 for experimental evidence). Specifically, it could be the case that this change in production reflects a change in how the northern [ŋɡ] form is evaluated, accruing

---

It could be argued that if the following word is temporally distant enough such that phrase-level resyllabification has not taken place, then the (ng) token may be treated as being pre-pausal and would therefore not undergo deletion at the phrase-level anyway; however, it is possible that for some tokens the juncture between the two words is long enough to block resyllabification but too short for the environment to be considered pre-pausal.
local prestige with this evaluation concentrated in environments where its presence is most salient. While it has been claimed that [ŋɡ] does indeed have local prestige in these northern communities (Foulkes & Docherty 2007: 64; Beal 2004: 127), this explanation of evaluation-driven change requires direct attitudinal evidence, which has not yet been provided.

The new pre-pausal behaviour could instead be seen as a prosodic strengthening mechanism, alongside other phonetic correlates such as pre-boundary durational lengthening (Delattre 1966; Lehiste et al. 1976; Turk & Sawusch 1997; Cho et al. 2013). Parallels can also be drawn with increasing rates of phrase-final ejectivisation for voiceless stops in Glasgow English (McCarthy & Stuart-Smith 2013): the fact that this highly-salient ejectivised release has been found to be most frequent for post-nasal [k] is particularly interesting as this is the voiceless counterpart to the [ŋɡ] environment discussed in this paper. Although such a change would need to be attested in these same varieties of British English, taken together, these two phrase-final phenomena could be considered to be part of the same boundary-marking ‘velar fortition rule’, which may in turn have discourse-pragmatic functions relating to turn-taking and the negotiation of conversation. In pursuing this possibility, future work is necessary to provide more insight into these pre-pausal tokens and their discursive properties, e.g. differentiating turn-final tokens from examples where a speaker pauses but then continues speaking.

That this heterogeneity is restricted to pre-pausal position may not even be that surprising given evidence of other processes that behave in similarly unpredictable ways before pause, particularly when compared to the consistency of how following vowels and consonants condition the variation. Take the example of /s/-debuccalisation in South American varieties of Spanish, which is activated pre-consonantally, blocked pre-vocalically, but shows inter-dialectal variation with respect to its application pre-pausally (Harris 1983; Kaisse 1996). A similar state of affairs has been attested for /t,d/-deletion across varieties of English: whilst the ranking of segmental constraints is consistent with respect to vowels and consonants, the effect of a following pause is dependent on the dialect in question (Guy 1980; also compare Tagliamonte & Temple 2005 on York English with Hazen 2011 on Appalachian English).

Exactly why pauses have such variable behaviour on probabilistic lenition processes is not clear. In discussing the sensitivity of /t,d/-deletion to the immediate phonological environment, Guy (1980) invokes the feature-dissimilatory effects of the OBLIGATORY CONTOUR PRINCIPLE (McCarthy 1986; Yip 1988), i.e. how the cline of deletion rates from pre-obstruent to pre-liquid/glide to pre-vocalic positions reflects
how featurally-similar those segments are to the coronal stop undergoing deletion. Crucially, pauses by their very nature do not fit into this typology and are therefore argued to be “susceptible to differing analyses by different speakers or dialects” (Guy 1980: 27). Whilst this is reflected by inter-dialectal differences in the effect of pause on /t,d/-deletion, in the case of (ng) it would appear to be registered in diachronic instability, assuming the contextual sensitivity of these processes is indeed driven by the Obligatory Contour Principle.

5.4 Homorganic nasal+stop clusters

As mentioned earlier in this discussion, [ŋɡ] is just one of a number of homorganic nasal+stop clusters present in historical varieties of English, alongside [mb] and [nt]/[nd]. All of these clusters have been subject to reduction at some point in the history of the English language, but the point at which this occurs, and the magnitude to which this occurs, is dependent on the place of articulation. We can therefore construct a markedness-driven implicational hierarchy of coalescence as follows:

- most marked: [mb], which is reduced to [m] in all varieties during the Late Middle English period, such that in Present Day English [mb] clusters only appear when tautosyllabic in monomorphemic or root-based items, e.g. bombard [ˌbɒmˈbɑːrd], cf. bomb [ˈbɒm], bombing [ˈbɒmɪŋ] (Borowsky 1993; Bermúdez-Otero 2011).

- less marked: [ŋɡ], which is reduced to [ŋ] in most regional varieties during the Late Modern English period (Bermúdez-Otero & Trousdale 2012), but still exhibits variation in the North West and West Midlands of England (Knowles 1973; Watts 2005; this paper).

- least marked: [nt]/[nd], which are reduced to [n] in almost all regional varieties of Present Day English as part of a widespread process of /t,d/-deletion in consonant clusters (see Tagliamonte & Temple 2005; Tanner et al. 2017; Baranowski & Turton forthcoming, all on British English). Crucially, the variation is stable with no evidence that this reduction process is running to completion.

Interestingly, parallels can be drawn between this ordering of environments and more general cross-linguistic markedness constraints on place of articulation; although it is a contested issue, there is fairly widespread agreement on coronals being universally less marked than labials and dorsals (see Prince & Smolensky 1993; Hume 1996; Wilson 2001), and in the history of English we observe that only coronal nasal+stop clusters survive in all varieties. Although Rice (1996) argues that both coronal and
dorsal are unmarked, it is nevertheless difficult to argue against labials having the weakest case for unmarkedness (though see Hume 2003 for an exception to this). It is therefore fitting that they should be the first to undergo nasal+stop coalescence in the history of English, and the only type to be lost in all dialects without exception.

6 Conclusion

In this paper it has been shown how the pattern of (ng) variation among speakers of North West British English is highly structured and predicted almost entirely by internal factors that fall out naturally from the architecture of grammar and the way that /ɡ/-deletion has progressed through it. In this way, the synchronic variation here reflects centuries-old linguistic change from the Late Modern English period, discussed in Bermúdez-Otero & Trousdale (2012), as well as a relatively recent innovation that has been explored in this paper.

The results of this study suggest that a change has been taking place in these North Western varieties of British English with respect to how (ng) behaves pre-pausally: this environment has changed from being [ŋ]-favouring to [ŋɡ]-favouring, and as a result we have seen a change from a purely cyclic, life-cycle compliant grammar, to a grammar that shows highly divergent behaviour at the phrase level. This contemporary system exhibits evidence of pre-pausal [ɡ]-insertion overlaid on the original cyclic deletion rule, which means that for these younger speakers there are two sources of [ɡ]-presence: it can either surface through resyllabification bleeding the cyclic deletion rule, or through some new prosodic privilege of being pre-pausal. Crucially, cluster analysis reveals that within this population sample there is evidence of both types of grammar without needing to extrapolate from the apparent time change, i.e. evidence that a purely cyclic system of /ɡ/-deletion was present at an earlier stage of the dialect prior to the actuation of this innovation.

The implications of these results range in scope from issues specific to the life cycle theory to more general considerations in diachronic and synchronic phonological analysis and variationist linguistics.

The theory of the life cycle predicts an ordered set of synchronic grammars which result from a pathway of change involving cyclic application of a phonological /ɡ/-deletion rule (Bermúdez-Otero 2013). However, the results presented in this paper indicate that patterns of variation that may on the surface appear to be incompatible with predictions may in fact naturally occur through later innovation. That is, just because a process is progressing along its life cycle does not mean that there will be no further innovation entering the grammar from below; these superposed
processes, which emerge through separate innovations, can distort what would otherwise appear on the surface to be a perfectly regular pattern of variation in line with predictions. This provides an important caveat for existing formulations of the life cycle and their application to diachronic change and synchronic variation (see e.g. Ramsammy 2015; Sen 2016; Turton 2016, 2017).

The variable patterning of (ng) largely fulfills the two predictions made explicit in Section 2.2, namely the way that [ɡ]-presence varies across different morphophonological environments and how the domain-specific rates of deletion reflect the age of the process in each cyclic domain. In showing this, it lends support not just to a theory of language change, but also to an architecture of grammar more generally, i.e. a modular architecture in which the phonological component is further stratified into stem-, word-, and phrase-levels, across which /ɡ/-deletion can apply cyclically in ways not dissimilar to classical Lexical Phonology (Kiparsky 1982a,b; Guy 1991a,b).

More generally this account of (ng) highlights the importance of amphichronic explanation in variationist linguistics: by considering the way in which /ɡ/-deletion has progressed through the grammar, specifically through narrowing of its morphosyntactic domain and cyclic application across a stratified phonological module, it provides an explanation for why these morphophonological environments pattern in the observed way. The reverse also applies; that is, not only does the diachronic account of this variable provide an explanation for its synchronic variation, this synchronic behaviour can be interpreted as further evidence for diachronic accounts of the life cycle of /ɡ/-deletion.
Bibliography


Turton, Danielle. 2016. Synchronic stratum-specific rates of application reflect
diachronic change: Morphosyntactic conditioning of variation in English /l/-
darkening. *Papers in Historical Phonology* 1, 130–165.

Turton, Danielle. 2017. Categorical or gradient? An ultrasound investigation of /l/-
darkening and vocalization in varieties of English. *Laboratory Phonology: Journal
of the Association for Laboratory Phonology* 8(1), 1–31.


Watts, Emma. 2005. *Mobility-induced dialect contact: A sociolinguistic investigation of
speech variation in Wilmslow, Cheshire.* Doctoral dissertation, University of Essex.

for a theory of language change. In Winfred Lehmann & Yakov Malkiel (eds.),
*Directions for historical linguistics: A symposium*, 96–195. Austin, TX: University of
Texas Press.

University Press.

*Phonology* 18(1), 147–197.

Language Series, No. 5.


Yip, Moira. 1988. The obligatory contour principle and phonological rules: A loss of
Paper II


N.B. This paper has been accepted for publication in *Laboratory Phonology: Journal of the Association for Laboratory Phonology*. When citing, please use the volume, issue, and page numbers of the published version once this becomes available.
Ki(ng) in the north: Effects of duration, boundary and pause on post-nasal [g]-presence

George Bailey
University of Manchester

Abstract

This paper highlights a hitherto unreported change in progress among northern speakers of British English, with increasing post-nasal [ɡ]-presence in words like *sing* or *wrong* pre-pausally. The factors that condition this innovation are unclear due to collinearity between various boundary phenomena. The right edge of phrasal prosodic categories may be associated with boundary tones, final lengthening, and pause; consequently, the variable presence of [ɡ] appears to be affected by prosodic boundary strength, segmental duration, and the presence and duration of a following pause. These factors are teased apart through analysis of an elicitation task from 30 northern speakers, which reveals that [ŋɡ] clusters are conditioned most strongly by pause. Post-nasal [ɡ]-presence is only licensed when the following consonant-initial word is temporally distant, showing only minimal sensitivity to prosodic boundaries directly. The surface effect of segmental duration arises only indirectly through its collinearity with pause duration. Current theoretical approaches to external sandhi emphasize a range of different factors, including phonological representations of prosodic constituency, phonetic parameters like segmental duration, and psycholinguistic mechanisms of production planning. This paper provides quantitative evidence from an under-reported feature of northern English that bears directly on these debates.

Keywords: external sandhi; boundary; prosody; duration; pause; velar nasal plus; variation; lenition; dialects of English
1 Introduction

External sandhi processes, where a phonological alternation is triggered across word boundaries, have been subject to extensive study, particularly with respect to locality restrictions on their application and the implications this has for theories of speech planning (see Wagner 2012a on the Production Planning Hypothesis, and more recently Kilbourn-Ceron 2017; Tamminga 2018). However, formal accounts of how these processes exhibit sensitivity to phrasal boundaries often fail to capture the various ways in which such an effect may be conditioned. The collinearity between boundary phenomena such as pause and phrase-final segmental lengthening poses a serious problem for research into the mechanisms conditioning such effects: are they conditioned directly by adjacency to prosodic boundaries of particular strengths, or do they reflect a more general sensitivity to segmental duration or pause? This study seeks to disentangle the close relationship between these factors, and does so by investigating one particular case of external sandhi that has been often overlooked in variationist linguistics.

The variable presence of post-nasal [ɡ] in words such as sing [sɪŋɡ]~[sɪŋ] and wrong [ɹɒŋɡ]~[ɹɒŋ] is a characteristic feature of the varieties of English spoken in the North West and West Midlands of England1. This phenomenon, which Wells (1982b: 365) refers to as ‘velar nasal plus’, has been documented in a number of dialectological handbooks (e.g. Wakelin 1984; Trudgill 1999; Hughes et al. 2012) but has scarcely been investigated under the variationist paradigm. As a result of this, while its diachronic pathway of change has been explored in detail (see Bermúdez-Otero 2011; Bermúdez-Otero & Trousdale 2012), synchronic patterns of variation in velar nasal plus remain comparatively understudied.

This paper provides evidence that variation in [ŋɡ] clusters, hereafter denoted by (ng) using standard sociolinguistic convention, is less stable than previously thought; specifically, the behaviour of (ng) in pre-pausal position appears to be undergoing change in apparent time, whereby younger speakers are reanalysing this environment as one that favours [ɡ]-presence. The primary goal of this paper is to investigate the mechanisms underlying this innovation, specifically to disentangle the collinearity between three factors that on the surface appear to condition this effect: segmental duration, prosodic boundary strength, and the presence/duration of a following pause. In doing so, this study adds to a growing body of evidence outlining how probabilistic lenition processes behave before phrasal boundaries, and its results have

1Whilst post-nasal [ɡ] also occurs in unstressed -ing clusters, leading to surface three-way variation between [ɪn]~[ɪŋ]~[ɪŋɡ], in this paper the focus is solely on the stressed (ng) clusters that are invariable in non-northern varieties.
implications for ongoing research into the conditioning factors of external sandhi.

Drawing upon production data from an elicitation task, it is shown that the probability of surface [ɡ]-presence is most strongly correlated with the duration of pause that follows it, independent of the word’s position in the utterance or intonational phrase. The presence of a following pause is also highly collinear with the duration of the preceding nasal due to the effects of pre-boundary segmental lengthening, but the former is a much stronger predictor of the variation in [ɡ]-presence. Thus, velar nasal plus in northern English dialects shows no evidence of direct reference to segmental duration (cf. Lavoie 2001), and there is only weak evidence of sensitivity to phrasal prosodic categories (cf. Nespor & Vogel 1986); rather, the results of this study emphasise the importance of the temporal relationship between the target and trigger in external sandhi processes.

The structure of this paper is as follows: Section 2.1 introduces velar nasal plus and outlines the current body of knowledge regarding how its patterns of variation are structured along social and language-internal dimensions; Section 2.2 provides a summary of the literature on how pausal boundaries affect other probabilistic external sandhi processes, and highlights a number of ways in which the conditioning of external sandhi has been accounted for in phonological theory; the discussion of pre-boundary lengthening in Section 2.3 foregrounds the collinearity issues explored in this paper, whose research goals are then re-stated in Section 2.4.

The methodology undertaken for this study is outlined in Section 3, detailing the methods of data collection and in particular how the elicitations were carefully designed in order to invoke different magnitudes of pre-boundary lengthening. The results of this study are split into two subsections: Section 4.1 presents evidence from sociolinguistic interviews of a change in apparent time with respect to the rates of [ɡ]-presence pre-pausally, and Section 4.2 addresses the primary goal of this paper by exploring how this innovation is represented in speakers’ grammars through analysis of an elicited reading task. Although the focus of this paper is to uncover the precise mechanisms that condition this innovation, discussed in Section 5.1, part of the discussion is also dedicated to addressing the social and/or internal factors that actually motivate this change; in Section 5.2, a number of possibilities are proposed, specifically whether this diachronic change reflects a shift in the social meaning and evaluation of the local form, or stems from the inherent variability of external sandhi processes compared with word-internal phenomena.
2 Background

2.1 Velar nasal plus

It should be pointed out that post-nasal [ɡ] was once present across all varieties of English, before it began to undergo deletion in the Late Modern English period. Drawing upon reports by eighteenth-century orthoepist James Elphinston (discussed in Garrett & Blevins 2009), Bermúdez-Otero & Trousdale (2012) provide a particularly enlightening account of this change. They show how the phonological /ɡ/-deletion rule progressed through the grammar such that in varieties of Present Day English, [ŋɡ] clusters are only ever present pre-vocally in monomorphemic or root-based items such as finger or elongate, in addition to a small set of lexically-listed exceptions (namely, the comparative and superlative forms of strong, long, and young).

Although this coda-targeting deletion rule ran to completion in most varieties of English, the non-coalesced [ŋɡ] form was not lost everywhere: variation in (ng) still exists today in these varieties spoken in the North West and West Midlands of England. Although we know very little about the synchronic variation of (ng) in these communities, the presence of post-nasal [ɡ] is well-documented in the dialectological literature (e.g. Wakelin 1984; Trudgill 1999; Hughes et al. 2012)). It has been documented in Birmingham (Thorne 2003), Cannock (Heath 1980), Liverpool (Knowles 1973), West Wirral (Newbrook 1999), Manchester (Schleef et al. 2015; Bailey 2015), and in Sandwell and the surrounding Black Country (Mathisen 1999; Asprey 2015). These areas all fall within the North West or West Midlands of England, corresponding with the Survey of English Dialects isogloss (Orton et al. 1978) as well as more recent dialectological surveys (MacKenzie et al. 2017). However, these studies do not go beyond pointing out the presence of this form, and many in fact do not acknowledge that its presence is variable in those communities in which it is attested, let alone explore the factors that condition such variation. With many of them also relying on impressionistic and auditory analysis, variation in (ng) has simply not been subject to the same sociophonetic scrutiny as many other probabilistic processes in Present Day English.

While variation in (ng) does historically stem from a deletion rule, it is possible that at this point in time the synchronic system does not work that way, and that some tokens of post-nasal [ɡ] surface instead from an insertion process. Determining whether or not this is the case is beyond the scope of this paper, and as such the subsequent discussion of (ng) variation will remain theory neutral, referring only to presence or absence of [ɡ] and not to the process assumed to underpin this variation.
The observation that [ɡ]-presence is favoured before pause, with which this paper is primarily concerned, has not been discussed explicitly in other studies. However, the observation that (ng) shows strong stylistic stratification could provide supportive evidence for this effect; both Mathisen (1999) and Bailey (2015) report high rates of [ɡ]-presence in word-list elicitations. The conventional and most immediate interpretation of this is of course that [ɡ]-presence is considered the ‘prestige’ form and that this style-shifting simply reflects adherence to this norm in more conscious speech styles. However, it should be noted that these word-list elicitation tasks conflate two things: formality, and phonological environment. In other words, do we find more [ɡ]-presence in word-list elicitations because this form is considered the standard and is therefore more frequent in formal discourse styles, or is it actually because in this style the tokens of (ng) are elicited with clear pauses and prosodic breaks between each item? It is of course possible that the high rate of word-list [ɡ]-presence is in fact attributable to both. The former explanation presupposes that forms with the post-nasal [ɡ] are indeed considered prestigious, but the only study to investigate the evaluation of [ŋɡ] shows no evidence that this is the case (Newbrook 1999).

A number of studies seem to suggest that the local non-coalesced form, in which post-nasal [ɡ] is present, is increasing in popularity with younger speakers, though few actually provide quantitative evidence in support of such claims. Asprey (2007: 90) reports that the presence of [ɡ] is “linked to the younger generations” in the Black Country, and this association between [ŋɡ] and youth speech is echoed by others (see Wakelin 1984; Chinn & Thorne 2001). Mathisen’s (1999) work in Sandwell in the West Midlands does, however, provide an empirical grounding to such claims; this increase, described as a “revitalisation” of this local form, is being led by young women and the working classes in particular. A preference for velar nasal plus among the working classes is corroborated by Thorne (2003: 121), and an increase in its use in apparent time is also found in the speech community of Wilmslow, Cheshire (Watts 2005: 173).

2.2 Boundary effects on other external sandhi processes

Since very little work has been carried out on the language-internal factors influencing (ng), specifically its sensitivity to phonological and prosodic environment and its behaviour pre-pausally, we can instead turn to comparable external sandhi processes that have been subject to more extensive variationist study. One such example is /t,d/-deletion in varieties of English: the reduction of word-final consonant clusters ending with a coronal stop e.g. just [dʒʌst]~[dʒʌs], proved [pɹuːvd]~[pɹuːv]. This is remarkably well-studied, having been attested across the world’s varieties of English, and its
variation shows similar patterning to (ng). Both involve segmental presence/absence in word-final consonant clusters, and both are sensitive to morphological and syntactic structure in ways consistent with a cyclic analysis. Guy (1991b) adopts a Lexical Phonology framework in accounting for the morphological effect on /t,d/-deletion, whereby deletion is less likely for past tense items due to the targeted /t,d/ segment appearing later in the derivation, while diachronic and synchronic accounts of (ng) have been combined under the life cycle of phonological processes (see Bermúdez-Otero & Trousdale 2012 on the diachronic process of /ɡ/-deletion, and Paper I of this thesis on the synchronic implications that follow from this analysis).

Most importantly for the present study, both processes show sensitivity to the immediate phonological context and do so in an identical manner: [ɡ]-presence is more likely pre-vocally than pre-consonantally (Knowles 1973; Upton et al. 1987; Watts 2005), and the same pattern of variation has been shown for /t,d/-deletion in countless studies (e.g. Tamminga 2016 on Philadelphia English, Baranowski & Turton forthcoming on Manchester English). In fact, Tagliamonte & Temple (2005) claim that this is consistently the strongest predictor of /t,d/-deletion in all varieties of English in which it has been studied. What is not so consistent, however, is how /t,d/-deletion behaves pre-pausally. In some varieties, following pauses are said to inhibit deletion even more so than following vowels – e.g. York (Tagliamonte & Temple 2005), Philadelphia (Guy 1980), and Chicano English (Santa Ana 1996) – while in others the deletion rate pre-pausally is higher (see Bayley 1994 on Tejano English and Hazen 2011 on Appalachian English). For some speakers, particularly those of African American Vernacular English, deletion in pre-pausal environments can be even as high, and sometimes higher, than in pre-consonantal position (see e.g. Fasold 1972 in Washington D.C. and Guy 1980 in New York).

More recent studies have done away with a categorical coding of pause presence/absence altogether, and instead incorporated pause duration as a gradient factor. Tanner et al. (2017) show how pause duration, used as a proxy of boundary strength, modulates the effect of following segments such that the influence of a following vowel or consonant on the application of /t,d/-deletion is neutralised when a long pause (100ms or greater) separates them from the preceding /t,d/ cluster. They argue that this behaviour lends empirical support to the production planning hypothesis (Wagner 2012a): the stronger the prosodic or syntactic boundary between constituents, the less likely it is that the following segmental material has been planned, and as such it can have no effect on the realisation of the preceding coronal stop. This has also been recently explored by Tamminga (2018), who finds a similar interaction between the magnitude of the following segment effect and the strength of the
syntactic juncture between the target and trigger.

Formal accounts of external sandhi conditioning, specifically the mechanisms that trigger this sensitivity to phonological environment, have also often focused on /t,d/-deletion. A number of explanations of the ‘following segment effect’ have been proposed, with the goal of capturing not just the consistent observation that deletion is more likely pre-consonantally than pre-vocalically, but also the variability of deletion pre-pausally. It has been argued (see e.g. Guy 1980) that this effect stems from the possibility of phrase-level resyllabification, where word-final pre-vocalic /t,d/ variably attaches as an onset to the following word and thus avoids deletion; however, this has been disputed on the grounds that the phonetic realisation of word-final pre-vocalic /t,d/ when present on the surface is not comparable to that of a canonical word-initial /t,d/ even though the former is argued to be in onset position (Labov 1997).

Alternative explanations make reference to the Obligatory Contour Principle (McCarthy 1986; Yip 1988) by highlighting differences in feature similarity between the /t,d/ and a following consonant, liquid, glide, or vowel. Crucially, the inter-dialectal variation in how /t,d/-deletion behaves pre-pausally stems from the fact that the pre-pausal environment by its very nature does not fit into the above hierarchy and is therefore “susceptible to differing analyses by different speakers or dialects” (Guy 1980: 27).

Coetzee (2004) offers yet another proposal, instead relying on licensing by cue (Steriade 1997) and how these perceptual cues for identifying the place and manner of articulation of stops (namely, their release and also the formant transitions into a following vowel) are realised before consonants, vowels, and pauses. Such an account simply has to stipulate inter-dialectal differences in the ranking of constraints, or alternatively in the phonetic realisation of pre-pausal consonants, to capture the difference between varieties in how pre-pausal /t,d/ behaves. Whatever the nature of this sensitivity to the phonological environment, there is ample evidence to suggest that the effect of a following pause on /t,d/-deletion is open to differing analyses between speech communities.

We also find similar inter-dialectal variation in /s/-weakening across varieties of Spanish. This rule is yet another example of a coda-targeting lenition process, in this case debuccalisation from /s/ to [h], where the effect of a following pause is not universal. In standard varieties of Argentinian Spanish, /s/-weakening is blocked pre-pausally (see Kaisse 1996), contrasting with Caribbean varieties where weakening shows no such sensitivity to pause (see Harris 1983).

These processes are uncontroversially leniting and therefore any comparison
with (ng), which as discussed earlier could conceivably be a case of synchronic [g]-insertion, should be taken with some degree of caution. However, it remains the case that there are clear parallels between these three processes with respect to their phrase-level conditioning: the varying segments ([ɡ], [t]/[d], and [s]) are present before vowel-initial words, absent before consonant-initial words, and show unusual and inconsistent behaviour pre-pausally. In the case of /t,d/-deletion and /s/-weakening, this registers itself as differences in pausal effects between different varieties, and in the case of (ng) as diachronic instability, as will be illustrated in Section 4.1.

2.3 Pre-boundary lengthening

One often over-looked aspect of how pauses influence variable linguistic phenomena is the way in which they affect suprasegmental features, specifically phonetic duration. It has been observed cross-linguistically that segments in pre-boundary position are longer in duration than those not adjacent to a prosodic boundary. Many reports focus on Indo-European languages (see Lehiste et al. 1976 on English; Lindblom 1968 on Swedish; Delattre 1966 on French, Spanish and German), but Hockey & Fagyal (1998) also report it for Hungarian of the Finno-Ugric family, despite this language having phonemic length distinctions.

It is generally considered that pre-boundary lengthening is triggered directly by finality in a prosodic constituent, with the magnitude of lengthening correlated with the size of the constituent in the phonological hierarchy (Gussenhoven & Rietveld 1992; Wightman et al. 1992). However, this has recently been disputed by Feldscher & Durvasula (2017), who instead propose that lengthening is triggered directly by pause. There is also evidence indicating language-specific implementations of lengthening, possibly influenced by the role of duration in other areas of the grammar, which suggests that the magnitude of pre-boundary lengthening is sensitive to factors other than just the prosodic hierarchy (Turk 2012; Cho 2016). The exact typology of constituents within this hierarchy is also subject to debate, but there is widespread agreement on the ‘major’ categories above the word-level, as well as their relative ordering: the Utterance (U) is higher than the Intonational Phrase (IP), which in turn is higher than the Phonological Phrase (PPH) (Selkirk 1978; Gussenhoven 2002).

Given that stronger boundaries elicit longer pauses and greater segmental lengthening, the collinearity between these three factors raises questions regarding the nature of these reported ‘pre-pausal’ effects. What if the effects of a following pause sometimes reflect something more granular, i.e. sensitivity to duration? This was explored by Sproat & Fujimura (1993) in their study of /l/-darkening: they argue that
contrary to earlier claims, /l/-darkening is gradient in nature and triggered by a purely durational mechanism in which the darkness of the /l/ is positively correlated with the duration of the rime. Although more recently it has been shown that this is an oversimplification, with ultrasound tongue imaging revealing both a categorical and gradient process of darkening (see Turton 2014, 2017), it is nevertheless the case that the gradient process of darkening is correlated with duration.

Much like with the parallels drawn between (ng) and (td) in the preceding section, I do not mean to suggest that (ng) and /l/-darkening are comparable variables; this is particularly important in the case of /l/ given that it is not only uncontroversially leniting but also of a non-deleting type. However, regardless of the similarities and differences between these processes, it is possible that the same durational mechanism applies in the case of (ng), even if it is interpreted as insertion. Given that such an insertion process only requires a slight change in gestural timing, where the velum is raised before rather than after cessation of the lingual gesture, it would hardly be surprising to find it showing sensitivity to the preceding nasal duration.

2.4 Research questions

In light of the current knowledge summarised in the previous section, this study aims to accomplish two things: to provide evidence of a hitherto-unreported change in progress towards increasing [ɡ]-presence among young speakers, restricted to prepausal contexts, and then to solve the collinearity between various boundary phenomena by investigating three related prosodic factors that potentially condition this change.

Solving this collinearity issue has wide implications for a number of theoretical approaches that make predictions with respect to what should condition such an effect. A classical Prosodic Phonology approach to external sandhi processes (Nespor & Vogel 1986) predicts that the conditioning environment is defined by the categories of the prosodic hierarchy itself, e.g. when final in the intonational phrase (IP) or utterance (U). If this were the case, we should find a stark contrast in [ɡ]-presence between domain-medial and domain-final positions, and a result in which this dichotomy provides the best fit to the observed variation would lend support to this theory.

On the other hand, there are theories of lenition (e.g. Lavoie 2001) that highlight the importance of durational factors over the abstract categories proposed by Prosodic Phonology. These accounts claim that the primary phonetic manifestation of weakening is shorter segmental duration; as such, they would predict that it is phonetic duration that directly influences (ng) variation, such that the probability of [ɡ]-presence is correlated with the duration of the syllable coda or rime.
Finally, recent years have seen a rise in theories that emphasise the psycholinguistic processing of language, such as the Production Planning Hypothesis (Wagner 2012a). Under these accounts, the most important factor in conditioning external sandhi processes is the temporal relationship between target (in this case, word-final [ŋg]), and trigger (the following consonant-initial word), thus motivating the inclusion of intervening pause duration in this analysis. It may be the case that this plays the biggest role in conditioning this case of external sandhi, whereby the presence (and possibly duration) of a pause following the (ng) token has a direct impact upon the probability that the [ŋ] is present on the surface, independent of prosodic position or phonetic duration.

It is important to consider these three factors separately at both the conceptual and empirical level, despite the strong collinear relationship between them. Although it has been shown that pause is the most important acoustic cue to the perception of intonational phrasing (see Swerts 1997; Mo 2010; Zhang 2012; Yang et al. 2014), it is not mandatory to mark a boundary with pause (Krivokapić & Byrd 2012), which results in cases where this collinearity breaks down. In showing how the effects of pause interact with prosodic position in Japanese high-vowel devoicing, Kilbourn-Ceron (2017) highlights the importance of teasing apart such factors for variables that show apparent ‘pre-pausal’ behaviour, and there is evidence from Argentinian Spanish that (a) IP boundaries can be produced without pause and (b) pauses can occur within IPs (Kaisse 1996).

This paper seeks to uncover the relative contributions made by these three boundary phenomena (prosodic boundary strength, segmental duration, and pause presence/duration) in boosting [ŋg]-presence in northern British English, and by doing so will contribute to our knowledge of how external sandhi processes operate in pre-boundary position.

3 Methodology

This study takes a two-pronged approach in answering the research questions posed in the preceding section, and it does so by drawing upon two complementary sources of data: a collection of sociolinguistic interviews, and a follow-up elicitation task with a similar population sample. These two methods of data collection have complementary strengths and weaknesses, and an analysis that makes use of both is therefore better-equipped to provide an accurate description of the variation. The interviews contain naturalistic language that is often the subject of variationist analysis; these will be used to provide empirical evidence of a change in progress among North West-
ern subjects. The elicitions, on the other hand, allow for careful control over the environments in which the dependent variable appears, which is a crucial component of this investigation into the behaviour of (ng) at various linguistic boundaries; these will be used to provide insight into factors that condition the afore-mentioned change.

All recordings were made using a Sony PCM-M10 recorder and a lavalier microphone attached to the participant, saved at a 44.1KHz sampling rate in uncompressed WAV format.

3.1 Sociolinguistic interviews

The naturalistic component of the data consists of 32 sociolinguistic interviews, most of which took place during a two-year period from 2015 to 2017. In the following subsections I provide detail on the demographics of the participants and the interview process itself.

3.1.1 Participants

The population sample consists of 32 speakers (17F; 15M), all of whom were born in the North West of England with a large majority born and raised in Greater Manchester. Their date of births range from 1907 to 1998, with two interviews conducted in 1971 included to provide extra time depth to the apparent time analysis of (ng). Socioeconomic status was controlled for by only interviewing upper working class speakers, with this classification based broadly on occupation following the methods of Baranowski (2017: 303) in Manchester.

3.1.2 Task

The sociolinguistic interviews were conducted one-on-one with the participants. Following guidance from Tagliamonte (2006: 37-49), the interviews typically lasted for approximately one hour and took the recommended form of “hierarchically-structured” topical modules such as childhood, work and family life, the local community, travel etc. (Labov 1984: 33). These topics consisted of open-ended questions, many of which were designed to elicit narratives of personal experience, said to provide the clearest access to a speaker’s vernacular and minimise the effects of the observer’s paradox (Labov 2010a). In total, these interviews yield 1526 tokens of (ng).

These interviews were conducted as part of a wider variationist investigation of (ng), and as such they were also coded for a number of linguistic factors such as the immediate preceding/following phonological context, word frequency, speech rate, and part of speech, among others. Tokens were coded as being ‘pre-pausal’ when final
<table>
<thead>
<tr>
<th>Age</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young</td>
<td>24 yrs</td>
<td>24 yrs</td>
</tr>
<tr>
<td></td>
<td>N=8</td>
<td>N=8</td>
</tr>
<tr>
<td>Old</td>
<td>64 yrs</td>
<td>60 yrs</td>
</tr>
<tr>
<td></td>
<td>N=6</td>
<td>N=8</td>
</tr>
</tbody>
</table>

Table 1: The age and sex distribution of informants for the elicitation task. Cells include the average age of each group, with N denoting number of subjects. 'Young' speakers are aged between 18 and 30; 'old' speakers are aged between 52 and 85.

in an ELAN breath group; broadly speaking these tokens are followed by a period of silence lasting approximately 100ms or longer.

3.2 Elicitations

While the conversational data provides a reliable insight into how (ng) varies in naturalistic speech, an elicitation task is required to tease apart the various collinear boundary phenomena that potentially condition this variation. For this elicitation task, subjects were asked to read out a list of sentences as naturally as possible and at their own pace. Each sentence contained exactly one token of (ng) before a particular linguistic boundary, detailed in Section 3.2.2, and they were presented one at a time on a laptop screen. In the following sub-sections I provide further information about the sample population and the design of elicitation stimuli.

3.2.1 Participants

The elicitation task was conducted with 30 speakers from the North West of England, many of whom were also subjects of the sociolinguistic interviews detailed in Section 3.1. These 30 speakers form a balanced population sample with respect to age and sex (see Table 1), and although many of the informants were born and raised in Manchester, the sample contains a number of speakers from other regions in the North West such as Blackburn, Widnes, Wigan, and Bolton. Efforts were made to ensure that all subjects who took part in this elicitation task were not only born and raised in the North West of England but also that they had at least one parent who was also a native British English speaker from the same region.

3.2.2 Stimuli

Five linguistic boundaries of different perceived ‘strengths’ are under consideration here, based primarily on the stimuli chosen by Sproat & Fujimura (1993) in their com-
parable study of coda-targeting /l/-darkening in English\(^2\). The aim of these carefully-controlled environments, which vary in their inherent ‘strengths’, is to elicit different magnitudes of pre-boundary segmental lengthening. Although later work suggests that the magnitude and implementation of pre-boundary lengthening is not universal and not solely a function of prosodic or syntactic boundary strength (Turk 2012; Cho 2016; Feldscher & Durvasula 2017), the results from Sproat & Fujimura (1993) nevertheless justify the adoption of similar stimuli here. It has already been shown that these elicitations result in a range of segmental durations, which will allow for an investigation into how well this phonetic property correlates with [ɡ]-presence. These boundaries, which are either syntactic or prosodic in nature, are detailed and exemplified below in increasing order of the strength of the juncture:

1. **NP-internal boundary**: Immediately followed by the head of an NP
   
e.g. *She was given [the wrong amount]\(_{NP}\)*

2. **VP-internal boundary**: Immediately followed by the direct object in a double object construction
   
e.g. *She gave [the ring]\(_{DO}\) [a quick polish]\(_{DO}\)*

3. **VP boundary**: Immediately followed by an NP/VP juncture
   
e.g. *[The sting]\(_{NP}\) became painful*

4. **Intonational phrase boundary**: Final in the intonational phrase
   
e.g. *[“It’s a traditional thing.”]\(_{IP}\) Patricia said*

5. **Utterance boundary**: Final in the utterance
   
e.g. *[The drink was surprisingly strong]\(_{U}\)*

Because the words under consideration contain two sonorous segments upon which the effects of lengthening can be registered (see Turk & Sawusch 1997; Turk & White 1999; Turk & Shattuck-Hufnagel 2007)), pre-boundary lengthening was operationalised as ‘sonorant duration’, encompassing both the vowel and nasal portion of

\(^2\)Word-medial tokens were also collected, but had to be discarded from the analysis due to a confound of epenthesis; the (ng) tokens before consonant-initial suffixes (e.g. *gangster*, *youngster*) resulted in nasal-sibilant clusters that are known to trigger excrecent stop insertion (Fourakis & Port 1986; Warner 2002). This is of course a productive and widely-attested phenomenon in English, sometimes referred to as the *prints-prince* merger, where gestural timing during the nasal+sibilant transition can lead to insertion of a stop that is homorganic with the preceding nasal, resulting in surface forms such as *[ˈpɹɪnts] ‘prince’, *[ˈhæmp.stə] ‘hamster’, and *[ˈjʊŋk.stə] ‘youngster’. Owing to the difficulty in determining whether a post-nasal stop in this environment arises through this process rather than being a genuine case of velar nasal plus, these examples had to be excluded.
the (ng) word. The phonetically-gradient relationship between this durational mea-
sure and the five boundary contexts included in this study is illustrated in Figure 1, highlighting the success of the chosen stimuli in eliciting various magnitudes of lengthening.

Each boundary context was represented eight times in the sentence list, equally
distributed by phonological context such that each boundary × preceding segment × following segment combination was represented by two example sentences. The segment immediately following the (ng) cluster was either a consonant or a vowel, with the following word also having non-initial stress, and the segment immediately preceding the (ng) cluster was either a low vowel or a high vowel. Controlling for vowel height is necessary because it presents a confound for our quantification of pre-boundary lengthening effects. Given that pre-boundary lengthening applies not just to the nasal segment of the (ng) word but also to the preceding stressed vowel, it is important to minimise the possibility of any other factors influencing the durational properties of these segments. The well-established correlation between vowel height and duration (see Lehiste 1970; Tauberer & Evanini 2009; Solé & Ohala 2010) is one such confound. Word token frequency is another potential confounding factor, and one that is less easily overcome given the small set of lemmas that actually contain a variable (ng) cluster. The impact of token frequency on phonetic implementation has been subject to extensive study, and one such surface manifestation of token frequency is registered in segmental duration, whereby less frequent words are often

\[\text{Figure 1: The impact of boundary strength on the sonorant (vowel+nasal) duration of (ng) tokens.}\]

\[\text{For the word-final (ng) tokens, the following consonant was almost always a non-lingual obstruent to prevent possible confounds of assimilation or phrase-level resyllabification of the word-final [ɡ]. The only exception to this is in the case of one IP-boundary elicitation, where the (ng) token is followed by a non-lingual sonorant /m/.}\]
longer than words that are frequent and more predictable in discourse (see Jurafsky et al. 2001; Aylett & Turk 2004). To minimise the impact of this confounding factor, efforts were made to avoid highly infrequent (ng) lemmas; with just one exception, all lemmas used in the stimuli range between 4.25-5.94 on the logarithmic 1-7 Zipf-scale (van Heuven et al. 2014).

In total, 40 sentences were elicited per participant (5 boundaries × 2 preceding segments × 2 following segments × 2 repetitions), yielding a total of 1,200 tokens; these sentences are given in full in the Appendix.

3.3 Data annotation

The recordings were all transcribed orthographically using ELAN and force-aligned with the FAVE suite (Rosenfelder et al. 2011) to facilitate a more efficient analysis. Forced alignment is a major methodological innovation in contemporary variationist linguistics, in which an audio file is time-aligned at the word- and phoneme-level with a corresponding orthographic transcription. Although recent work has probed the ability of forced alignment to also automatically code for linguistic variation (e.g. Yuan & Liberman 2011; Bailey 2016), a manual method of coding was employed here. Coding of the dependent variable was carried out using a combination of auditory analysis and visual inspection of the spectrogram in Praat (Boersma & Weenink 2017). For ambiguous tokens where the presence/absence of [ɡ] was not clear (approximately 3% of the entire sample), a second round of coding was carried out independently by another phonetically-trained researcher, and any tokens for which there was disagreement were subject to further inspection. These cases were extremely rare, and there is in fact relatively little variation with respect to the phonetic realisation of post-nasal [ɡ], which is almost always released and very often devoiced in phrase-final position. In light of this, a binary coding scheme was used based on the categorical presence/absence of a post-nasal stop. Prototypical examples of a [ɡ]-ful token and a [ɡ]-less token are given in Figure 2.

Although the stimuli were designed to control intonational phrasing, with boundaries 1-3 intended to elicit IP-medial tokens and boundaries 4-5 IP-final tokens, this was independently clarified through intonational analysis. Pitch contours consisting of 64,644 dynamic pitch measurements were extracted, manually corrected, and smoothed using the 
mausmooth
Praat script (Cangemi 2015). The elicited sentences were then annotated by the author in the ToBI framework (Beckman et al. 2005) for nuclear accent placement and presence of phrase accent and boundary tones, the latter providing a more reliable annotation of intonational phrasing. This manual
annotation is of paramount importance for tokens where the phonetic cues to intonational phrasing are only partially present: tokens of (ng) produced in the IP-medial context that are followed by a pause – and conversely tokens in the IP-final context that are not – are crucial to the analysis and in these cases the manual annotations were compared with those of another researcher trained in phonetics to ensure the reliability of the coding.

4 Results

The results of this study will be presented in two complementary sub-sections: The first part of the analysis draws upon sociolinguistic interview data to provide evidence of change in apparent time (Section 4.1). In the second part of this analysis (Section 4.2), attention turns to the follow-up elicitation task with the goal of probing the precise factors that condition this innovation.

All logistic regression models reported in this section were fit using the `glmer` function in the `lme4` R package (Bates et al. 2015). All models include random intercepts of speaker and word.

4.1 Change in apparent time

Although there have already been reports that rates of [g]-presence are increasing in a number of communities, summarised earlier in Section 2.1, the interaction between age (or date of birth) and phonological environment with respect to [g]-presence has yet to be investigated. Figure 3 plots the pre-consonantal and pre-pausal rates of velar nasal plus by date of birth for this sample of speakers from the North West of England, where ‘pre-pausal’ refers to tokens followed by a period of silence lasting around 100ms or longer. The results indicate that the increase in [g]-presence over
the 91-year time span covered by this sample of speakers is largely confined to the pre-pausal environment.

It should be noted that there also seems to be a slight increase in [ɡ]-presence pre-consonantally, but this trend is much less dramatic and the correlation is not statistically significant (Spearman’s $r_s = 0.23, p = 0.21$); the trend in pre-pausal environments, however, is strong and highly significant ($r_s = 0.70, p < 0.001$). The favourable effect of a following pause on the probability of [ɡ]-presence is particularly evident for speakers born after 1975, many of whom show categorical use of the local form in this particular environment.

This apparent diachronic change affecting pre-pausal velar nasal plus finds statistical support from the results of mixed-effects logistic regression, where the interaction between phonological environment and date of birth is significant for the pre-pausal tokens but does not pass the threshold for significance pre-consonantally (see Table 2). Furthermore, conducting an ANOVA comparison between nested models confirms that adding an interaction with date of birth leads to a statistically-significant decrease in AIC (935.31, cf. 958.29; $p < 0.001$) and, therefore, a better fitting model.

Although not plotted in Figure 3, there is also no evidence for a change in progress involving the pre-vocalic environment ($r_s = 0.04, p = 0.73$).
Table 2: Mixed-effects logistic regression model for the interaction between following segment and date of birth; includes random intercepts of speaker and word. [ɡ]-presence as application value. Vowel as reference group.

### 4.2 Elicitation task

Although there are a number of benefits to analysing the conversational data discussed in the previous section, most notably the fact that this is a naturalistic speech style and therefore more representative of the speakers’ vernaculars, it is not without fault. One particular limitation is that a dichotomy between whether or not a token of (ng) is followed by a pause actually conlates a number of prosodic environments and interactional situations; in reality, these pre-pausal tokens may encompass a wide range of contexts, e.g. turn-final, utterance-final, IP-final etc. Is there an absence of segmental material following the (ng) token because the speaker was interrupted, or was the pause just temporary, with the speaker then resuming their turn? Pauses may arise for a number of different reasons, whether they be cognitively or interactionally motivated (see Kendall 2013 for an exploration of the factors that condition pause production).

To combat this shortcoming, the second part of this analysis focuses on the follow-up elicitation task where the exact environments in which (ng) clusters appear can be carefully controlled using reading passage stimuli. In this way, efforts can be made to disentangle the collinearity between the three factors that on the surface appear to be boosting [ɡ]-presence: phonetic duration, prosodic boundary strength, and pause presence/duration.

Figure 4 illustrates an interaction between boundary strength and following segment with respect to rates of [ɡ]-presence. When the (ng) cluster is pre-vocalic, rates of [ɡ]-presence remain high irrespective of the type of boundary; however, in pre-consonantal position the variation clearly shows sensitivity to boundary strength in a
Figure 4: Rate of post-nasal [ɡ]-presence by boundary strength and the type of following segment.

much more striking manner, such that the rate of [ɡ]-presence in the weakest boundary context is as low as 4%.

The relative lack of variation pre-vocally is no great surprise, and is likely due to the fact that there are two competing forces promoting the presence of [ɡ] in this environment: one that favours [ɡ]-presence before stronger boundaries, as we can see with the pre-consonantal tokens, but also one that favours [ɡ]-presence in weaker boundary contexts; crucially, this latter effect is confined to the pre-vocalic environment. If we assume that variation in (ng) is derived from a coda-targeting deletion rule, and that the promoting effect of following vowels on the probability of [ɡ]-presence stems from phrase-level resyllabification of the [ɡ] into onset position, it logically follows that this effect is more likely when the juncture between the words is weaker. Because weaker boundaries favour resyllabification, they consequently also favour [ɡ]-presence, but crucially this applies only in pre-vocalic environments. These two antagonistic effects cancel each other out pre-vocally, where rates of [ɡ]-presence are high across the board, whereas in pre-consonantal environments only the former, more general effect is present. Given then that (ng) only shows sensitivity to boundary strength in pre-consonantal environments, the subsequent analysis will focus on this subset of the data, discarding the pre-vocalic tokens that are largely invariable.

In the pre-consonantal environment, we do clearly see a monotonic increase in [ɡ]-presence correlated with the strength of the following juncture. However, the

5The likelihood of phrase-level resyllabification being the source of this phonological environment effect is further increased by the independent observation that resyllabification across word boundaries is more likely when the following word, to which the [ɡ] is attaching, has non-initial stress (Bermúdez-Otero & Trousdale 2012: 697); recall that in this study, all word-final tokens of (ng) are elicited before words with non-initial stress.
rates of [g] do not increase in a gradual manner parallel to the phonetically-gradient relationship between boundary strength and segmental duration; instead, we see a stark contrast between boundaries 1–3 and 4–5, suggesting that the meaningful contrast is between IP-medial and IP-final position. However, sensitivity to IP boundaries alone would not account for the contrasting behaviour of (ng) clusters between boundaries 4 and 5; in the former (utterance-medial IP boundary) we see 53% [g]-presence, whereas in the latter (utterance-final IP boundary) we find rates almost at ceiling level (96%).

Pause presence/duration provides a possible explanation for this contrast, in addition to showing the strongest correlation with the presence of post-nasal [g]. The use of [ŋɡ] is more variable at the utterance-medial IP boundary (i.e. boundary 4) because here the prosodic phrasing and, in particular, presence of pause is also variable (cf. the utterance-final IP boundary which is always pre-pausal). This is shown in Figure 5, which illustrates the relationship between pause duration, pre-boundary lengthening, prosodic position, and the realisation of (ng).

What is perhaps most interesting to note from Figure 5 is that there is much clearer separation along the x-axis than along the y-axis with respect to [g]-presence. In other words, following pause duration is a much stronger predictor than sonorant duration: we can draw a cut-off point around 4.6 (~100ms) on the x-axis, such that any pause longer than this is enough to result in [g]-presence almost without fail. Reassuringly, this is the same value as the cut-off point used to identify pre-pausal tokens in the conversational data when establishing the change in progress, as de-
scribed in Section 4.1.

To investigate the effects of pause and IP position independently, there need to exist tokens of (ng) where the pausal cue to major prosodic boundaries is absent on the surface. Figure 5 highlights an overlap between the IP-medial and IP-final tokens with respect to the following pause duration, suggesting that this is the case. In total, 65 of 120 tokens in the IP-final context surface without a pause. However, it is entirely possible that the intonational phrasing this stimuli intended to elicit was not actually produced, and that these cases of IP-final tokens without pause are in fact medial in the IP. To combat this, we need independent evidence, specifically based on the pitch contour, of the presence of IP boundaries. Intonational analysis, combining visual inspection of the pitch contours with manual ToBI annotation, reveals that none of these 65 tokens show convincing evidence of a prosodic boundary after the (ng) word.

However, the ability to tease apart these two collinear effects does not rest solely on the presence of such tokens, where IP boundaries are not marked by pause. If there are cases of IP-medial tokens that are followed by pause, and crucially exhibit [ɡ]-presence, this would provide strong evidence that the variation is conditioned most strongly by pause rather than the presence of a prosodic boundary. 25 of the 360 tokens (~7%) elicited in the IP-medial context are produced in such a way. Based on the intonational analysis, 14 of these 25 (56%) are genuinely in IP-medial position with no evidence of pitch reset or boundary tone, i.e. there is a brief juncture before resumption of the same pitch movement. An example of this is given in Figure 6a. It is also important to note that in this example, the hiatus in the pitch contour does not reflect devoicing of the /ɡ, b/ sequence in Spring began but rather a genuine period of silence, as shown in the waveform and spectrogram. A counter-example is presented alongside this in Figure 6b, where the pause is clearly a phonetic cue to an IP boundary tonally marked with a fall-rise phrase accent and boundary tone combination. Crucially, post-nasal [ɡ] is present in all of these 14 genuine cases where (ng) occurs before an IP-medial pause.

Mixed-effects logistic regression lends further support to the idea that the presence and duration of pause is the primary conditioning factor of (ng). Three individual models were initially fit, including a main predictor of either: (1) sonorant duration, (2) IP position (based on what was actually produced, rather than what was intended by the stimulus), or (3) following pause duration, in addition to random intercepts of speaker and word. These models were compared for ‘goodness of fit’ based on their AIC values to determine which predictor explains most of the variation in (ng), where lower values correspond to a better model. These comparisons, summarised in Table 3, indicate that sonorant duration explains the least amount of variation (AIC:
They promised the King possession of their land.

The Spring began with a rainy day.

Figure 6: Pitch contours and ToBI annotation for two pre-pausal tokens in the IP-medial context: one genuinely IP-medial in (a) and one produced with phrase-final intonation in (b).
Table 3: AIC comparison between models, all of which include random intercepts of speaker and word. All additions to the base models lead to significant increases in model fit by ANOVA comparison ($p < 0.001$), with the exception of the model containing pause and sonorant duration ($p = 0.07$).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>sole predictor</th>
<th>with sonorant dur.</th>
<th>with pause dur.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sonorant duration</td>
<td>562.30</td>
<td>NA</td>
<td>271.72</td>
</tr>
<tr>
<td>(continuous; scaled)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position in IP</td>
<td>358.80</td>
<td>348.50</td>
<td>267.44</td>
</tr>
<tr>
<td>(medial vs. final)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pause duration</td>
<td>272.70</td>
<td>271.72</td>
<td>NA</td>
</tr>
<tr>
<td>(continuous; log-transformed)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Best-fitting logistic regression model; [g]-presence as application value, with random intercepts of speaker and word.

| Fixed effects | Estimate | Std. Error | z-value | Pr(>|z|) |
|---------------|----------|------------|---------|---------|
| (Intercept)   | -10.5752 | 1.7062     | -6.198  | <0.001 *** |
| IP position   |          |            |         |         |
| final         | 1.8094   | 0.6845     | 2.644   | 0.008 ** |
| Pause duration|          |            |         |         |
| (log-transformed) | 1.9990 | 0.3574     | 5.593   | <0.001 *** |

562), and that IP position fares a little better (359). Pause duration (273) is by far the strongest predictor.

Models were then fit with a combination of these predictors to investigate the possibility of additive effects, which could be the case if (ng) is conditioned by multiple phonetic cues to boundary strength. ANOVA comparisons between nested models were conducted to quantify whether or not the increase in the amount of variation explained by these additional predictors offsets the cost of a more complex model. In doing this, it became apparent that the strong predictive power of pause duration does not mean that the other collinear variables play no role; adding IP position to a model with pause duration leads to a decrease in AIC that, although small, is statistically significant (267.44, cf. 272.69; $p = 0.007$). The fact that IP position explains a significant portion of the remaining variation suggests that (ng) may be sensitive not only to pause but also to the prosodic phrasing. That is, although the probability of [g]-presence is most strongly influenced by pause, it is also boosted when final in an intonational phrase. This best-fitting model is given in full in Table 4.
5 Discussion

I would now like to address two separate aspects of the variation discussed here: the mechanisms of this innovation in velar nasal plus, and the implications this has for our understanding of how external sandhi processes are conditioned in pre-boundary environments, and also the possible motivations driving this change.

5.1 Mechanisms of innovation

Having successfully ‘disentangled’ the collinearity between what on the surface appeared to be effects of nasal duration (with increasing [g]-presence after longer nasals), prosodic position (with more [g]-presence IP-finally), and following pause (with higher rates of [g]-presence pre-pausally), the results indicate that this innovation in (ng) is conditioned most strongly by the presence/duration of a following pause. This would of course suggest that the apparent relationship between post-nasal [g]-presence and sonorant duration is indirect and stems only from the fact that segmental duration is increased pre-pausally. There is limited evidence to suggest that (ng) is also directly sensitive to prosodic boundary strength; IP position explains a small amount of variation independently of pause, but comparisons between these two predictors suggests that this effect is much smaller in magnitude.

The important role of pause is perhaps best visualised abstractly, as in Figure 7. There is naturally a great deal of overlap between pre-pausal tokens and IP-final tokens, given that the presence of a following pause is one of the major phonetic cues to prosodic boundaries; these tokens that are both IP-final and pre-pausal exhibit post-nasal [g] almost without fail. The non-overlapping portion of Figure 7 reflects the existence of tokens that are pre-pausal but actually medial in the IP; the fact that in these cases [g] is still ever-present provides strong evidence to suggest that only pause is necessary for [g] to surface in these environments, regardless of the prosodic structure.

That this innovation seems to be conditioned most strongly by the presence/absence of pause, rather than by segmental duration or prosodic boundary strength, is rather interesting in light of previous studies that have also attempted to tease apart these factors for other external sandhi processes, e.g. /s/-debuccalisation in Spanish. Kaisse (1996) shows how in the Buenos Aires variety of Argentinian Spanish, word-final coda /s/ does not weaken to [h] when the following segment is ‘temporally distant’, i.e. /s/ is saved from weakening when in pre-pausal position. Much like the argument presented here for (ng) variation, Kaisse claims that this blocking of debuccalisation is triggered on the temporal domain, and is independent of prosodic
Figure 7: Visualisation of the distribution of pauses and IP boundaries in this study’s dataset, and the effect they have upon [ɡ]-presence.

position; this claim is based on the fact that IP-final tokens in fast speech, where the speaker does not pause, still undergo weakening, and that IP-medial tokens where the speaker pauses before resuming with the same intonational contour do not undergo weakening. Comparisons are also drawn with final /r/-devoicing in Turkish, which shows similar behaviour in that devoicing only occurs IP-finally if the IP boundary is marked by a pause (Kaisse 1990).

More recently, it has been shown that other processes exhibit rather more complex behaviour in pre-boundary environments. Kilbourn-Ceron (2017) investigates the conditioning of high vowel devoicing (HVD) in Japanese and addresses the same collinearity issue highlighted in this paper; the results indicate that all three boundary phenomena play a joint role in conditioning HVD, with an interaction between prosodic position and pause presence such that pauses inhibit HVD phrase-medially but promote it phrase-finally. These results paint a much more complex picture relative to earlier claims that suggest the pre-boundary effect is triggered in utterance-final position (Kondo 1997).

No other putative pre-pausal effects have, to this author’s knowledge, been investigated from this perspective; however, this interplay between prosody, pause, and segmental duration does raise questions about the nature of similar effects that have been reported for other external sandhi processes (most notably /t,d/-deletion, as discussed in Section 2.2), which could form a fruitful avenue of further research.

5.2 Motivations of innovation

While the quantitative analysis discussed in this paper has made it possible to determine the mechanisms of this innovation, the motivations behind pre-pausal [ɡ]-
presence have thus far been neglected.

This relatively recent pre-pausal innovation could have been triggered by language-internal factors, specifically by the very nature of how synchronic ‘following segment effects’ are stored and processed in speakers’ grammars. As discussed in Section 2.2, the effect of following consonants in promoting /t,d/-deletion, and of following vowels in inhibiting it, has been analysed under the Obligatory Contour Principle (Guy 1980); the same framework can apply here with (ng) variation. Under this analysis, the effect stems from the avoidance of similar adjacent segments, with following consonants sharing more features with the post-nasal [ɡ] than following vowels. This also accounts for the intermediate effect of following liquids on /t,d/-deletion, which share more features with the preceding /t,d/ than a following vowel but fewer than a following consonant. Crucially, pauses by their very nature do not fit into this typology and could therefore be left open to interpretation with respect to their effect on probabilistic external sandhi processes such as these. The possible consequences of this ‘instability’ could be registered synchronically through inter-dialectal differences (e.g. how the behaviour of pre-pausal /t,d/-deletion differs between speech communities), but also diachronically, as reported here for (ng) in Section 4.1, with changes in pre-pausal behaviour over successive generations of speakers.

We can also turn to an entirely different process, of voiceless stop ejectivisation, in search for possible explanations. Ejectivisation of the English voiceless plosive set /p, t, k/ has been attested by many scholars (Fabricius 2000; Ogden 2009; Gordeeva & Scobbie 2013) and has been said to be increasing over time for /k/, for which it is most frequent (McCarthy & Stuart-Smith 2013). In the same paper, McCarthy & Stuart-Smith explore the factors conditioning /k/-ejectivisation in speakers of Glasgow English, and find that it is favoured not only phrase-finally but also when preceded by a nasal consonant. That is, the words most susceptible to ejectivisation are sink, rank, hunk etc.; these phrase-final /ŋk/ clusters that are frequently ejectivised are essentially the voiceless counterparts to the (ng) clusters that so frequently exhibit post-nasal stop presence in phrase-final position, e.g. sing, rang, hung. These conditioning factors would need to be independently attested in the North West of England, but assuming ejectivisation shows comparable behaviour for the speakers recorded here, these two processes could conceivably be seen as part of the same wider phenomenon: a boundary-marking ‘velar fortition rule’, with parallel changes involved in strengthening voiceless velar nasal+stop clusters through ejectivisation and voiced velar nasal+stop clusters through presence of [ɡ].

This innovation could also be driven by external factors; that is, it could be
socially-motivated. Utterance/phrase-final position is highly salient\(^6\), and as such we may expect the influence of social evaluation to be registered most strongly in this environment. This explanation presupposes two things, however: that (ng) variation is sufficiently salient such that it is subject to social evaluation, and if so, that the presence of \([g]\) carries local prestige in these north western communities, rather than the more regionally standard \([ŋ]\). If this is indeed the case, perhaps this diachronic change in production, with increasing rates of \([g]\)-presence in highly salient pre-pausal and phrase-final environments, actually reflects a perceptual shift within these communities. Younger speakers may well be attaching local prestige to this post-nasal \([g]\), and actively using this vernacular feature to project a northern identity and align themselves with this dialect region (similar to the use of centralised diphthongs by Martha’s Vineyard residents in Labov 1963, or the realisation of goat vowels as \([e:]\) by speakers of Tyneside English, in Watt 2002). The results from Newbrook’s 1999 perception task, discussed briefly in Section 2.1, potentially reflect such a change in evaluation, but further research is needed to provide direct evidence of this perceptual shift.

6 Conclusion

The goal of this study was to investigate the mechanisms of a recent innovation in post-nasal \([g]\)-presence, and in doing so explore the collinear relationship between prosodic boundary strength, pause, and segmental duration in conditioning external sandhi processes. In teasing apart these three factors, all of which on the surface appear to affect the probability of \([g]\)-presence, this study has shown that their relative contributions in conditioning (ng) variation are far from equal. The presence and duration of a following pause provides the strongest explanation of probabilistic \([g]\)-presence; the additive effect of IP position overlaid on this is much weaker, despite theoretical frameworks that foreground the importance of prosodic categories in conditioning external sandhi (e.g. Nespor & Vogel 1986). The apparent relationship between segmental duration and (ng) variation, such that \([g]\)-presence is more likely after longer nasals, arises only because duration is itself correlated with prosodic boundary strength and pause through the process of pre-boundary lengthening. That is, these results suggest that (ng) shows no direct sensitivity to segmental duration.

The results of this study add to a growing body of knowledge about how probabilistic lenition processes behave in pre-boundary position, and in doing so raise questions regarding the nature of similar effects that have previously been attributed

\(^6\)For psycholinguistic evidence of the salience of utterance-final position, see Sundara et al. (2011) (visual-fixation task) and Dube et al. (2016) (EEG experiment).
to one of these collinear factors without due consideration of the others.

The process that gives rise to (ng) variation, whether that be a synchronic deletion or insertion rule, bears a striking resemblance to /t,d/-deletion in a number of ways, most notably the strong effect of following segment which sees the word-final consonant cluster licensed pre-vocally but not pre-consonantally. However, where /t,d/-deletion shows inter-dialectal variation with respect to its behaviour pre-pausally, [ŋɡ] clusters instead show inter-generational variation, with younger speakers re-analysing this pre-pausal environment as one that favours use of the local form with [ɡ]-presence.

This shibboleth of north western dialects, the variable presence of a feature that has been lost in almost all other varieties of English spoken throughout the British Isles, is yet another example of the oft-discussed linguistic conservatism of the north of England. However, in light of the results reported here, this feature is clearly less stable than previously thought: although this variation in (ng) began some time in the Early Modern English period, it still exhibits interesting behaviour today, and even appears to be undergoing a revitalisation in this community. Far more than a mere relic of traditional northern dialects, this variation in (ng) clusters offers valuable insight into the diachronic trajectory of phonological processes (Bermúdez-Otero & Trousdale 2012) and, as revealed in this study, how external sandhi processes are conditioned in pre-boundary environments.

Acknowledgements

This work would not have been possible without the generous funding of the Economic and Social Research Council [NWDTC studentship, grant number ES/J500094/1], or the speakers who were equally generous in giving up their time for this research. I am also indebted to Ricardo Bermúdez-Otero, Maciej Baranowski, and Laurel MacKenzie for their guidance and insightful feedback, to the audiences at FWAV4, NWAV45, and the 25th Manchester Phonology Meeting for their questions, and finally to the LabPhon editors and reviewers, whose comments were instrumental in shaping the final version of this paper. All remaining errors are entirely my own.
Appendix

<table>
<thead>
<tr>
<th>(Suffix boundary)</th>
<th>(Suffix boundary)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The singer wasn’t very talented</td>
<td>The old banger broke down on the motorway</td>
</tr>
<tr>
<td>The youngest was the most loved</td>
<td>The hanger came free with the dress</td>
</tr>
<tr>
<td>The youngerster hated wearing glasses</td>
<td>The gangster was soon behind bars</td>
</tr>
<tr>
<td>The youngest wasn’t very well-behaved</td>
<td>The wrongful accusation was very insulting</td>
</tr>
</tbody>
</table>

1. NP-internal boundary

<table>
<thead>
<tr>
<th>He smiled at the young apprentice</th>
<th>She was given the wrong amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>They had to cancel the Spring event</td>
<td>He had lots of strong opinions</td>
</tr>
<tr>
<td>It was a test for the young professor</td>
<td>It was a very long performance</td>
</tr>
<tr>
<td>He liked feeding the young baboon</td>
<td>She was bored by the long parade</td>
</tr>
</tbody>
</table>

2. VP-internal boundary

<table>
<thead>
<tr>
<th>She gave the ring a quick polish</th>
<th>He offered the gang an apology</th>
</tr>
</thead>
<tbody>
<tr>
<td>He gave the wing a lick of paint</td>
<td>She gave his slang a disapproving look</td>
</tr>
<tr>
<td>They offered the King protection from enemies</td>
<td>He granted the gang permission to stay</td>
</tr>
<tr>
<td>They promised the King possession of their land</td>
<td>She sent the gang potential targets</td>
</tr>
</tbody>
</table>

3. VP boundary

<table>
<thead>
<tr>
<th>The ring adorned her finger</th>
<th>His slang annoyed the teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td>The wing appeared to be loose</td>
<td>His tongue attracted unwanted attention</td>
</tr>
<tr>
<td>The sting became painful</td>
<td>The bang produced a loud noise</td>
</tr>
<tr>
<td>The Spring began with a rainy day</td>
<td>The song began on a high note</td>
</tr>
</tbody>
</table>

4. Intonational phrase boundary

<table>
<thead>
<tr>
<th>“Please pass me the string,” Alicia said</th>
<th>“He doesn’t look very strong,” Adele said</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I think I was stung,” Estelle said</td>
<td>“I burnt my tongue,” Elaine said</td>
</tr>
<tr>
<td>“I don’t know what to bring,” Petunia said</td>
<td>“The price is wrong,” Penelope said</td>
</tr>
<tr>
<td>“It’s a traditional thing,” Patricia said</td>
<td>“The film was too long,” Michelle said</td>
</tr>
</tbody>
</table>

5. Utterance boundary

<table>
<thead>
<tr>
<th>He was too nervous to get up and sing.</th>
<th>He couldn’t believe the answer was wrong.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The bird had damaged its wing.</td>
<td>The drink was surprisingly strong.</td>
</tr>
<tr>
<td>She couldn’t wait for the first few days of Spring.</td>
<td>Her fans didn’t like the new song.</td>
</tr>
<tr>
<td>Her holiday was ruined by a jellyfish sting.</td>
<td>The ceremony dragged on too long.</td>
</tr>
</tbody>
</table>

Table 5: List of elicited sentences by boundary category, vowel height (left column: high vowel, right column: low vowel), and following segment (top rows: vowel, bottom rows: consonant).
Bibliography


Paper III


N.B. This paper has been published online in Journal of Sociolinguistics: https://doi.org/10.1111/josl.12307. When citing, please use the volume, issue, and page numbers of the print version once this becomes available.
Emerging from below the social radar: Incipient evaluation in the North West of England

George Bailey

Abstract

This paper investigates the social meaning of post-nasal [ŋ]-presence, a dialectal variant characteristic of North Western varieties of British English that is claimed to have local prestige. Using a matched-guise approach, this study reveals the absence of a community-wide norm with respect to how [ŋɡ] clusters are evaluated as well as diachronic change in the level of awareness speakers have of this variable. Older subjects are not sensitive of the dialectal status of [ŋɡ] and as a result do not evaluate it differently from [ŋ]; the local form is more accessible to evaluation among younger subjects, for whom the northern indexicality is stronger, but at this incipient stage of social meaning there is no agreement on what the content of this evaluation should be. The results speak to questions regarding the development of shared norms, their role in the speech community, and the granularity of social meaning more generally.

Keywords: social meaning, indexicality, variation, community, phonetics and phonology, velar nasal
1 Introduction

According to a foundational conceptualisation of the speech community, groups of speakers are defined not only by the use of certain linguistic features but also by adherence to a set of shared evaluative norms (Gumperz 1968; Labov 1972a). More recent work has emphasised how normative regularity is constructed locally within particular communities of practice and how individuals exhibit stylistic agency within those communities (e.g. Eckert & McConnell-Ginet 1992; Eckert 2000), but it remains the case that a thorough understanding of a particular linguistic variable requires an account not only of its patterns in production but also of its broader social meaning that emerges through speaker-hearer interactions (Eckert & Labov 2017). Advances along these two lines of research can be mutually informative, and in light of this a number of well-studied linguistic variables have been subject to extensive analysis in the domains of both speech production and perception, e.g. ING (see Trudgill 1972; Labov 1989 on production, and Labov et al. 2006, 2011; Campbell-Kibler 2011b on perception), TH-fronting (see Baranowski & Turton 2015 on production, and Levon & Fox 2014 on perception), and T-glottalling (see Fabricius 2000; Straw & Patrick 2007 on production, and Schleef 2017 on perception).

This paper addresses the social evaluation of one particular feature of dialects spoken in the North West of England. ‘Velar nasal plus’ (hereafter VNP) refers to the variable presence of post-nasal [ɡ] both word-finally as in wrong [ɹɒŋ]-[ɹɒŋɡ] and word-medially as in singer [sɪŋə]-[sɪŋɡə]; despite recent advances in our understanding of how this variable behaves in speech production (e.g. Watts 2005), we still know relatively little about how this dialectal form is actually evaluated by listeners. VNP is an interesting case study of social evaluation for a number of reasons: it is a rare case of a regional variant being favoured in more formal elicited discourse styles, leading to claims that it has overt local prestige (Beal 2004: 127). On the other hand, there are conflicting reports with respect to its social and stylistic stratification and the way in which such patterns reflect linguistic norms and standard language ideologies; the local [ŋɡ] form is favoured in more formal discourse styles but is also more frequent among working class speakers (Watts 2005). This of course contrasts with the widely-attested correlation between social and stylistic variation where the form favoured by higher socioeconomic classes is also favoured in more careful styles of speech (Labov 2001). Finally, research into the social meaning of VNP may provide an explanatory mechanism behind a recent change in progress that sees increasing rates of [ɡ]-presence before pause (see Papers I–II of this thesis); this could be an example of evaluation-driven change, particularly considering the salience of phrase-final
environments.

In this paper, a matched-guise task is used to probe listener evaluations of [g]-presence, and the results indicate a striking lack of community-wide agreement. The evaluation of VNP is currently undergoing change: older subjects do not recognise [ŋɡ] as a northern form and do not evaluate it differently from [ŋ] in terms of perceived professionalism; although there is heightened awareness of the dialectal status of [ŋɡ] among younger subjects, this does not translate to uniform evaluation. I interpret this as reflecting an early stage of incipient social meaning; the awareness of [ŋɡ] as a marker of northern dialects is increasing but not yet strong enough to result in the shared community norm predicted by Labov’s Principle of Uniform Evaluation (2001: 214). This inter-speaker variation with respect to how [ŋɡ] is evaluated could be further motivated by orthographic factors, with the presence of <g> in the spelling overriding the traditionally-negative indexicalities of northern accents.

The results of this study have implications for our understanding of how shared norms are developed, their role in the speech community, and more generally for theories of indexicality with respect to non-standard or dialectal variants; the results are also pertinent to a number of current theoretical issues such as the extent to which social evaluation can act as a driving force in the incrementation of sound change.

2 Velar nasal plus

‘Velar nasal plus’ is a term coined by Wells (1982b: 365) to refer to the variable presence of post-nasal [g] in dialects spoken across the North West and West Midlands regions of England; the geographic distribution of this non-coalesced [ŋɡ] form was established in the 1950s by the Survey of English Dialects (Orton et al. 1978), and its presence has been independently attested in a number of varieties spoken throughout these regions: Liverpool (Knowles 1973), Cheshire (Watts 2005), Birmingham (Thorne 2003), Cannock (Heath 1980), the Black Country (Asprey 2015) etc.

The envelope of variation of VNP actually encompasses two environments, denoted hereafter by (ing) and (ng): the former refers to unstressed -ing clusters that in all other varieties of English exist as an alternation between [ɪn] and [ɪŋ] in words such as walking or building, whilst the latter refers to stressed ng clusters in words such as thing or hang, which are of course invariably realised with the bare velar nasal form by speakers of other dialects. Although the social meaning of [g]-presence in (ing) has been subject to recent investigation (Schleef et al. 2015), there is no reason to believe that this meaning is shared across both environments, particularly as they differ so much in production. Where reported, rates of VNP as a realisation of the
(ing) variable are as low as approximately 1% in conversational styles, contrasting with the two-way alternation in (ng) which sees rates of [ɡ]-presence as high as 40% (Watts 2005).

2.1 Existing perceptual evidence

To this author’s knowledge, the only study to explicitly address the evaluation of [ɡ]-presence in (ng) is Newbrook’s 1999 work in West Wirral, a region of Merseyside in North West England, where there exists a system of three competing influences from RP, Cheshire, and Liverpool (or ‘Scouse’) varieties of English.

In one task, Newbrook’s informants were asked which variant they consider to be the norm, where norms are described as “those forms regarded as prestigious and/or as targets for emulation (in, for instance, formal settings)” (1999: 100). For thirteen of the seventeen variables included in this study, almost 100% of the informants identify the RP form as the norm; crucially, word-medial -ng- and word-final -ng are among the four exceptions to this pattern. In the case of the former, a majority of informants orientate towards the local [ŋɡ] form, while for the latter the pattern of responses was more evenly split. Newbrook claims that cases where the RP form was rejected as a norm can be attributed to either ‘conscious rejection’, where speakers are fully aware of the dialectological facts but still reject the RP form as a prestigious variant, or ‘confusion/ignorance’, claimed to be more common, where speakers accept RP as the standard but aren’t actually aware of what the RP form is.

In the second task, subjects were asked which form they use most often; self-report tasks are well-established measures of sociolinguistic prestige, with their results said to “reflect adherence to the social norms of how speech should be pronounced rather than a report of actual use” (Labov 2001: 194). Although the results from this are less categorical in nature, there is still a strong preference for the RP form across most of the dialectal features under study. Once again, however, -ng- and -ng are among the exceptions to this pattern, with only 19% of subjects reporting use of [ŋ] for the former; much like the norm-identification task, the responses were much more variable for word-final -ng. Results from these two complementary tests are illustrated in Figure 1.

The overall picture is one in which the community are in strong agreement on the evaluation of -ng-, with the local VNP variant almost unanimously endorsed as a norm and claimed to be the most frequently used form; in contrast to this, word-final -ng exhibits a great deal of within-community variation with respect to its overt sociolinguistic evaluation.
Figure 1: Results from the norm-identification and self-report tasks in West Wirral, Merseyside, including -ng (e.g. sing) and -ng- (e.g. singer). Based on Tables 5.2 and 5.3 from Newbrook (1999: 100–102).

2.2 Inferences from production studies

Given our understanding of how patterns of variation in speech production can indirectly reflect the social evaluation of language, we can also look at how VNP is stratified along stylistic and sociodemographic dimensions for additional insight into its social meaning.

Stylistic stratification is often a strong indicator of how linguistic variants are evaluated, where increased attention paid to speech in formally-elicited styles results in decreased use of non-standard forms and an increase in forms considered to be overtly prestigious (Labov 1972a: 208). Based on this, it would seem that the non-coalesced [ŋɡ] form does carry local prestige in these dialect regions given the attested pattern of increased use in more conscious styles (Mathisen 1999; Schleef et al. 2015). However, this interpretation of style-shifting has been subject to criticism for being a simplistic and reactive-orientated view to intra-speaker variation (see Coupland 2007 for discussion); the unnatural and performative nature of word list elicitations presents speakers with an opportunity to project their identity through linguistic means à la Schilling-Estes’s (1998) proactive model of style-shifting. Interpreting stylistic behaviour in this way is further confounded by the collinearity between discourse formality (ranging from casual conversation to formal elicitations) and prosodic factors such as speech rate and intonational phrasing. This could be
particularly problematic in the case of (ng) variation due to its sensitivity to pause (see Papers I–II of this thesis).

Notions of prestige and ‘standard’ language use are not only reflected stylistically but also by stratification along sociodemographic lines. Use of [ŋɡ] reportedly correlates with socioeconomic status; using neighbourhood as a proxy for social class, Watts (2005) compares the rates of VNP in Wilmslow (typically middle class) and Coshaw (typically working class) and finds that use of the local [ŋɡ] form is much more frequent among speakers living in the latter, working class neighbourhood. While this is only an indirect measure of socioeconomic status, this observation is at least backed up by Mathisen (1999) who reports that working class speakers are leading the ‘revitalisation’ of VNP in Sandwell. However, far from being a defining feature of working class vernacular, there are numerous claims throughout the dialectological literature that [ŋɡ] is actually used throughout the social scale. Wells (1982b: 365) reports its presence for all but the “very small layer of RP speakers at the top”, and this is echoed by Wakelin (1984) and Heath (1980) in the Midlands, and by Knowles (1973: 295) in the North West city of Liverpool, who describes (ng) as a “conflict of local and national norms”. It has also been described as having a relatively low social profile and not being a “crashing local-accent feature which ambitious upwardly-mobile northerners might want to try to modify or eliminate” (Wells 1997: 43).

In summary, the production studies of VNP do not provide a clear consensus regarding the social groups that favour use of [ŋɡ]. The observation that it is favoured by working class speakers would suggest that it does not have overt sociolinguistic prestige in these communities; however, if the observations of stylistic stratification are taken at face value (i.e. as a reflection of style-shifting triggered by formality of discourse), this would indicate that [ŋɡ] does indeed have local prestige. As a result of this contradiction, and the relative lack of direct perceptual evidence, we still do not know which of [ŋ] or [ŋɡ] carries local prestige in these speech communities nor the exact social attributes indexed by VNP, thus motivating the present perception study.

3 Methodology

3.1 Experimental paradigm

This study adopts a matched-guise approach, following on from perceptual work on (ing) by Labov et al. (2006, 2011) in the US and Levon & Fox (2014) and Schleef et al. (2015) in the UK. Specifically, the ‘newscaster’ paradigm is used, where subjects are told they will hear different audition tapes from a speaker applying for a role as a
news broadcaster. These tapes differ only in the presence/absence of post-nasal [ŋ] and any differences in how they are rated by subjects can therefore be attributed to the variable presence of this dialectal feature. The newscaster context is important as it has been shown to be particularly effective in priming “overtly prestigious sociolinguistic norms” (Levon & Fox 2014: 189) and will therefore reveal whether or not VNP has the overt local prestige that has previously been ascribed to it (Mathisen 1999; Beal 2004). This methodological approach is particularly applicable in a British context in light of actual cases of linguistic prejudice against northern news presenters such as the BBC’s Steph McGovern (Furness 2013), and there is further evidence that roles of a similar professional nature are subject to the same linguistic scrutiny, particularly in the case of northern accents, e.g. in teaching (Baratta 2017). It has also been shown recently that matched-guise techniques are not confounded by speech style, with subjects reacting similarly to the use of non-standard forms in conversational speech as in elicited speech (Tamminga 2017). This is an important finding and one that allows us to generalise results beyond this specific experimental context with a reasonable degree of confidence.

3.2 Rating scales

After listening to each ‘audition tape’, subjects are asked to rate the speaker on a number of 7-point Likert scales. As the concept of ‘prestige’ can be operationalised in a number of ways, these descriptive scales were carefully selected according to the success or otherwise of the afore-mentioned matched-guise studies of (ing).

In this study, the ‘professionalism’ scale is used, following on from Labov et al. (2006, 2011) in the US and Levon & Fox (2014) in the UK, but is supplemented by other descriptors that were found by Schleef et al. (2015) to be relevant in the evaluation of VNP as a realisation of (ing), namely ‘education’ and ‘formality’. Finally, a ‘northern’ scale is included to investigate the extent to which subjects recognise this as a dialectal form; the inclusion of this scale could prove to be important given the speculative comments made by Newbrook (1999) regarding whether subjects who endorse [ŋ] are simply unaware of the dialectological facts (i.e. they fail to realise that this is a local vernacular form of north western dialects) or they are aware but still accept it as an overtly prestigious form.

3.3 Stimuli

The recordings were made by a 56 year-old female speaker of Manchester English, who was asked to read out a number of fictional news headlines in a manner imitative of traditional news broadcasts. The headlines were read out once with [ŋ] present for
all (ng) tokens, and then once again with [ɡ] absent for all tokens. These recordings were then cross-spliced using Praat (Boersma & Weenink 2017) to produce two guises for each headline group, which differ in the presence/absence of post-nasal [ɡ] but are identical in every other respect. The recordings were carried out in a sound-attenuated booth, using a Sony PCM-M10 recorder and saved at a 44.1KHz sampling rate.

There are three passages, each containing two tokens of (ng). The three groups differ with respect to the morphophonological context in which (ng) appears, with one containing two pre-consonantal tokens (e.g. *sing tunes*), one containing two phrase-final tokens (e.g. [...] *sing*), and one containing two word-medial pre-vocalic tokens (e.g. *singer*). These three environments were chosen because they have been shown to favour [ɡ]-presence to differing degrees in production data, which could result in a context-dependent evaluation of VNP: it is used in pre-vocalic environments at much higher rates than in pre-consonantal contexts, and there are suggestions that its use is increasing over time phrase-finally (see Papers I–II of this thesis). The three headline groups are given below.

**Phrase-final:**

1. Scientists working on the Large Hadron Collider have today found new evidence that reveals what the universe was like at the time of the Big Bang.

2. In other news, weather experts warn that increased levels of global warming have led to the highest temperatures ever recorded in Spring.

**Word-medial, pre-vocalic:**

1. Justin Bieber came under fire yesterday after pictures surfaced online that show him spitting at a fan. The latest scandal has prompted widespread criticism of the Canadian singer.

2. In sport, Liverpool today dropped more points in the absence of their star player Sadio Mané, leading to claims that the club are too reliant on the right winger.

**Word-final, pre-consonantal:**

1. The government is demanding that zoos increase security after the latest incident saw an escaped gorilla attack a young child.

2. In politics, Theresa May has warned that Britain may not see the benefits of Brexit for many years, admitting that negotiations would be a long process.
Each passage also contains two tokens of unstressed (ing), which could present a possible confound as this environment overlaps with stressed (ng) in being a possible context for post-nasal [ŋ]-presence. It is important to note, however, that outside of word list elicitations unstressed (ing) almost never surfaces with this non-coalesced realisation (rates as low as 1% are reported by Watts 2005 and Schleef et al. 2015). These tokens were produced with the plain velar nasal by the speaker, and were left unaltered for the survey stimuli as this would be the most unmarked realisation for this style of read speech. That is, a token of [ɪŋɡ] would sound somewhat unnatural given its rarity in speech and would potentially attract attention away from the target (ng) segment, but equally a token of [ɪn] would likely be perceived as far too informal for newscaster speech, resulting in the same effect. Additionally, any influence of the velar nasal in (ing) would simply be to decrease sensitivity to (ng) (see Campbell-Kibler 2009 and Watson & Clark 2013 on ‘bullet-proofing’), as it is balanced across subjects and guises; it wouldn’t impact the directionality of the evaluation, just its magnitude.

Four other headline groups were included as distractors; these contain no tokens of (ng) and are again represented by two guises, although in this case the guises are simply two different readings of the same passage with no systematic differences between them. The only exception to this is the final distractor pair, which contains two tokens of intervocalic /t/ (in the words water and better) realised as a glottal stop [ʔ] in one guise and a canonical alveolar plosive [t] in the other. This was included to provide a baseline against which the evaluation of (ng) can be compared, given the salience and well-established stigma of intervocalic /t/-glottalling in British English (Fabricius 2000).

### 3.4 Subjects and experimental design

In total, responses were collected from 35 subjects, all native speakers of British English born and raised (at least up to the age of 13) in the North West of England. This population sample is balanced by age, consisting of 18 younger subjects (with a mean age of 23) and 17 older subjects (mean age of 58). The survey, which was hosted on LimeSurvey using embedded sound files, was distributed on university mailing lists and social media platforms, and through friends and acquaintances.

This study employs a repeated measures design, where all subjects are exposed to all guises. It is counter-balanced to prevent potential confounds relating to the order of stimuli; for each [ŋ]-[ŋɡ] pair, half of the subjects are exposed to the [ŋ] guise first and the other half hear the [ŋɡ] guise first. The same counter-balancing procedure is applied to the environments themselves such that a third of subjects
hear the pre-consonantal pair first, a third hear the pre-vocalic pair first, and a third hear the phrase-final pair first. The distractor pairs are presented between each (ng) pair, and the [t]-[ʔ] guises always come at the end of the experiment to ensure the highly-informal [ʔ] guise doesn’t impact subjects’ later ratings for the target stimuli.

At the end of the survey, subjects were asked a number of biographical questions to collect information about their age, gender, and where they grew up; they were also asked if the words singer and finger rhyme in order to establish whether or not they actually have VNP in production. With this information, we can look for possible change over time with respect to the social meaning of (ng).

4 Results

There are two ways of approaching the analysis of subjects’ responses: using the absolute ratings awarded to the [ŋ] and [ŋɡ] guises, or calculating difference scores on a speaker-by-speaker basis (i.e. for each speaker calculate the difference in rating for each [ŋ]-[ŋɡ] guise pair). In the first part of this analysis, the absolute ratings will be used to provide an overview of the scores awarded to the VNP guises; the subsequent section will investigate the difference scores to provide a closer look at the extent to which the [ŋ]-[ŋɡ] guises elicit contrasting evaluation on an individual basis.

4.1 Absolute ratings

Figure 2a shows the distribution of ratings by guise and environment, for all four evaluative scales. What is immediately striking is how similar all of the distributions are with respect to their shape and the location of their peak; it seems to suggest that subjects are not evaluating the [ŋ] and [ŋɡ] guises differently, or at least that they are doing so to a very small degree.

The direction of this difference is, however, consistent across all environments. It is also uniform across scales in the case of professionalism, formality and education. For these three scales, which can all be understood as measures of overt sociolinguistic prestige, it is actually the standard [ŋ] guise that receives the slightly more positive evaluation. In the case of the northern scale, the direction of this difference is unsurprisingly inverted, with [ŋɡ] heard as more northern than [ŋ]. The difference in northern ratings is also fairly small, but this is likely due to the fact that, owing to the presence of other northern features in the speaker’s voice\(^1\), even the [ŋ] guise

\(^1\)The passages contained a number of instances in which the speaker exhibited other northern features, such as the lack of distinction between the foot–strut and trap–bath lexical sets in words such as young and after, respectively. Other than these supra-regional features, there was an avoidance of non-standard forms such as vowel monophthongisation, post-vocalic /i/, fronted /θ, δ/, or dropped
Figure 2: Distribution of ratings by guise and environment for all four descriptive scales. Vertical lines represent mean ratings. Velar nasal plus guises in (a); /t/ distractor guises in (b).
receives high ratings on this scale (averaging over 5 on the 7-point Likert scale across all environments).

Crucially, the small effect size evident in the case of VNP guises does not result from failings in the experimental design; in the case of the distractor [t]-[ʔ] pair, the matched-guise task elicits drastically different responses as one might expect. This is illustrated in Figure 2b, where the guise featuring glottal replacement of /t/ has average ratings as low as 2.2–2.9 on the professional, educated, and formal scales. The contrast between how these two variables elicit differences in evaluation suggests that VNP does not attract the same magnitude of social meaning as intervocalic /t/-glottalling, which is known to be widely stigmatised (Fabricius 2000).

Although these results seem to indicate that VNP is not subject to a great degree of social evaluation, the indexical properties of these two forms becomes clearer when (a) old and young age groups are considered separately, and (b) difference scores are used for greater insight at the level of the individual.

4.2 Difference scores

The within-subjects design of this experiment, where each subject is exposed to both guises in all environments, facilitates an investigation of social meaning on a subject-by-subject basis. For each subject-guise-environment pair, the rating given to [ŋ] is subtracted from the rating given to [ŋɡ], such that a positive difference score corresponds to a higher rating for the local [ŋɡ] form, a negative score corresponds to a higher rating for [ŋ], and a score of 0 means that the subject rated both guises the same for that particular scale and environment.

As one might expect from the results already presented, many of these difference scores are indeed 0, reflecting the lack of social meaning ascribed to VNP. To be exact, 59% of the 420 difference scores are 0, and 86% of them fall within the ±1 range. Interestingly, the deviations from 0 are not equally distributed across young and old subjects. When we consider these two social groups separately, as in Figures 3 and 4, it becomes apparent that not all members of the speech community are in agreement with respect to the social meaning of VNP.

The most immediately-apparent pattern is the variation in difference scores among young subjects relative to older subjects. Among the latter, [ŋ] and [ŋɡ] do not differ with respect to how they index professionalism, education, formality, or even northernness. Younger subjects, however, are much more likely to rate these two variants differently on these scales, suggesting that the variable has greater so-
Figure 3: Difference scores by environment, scale, and age group. Diamond indicates mean difference score. Dotted line indicates neutrality (difference score = 0), where points to the right reflect higher ratings for [ŋɡ] and points to the left reflect higher ratings for [ŋ].

Figure 4: Distribution of difference scores by scale and age group. Dotted line indicates mean difference score.
cial meaning among the younger generations. This is further quantified by comparing measures of central tendency between the two age groups (SD = 0.54 for old subjects’ difference scores, cf. 1.43 for young). In terms of the direction (i.e. polarity) of these difference scores, what is perhaps most curious is that even among young subjects there is disagreement; there is only a slight trend towards negative values for the scales that measure overt prestige (i.e. professionalism, education, and formality), reflecting a more positive indexicality of [ŋ] relative to [ŋɡ], but overall the variation in values is striking.

For the northern scale, there is more consistency among young subjects in terms of the spread of negative and positive difference scores (25 positive, cf. 6 negative); this suggests that they at least agree on [ŋɡ] as a feature of northern dialects, but crucially this does not translate to uniform evaluation along the scales that prime overt prestige. There is also no correlation between the northern and professional difference scores among young subjects (Spearman’s $r_s = -0.09, p = 0.53$); that is, it is not the case that higher northern scores predict lower professional ratings and lower northern scores predict higher professional ratings.

It should also be pointed out that, as Figure 3 illustrates, this contrasting behaviour of young and old subjects is not restricted to a particular environment; in fact, all three morphophonological environments included in this study behave in largely the same way, despite the prior prediction that pre-pausal [ŋɡ] may behave differently due to the ongoing change in production.

To test the statistical significance of these effects, two mixed-effects linear regression models were fit to the difference scores: one to test the effect of age on professionalism (specifically the increase in variation), and one to test the effect of age on northernness (specifically the increase in value).

Because in the former case the hypothesis under consideration is simply an increase in the magnitude of the difference score, regardless of its direction (i.e. whether it’s negative or positive), the polarity was removed from the dependent variable. That is, the model does not distinguish between a value of -3 (rating [ŋ] as more professional than [ŋɡ]) and a value of +3 (rating [ŋɡ] as more professional than [ŋ]), since in both cases the magnitude of the difference score (i.e. deviation from 0) is the same. In the case of the northern scale, we are testing for a specific direction of the effect (that younger subjects are more likely to rate [ŋɡ] higher than [ŋ] relative to older subjects); as a result, this model uses the unchanged values. Both models contain fixed effects of age group, environment, and their interaction; they also include a random intercept of subject to account for the within-subjects experimental design.

The results in Table 1 show that for both the professional and northern scales there
Table 1: Mixed-effects linear regression coefficients for (a) the northern scale, and (b) the professional scale; difference score as the dependent variable in (a), deviation from 0 as the dependent variable in (b). Random intercept of subject. Reference level of old and pre-consonantal.

is a significant effect of subject age. The positive coefficients indicate that in the case of the former, there is a greater difference between [ŋ] and [ŋɡ] on the professional scale for younger subjects; in the case of the latter, the extent to which [ŋɡ] is heard as more northern than [ŋ] is greater among young subjects relative to older listeners.

In neither case is there a significant interaction between age and environment, nor a significant main effect of environment itself (even when removing the interaction and including them as independent predictors). The evaluation of VNP, and in particular this change in its evaluation across generations, is uniform across all morphophonological environments — an important point that will be further discussed later in this paper.

5 Discussion

The results of this study reveal a great deal about the salience of VNP, and how this has increased over time to give rise to social meaning. Furthermore, the lack of agree-
ment with respect to certain aspects of this meaning, in particular the overt evaluation of [g]-presence, speaks to questions about the indexicality of this variable and of northern accents more generally. The uniformity across morphophonological environments also has implications for our understanding of the granularity of evaluation. The following discussion will explore all three of these points in closer detail.

5.1 Salience

Generally speaking it is clear to see that VNP is not a salient feature, and that as a result of this it does not evoke the same degree of evaluation as more salient phonological variables such as /t/-glottalling, which was included not just as a distractor but to provide a baseline for such comparisons.

However, despite the popularity of ‘salience’ in the sociolinguistic literature, it is not straightforward to operationalise or indeed quantify this concept. Jaeger & Weatherholtz (2016: 3) define the salience of a linguistic variable as “a function of its (perceived) informativeness about social group membership”, and in doing so propose a measure to quantify this concept in terms of surprisal and new information content. Alternatively, Kerswill & Williams (2002b: 81) defines it as the extent to which a linguistic form is “in some way perceptually and cognitively prominent”, and a thorough discussion of other definitions is provided by Honeybone & Watson (2013). Trudgill (1986: 11) proposes four criteria that linguistic variables should meet in order to be considered salient; they must:

1. be taking part in ongoing linguistic change
2. involve an alternation between forms that are perceptually very different with respect to their phonetic realisation
3. be phonologically contrastive
4. involve a ‘standard’ variant, particularly one that is reflected orthographically

Applying these conditions to the case of (ng), it is clear to see why this variable might not have the same degree of salience as other variable phenomena. It is undergoing change in progress but this is restricted to one prosodically-defined environment, it is not contrastive but rather purely allophonic, and with respect to (4) it is actually the local [ŋɡ] variant that arguably more closely resembles the orthography.

There are a number of other factors that have been argued to influence the salience of a linguistic variable, such as frequency of occurrence (Bardovi-Harlig 1987 claims that more frequent forms have greater salience) or prosodic prominence (Yaeger-Dror 1993 claims that forms in prosodically strong positions, such as initial in the word,
are more salient). Both lend further support to the proposed lack of salience of (ng), which is relatively infrequent in conversation and never occurs word-initially.

However, all of these testable criteria presuppose that salience is a static property of linguistic variation, which is not necessarily the case. More recent work has argued for a dynamic approach to sociolinguistic salience, determined by contextual factors such as when and where a variant is used in discourse, the interlocutor who uses it, and by the listener’s past experience and prior expectations of the distribution of the variant in question (Drager & Kirtley 2016; Hay et al. 2018). In this way, salience is intrinsically linked to context-specific probabilities of use and therefore the extent to which a variant is unexpected in a stretch of discourse, with respect to both social and linguistic factors.

Despite the importance of approaching salience as a dynamic property, it remains the case that certain linguistic variables are generally more salient in a speech community than others. This is reflected directly by Labov’s marker-indicator-stereotype typology, determined by the degree to which a linguistic variable exhibits stylistic/social stratification and the level of metalinguistic commentary it receives (Labov 1972a). As such, I suggest a distinction between ‘global’ and ‘local’ salience, both of which contribute to the overall degree of awareness of each instance of a particular sociolinguistic variable. With high global salience, combined with strong indexical value, a variable may eventually reach the point of enregisterment (Agha 2004; Johnstone 2016).

Since local salience is not static, it does not make sense to talk about it undergoing community-wide change. Rather, an apparent time interpretation of this paper’s results suggest a change in the global salience of this variable. The variable as a whole is becoming a more salient dialectal feature and as such is becoming increasingly associated with northern accents over time. This heightened sensitivity to the dialectal status of (ng) could arise from a number of mechanisms, such as increased mobility and therefore more contact with speakers of non-northern varieties. It could also be the case that this alternation is more salient for young speakers because among the younger generations the rates of [ɡ]-presence are much higher; that is, by using it more in production northern speakers may now be more aware of its absence among their non-northern peers.

5.2 Variation in meaning

Although the increase in sociolinguistic salience has made this variable more accessible to social meaning, there is no widespread agreement on what this meaning should be. At a community-wide level, the evaluation is trending very slightly towards a
standard-driven norm where subjects penalise the local form in terms of perceived professionalism; however, at the level of the individual there is still a great deal of variance even among those who agree that [ŋɡ] is northern-sounding. It is important to consider why this is the case.

It is relatively straightforward to understand why subjects may hear [ŋɡ] as more northern and then rate it lower on the professional scale; regional varieties of British English are often stigmatised, and this is particularly the case for northern dialects (Furness 2013; Baratta 2017). The survey conducted by Coupland & Bishop (2007) actually suggests that Manchester and Liverpool English, both of which fall within the VNP isogloss, are particularly stigmatised; they are ranked 27th and 30th respectively for social attractiveness out of the 34 varieties included in the study2. Therefore, the evaluation of VNP likely arises through second-order indexicalities with northernness (Silverstein 2003) and the fact that in England an RP-norm still pervades professional settings such as politics or, in this case, newsreading.

It is somewhat more difficult to account for the cases where subjects hear [ŋɡ] as not only more northern but also more professional and/or educated than the standard [ŋ] form. In the case of this particular variable there are a number of possible explanations however. There could be orthographic influences due to the presence of <g> in spelling; subjects may be of the belief that a realisation with [ɡ]-presence is more ‘correct’ because it more closely reflects the orthographic representation of these clusters. Related to this, it should also be noted that [ŋɡ] is the more historically-conservative form, once present across all dialects of British English, and that those varieties without post-nasal [ɡ] (including RP) are actually featuring the innovative variant.

The generalisation also exists that in many cases dialectal or variable processes of segment lenition (particularly deletion) carry social stigma, e.g. /h/-dropping, /t/-glottalling, and (ɪŋ)3; that is, the prescriptivist idea exists that dropping sounds is characteristic of ‘lazy’ or ‘incorrect’ speech. It is possible that, because of this pressure, some of these young speakers in the North West still think of [ŋɡ] as the ‘correct’ pronunciation, despite knowing that it is a non-standard dialectal variant. This is made more likely by considering its frequent use in the kind of clear speech styles elicited by word lists and other such tasks; the notion of clear speech variants does not completely overlap with concepts of overt prestige or standard language, despite

---

2 Velar nasal plus is also a feature of Birmingham English and the varieties spoken in the Black Country, which are actually the most stigmatised dialects in the entire study by Coupland & Bishop (2007); they are ranked 34th and 33rd respectively.

3 Of course variation between [ɪn] and [ɪŋ] is not actually a case of segment deletion, but it remains the case that most non-specialists refer to the former as ‘g-dropping’.
frequent conflation of the three. This argument receives further support from the ob-
servation that [ŋɡ] is becoming increasingly frequent in pre-pausal environments (see
Papers I–II of this thesis), which are likely to be associated with the kind of citation
forms that are characteristic of ‘clear speech’.

An anonymous reviewer raises the possibility that the lack of agreement among
younger subjects arises because they are in fact members of different speech com-
minities within the North West, which evaluate (ng) in different ways. If this is the
case, and there is lack of agreement because these north western subjects belong to
separate communities, this would mean that the observed developments reflect in-
dependent but simultaneous changes in opposite directions, which would be highly
unlikely; while this possibility cannot be ruled out directly, there are a number of
arguments against this explanation.

Firstly, descriptive studies of this variable across the North West, where they test
the same predictors, largely report similar effects with respect to social and inter-
nal factors (e.g. Knowles 1973 in Liverpool; Watts 2005 in Cheshire; Bailey 2015 in
Manchester and Blackburn; see also Mathisen 1999 in Sandwell, West Midlands). It
would be unlikely that the evaluation of this variable exhibits diversity within the
North West region given that this region behaves as a homogenous community with
respect to the production of (ng). It is also important to recall that there is eval-
uative uniformity across the older generations, which would not be predicted if the
divergence among younger subjects actually reflects their belonging to different com-
munities.

Furthermore, this regional uniformity is confirmed with independent perceptual
evidence from the Newbrook (1999) study discussed earlier; the results point to a
similar state of affairs in one specific community, West Wirral, where there is a clear
lack of agreement in the evaluation of word-final (ng). Given the forced choice nature
of the task, this either reflects a lack of uniformity in evaluation, or that there is
absence of any evaluation and as such subjects are responding randomly. Either way,
the results corroborate the argument made in this paper, providing further evidence
that the results described here are generalisable to the whole of this region.

What do these results mean, then, for the concept of the speech community?
This paper addresses only one variable, and as such the lack of normative regular-
ity should not be taken as strong critique against Labov’s (1972a) definition of the
speech community as a group of individuals with shared evaluative norms. A speech
community is a difficult concept to define, particularly with the move towards more
locally-relevant categories of group membership in the second and third waves of
variationist study, but it is clear that it cannot be defined by one single property.
Instead, it should be viewed as a combination of factors, including uniform evaluation, geographic proximity, density of contact (e.g. Bloomfield 1933; Gumperz 1971), shared patterns of production (e.g. structured differentiation of stable variables along sociodemographic lines), and shared participation in ongoing change.

However, the results do suggest that the Principle of Uniform Evaluation (Labov 2001: 214) should be taken with caution, since salience and social meaning are clearly dynamic properties of linguistic variation that are subject to change. It may be the case that uniform evaluation only holds when the social meaning of a variable is itself stable. Because the social meaning of VNP is relatively new and still being developed — with ongoing change in its social salience and the degree to which it indexes group membership among northern speakers — it has not yet reached a stable stage of uniform evaluation. It should also be noted that Labov’s Principle of Uniform Evaluation is based primarily on uniform patterns of synchronic style-shifting, which this variable does in fact exhibit (Mathisen 1999; Schleef et al. 2015); as discussed earlier, however, this is an indirect proxy of evaluation confounded by prosodic differences in elicitation tasks, and this paper shows that style-shifting is not always corroborated by direct attitudinal evidence.

5.3 Indexicality of (ŋɡ)

The results presented in this paper point to some degree of directionality and order of causality in the development of multi-layered indexical fields. The move towards community-level agreement with respect to the northern status of [ŋɡ] suggests that this stage is a prerequisite for the other social meanings that follow from this association, which are still being developed towards a community-wide norm.

This evaluative behaviour can be interpreted under the indexical frameworks proposed by Ochs (1992, 1996), who distinguishes direct and indirect indexicality, and Silverstein (2003), who proposes multiple orders of indexicality.

The results suggest that higher order indexicalities are perhaps more individualistic and therefore more likely to exhibit inter-speaker variation. That is, the subjects from this study who show clear awareness that [ŋɡ] is a dialectal variant are all in agreement that it is a northern form – which is unsurprising, given that this is based on objective dialectological facts – but as one climbs up into higher order indexicalities, there is more potential for individual ideologies to shape the indexical field.

This is particularly applicable to (ŋɡ), hence the complete absence of agreement in evaluation, due to what appears to be its incipient stage of social meaning. We are beginning to see an association emerge between this regional form and the varieties of English spoken in the North West of England; however, the fact that older generations
within this community are still blind to the dialectal status of this feature suggests that this association has not long been established. Over time, particularly when this first-order indexicality is shared among all members of the community, there may also be a move towards community-wide agreement at higher orders; however, Eckert (2008: 467) has previously argued that variation in the content of such evaluation is still possible even when there is uniformity at lower orders, arguing that “while the entire population might agree on first-order indexicality — who uses what variant — the evaluation of that differentiation can differ across the population”.

It is also likely that the indexical field of this variable is much more complex than a simple association with northernness and a perception of prestige that stems from this association; as Schleef et al. (2015) suggest for [ɡ]-presence in the unstressed (ing) environment, for some subjects use of this local variant may index a type of ‘over-articulate’ stance, which may overlap with the indexical field of its presence in the (ng) environment, thus resulting in some of the negative evaluations found in this paper and contributing further to the inter-speaker variation with respect to how this variable is evaluated.

The normative irregularity of (ng) adds to a growing body of evidence suggesting that indexicality is not a ‘fixed property’ of linguistic variables but is in fact a dynamic semiotic field prone to change and reinterpretation: Bucholtz (2009) reports indexical change in the use of guey in Spanish, and parallels can be drawn with a number of studies that have shown distinct indexical fields tied to a variable’s use in particular registers and dependent on listener-specific interpretations and ideologies (Campbell-Kibler 2008, 2011a; Podesva 2008; Moore & Podesva 2009; Pharao et al. 2014).

5.4 The granularity of social meaning

A notable result is the uniformity with which VNP is evaluated across different phonological environments. The lack of interaction between age group and environment (differentiating between pre-consonantal, pre-pausal, and word-medial intervocalic contexts) indicates that whilst the overall alternation between [ɡ]-presence and absence has accrued social meaning over time, this meaning is not concentrated on its use in a specific environment. This is particularly interesting in light of the ongoing change taking place pre-pausally; in Paper II of this thesis it is shown that rates of [ɡ]-presence are increasing over time in a ‘pre-pausal’ environment conditioned by temporal and possibly prosodic factors. The observation in this study that [ɡ]-presence clearly does not carry overt or local prestige, coupled with the fact that its evaluation in pre-pausal contexts is comparable to all others, suggests that this change is not evaluation driven. Rather, it is proceeding fully under the radar.
This observation that the change itself is not subject to evaluation has further implications for our understanding of social meaning and the granularity with which it applies. An important question in the study of evaluation and its role in the incrementation of sound change, most recently addressed in Eckert & Labov (2017), is what objects of linguistic variation are actually subject to evaluation. Eckert & Labov argue, based on the Northern Cities Shift, that whilst evaluation can attach to the realisations of individual phonological units, it is blind to the more abstract components of linguistic variation such as phonological systems of chain shifts and mergers that concern the relationship between phonemes. This finds further support in the present study of VNP, where the overall alternation between [ɡ]-presence and absence (i.e. the concrete phonetic element) is beginning to accrue social meaning, but the more fine-grained change conditioned by pause and prosody is not.

These levels of granularity are represented diagrammatically in Figure 5, moving from more coarse-grained to more fine-grained from top to bottom: the most coarse-grained level is simply the phonetic [ɡ] sound, but as Eckert & Labov (2017: 481) point out, it isn’t the sound itself that is evaluated but rather “the use of that sound as the particular allophone representing a certain phoneme”. In this case, we see that it’s the use of [ɡ] in underlying /ŋɡ/ clusters, represented using traditional sociolinguistic variable notation as (ng), which accrues meaning. The most fine-grained level, the behaviour of (ng) in specific environments differentiated by segmental and prosodic factors, is also not subject to evaluation. In sum, these results suggest that social meaning is somewhat cruder than linguistic variation in production in that it applies to the presence/absence of particular sounds but not to the ways in which their variation is conditioned; it often attaches to some intermediate level that is neither too
coarse nor too granular.

While social evaluation may play a role in processes of linguistic change, this issue of granularity suggests that its ability to act as a driving force in the intergenerational incrementation of sound change is highly dependent on the variable in question (see Bermúdez-Otero forthcoming for further discussion of social evaluation and its limitations in accounting for patterns of sound change).

6 Conclusion

This study marked the first experimental investigation into the social meaning of (ng). Like many other sociolinguistic variables, our prior understanding of how listeners evaluate this dialectal form was speculative at best due to the reliance on inferences made from production studies and how patterns of variation can reflect sociolinguistic norms and standard language ideologies. These earlier studies argued that [ɡ]-presence is a local prestige variant (Mathisen 1999; Beal 2004), but the results of the matched-guise task reported here casts doubt on these claims.

The observed differences between young and old subjects suggests that VNP is a case of incipient social meaning, with evaluation present among young subjects but not so for the older generations. Members of these north western speech communities are becoming increasingly aware of this variable post-nasal [ɡ], specifically its status as a marker of northern dialects. However, although the increased salience of [ŋɡ] makes this variable more accessible to social evaluation, there is no consistency with respect to the polarity of responses; while some subjects rate [ŋ] as more professional, used as a proxy for overt sociolinguistic prestige, others actually rate [ŋɡ] more positively. Whether or not this lack of community-wide agreement is a characteristic feature of incipient social meaning (i.e. an early stage of developing shared sociolinguistic norms) remains to be seen, and can only be ascertained through further study of comparable variables. It may be that such disagreement only arises in the case of this variable where there are clearly antagonistic forces promoting both variants in the alternation: on the one hand [ŋɡ] is associated with northernness and, by association, decreased professionalism, but on the other hand it can also be perceived as a clear-speech variant closer to the orthographic norm, in part due to its frequency of occurrence in citation environments. If this lack of uniform evaluation is indeed characteristic of the early stages of social meaning more generally, it naturally follows that such disagreement should be ephemeral and that with time the scales should tip in favour of one form over the other. As such, it would be fruitful to return to this variable in future work and provide longitudinal insight into the
diachronic development of this evaluation.

The fact that listener attitudes towards [ɡ]-presence are not sensitive to contextual factors also indicates that the ongoing change taking place pre-pausally is not evaluation-driven; the implications of this finding are not limited to this particular variable, however, as it also lends further empirical support to recent claims from Eckert & Labov (2017) that not all objects of linguistic variation are accessible to social evaluation. Although linguistic variation is fine-grained at the level of production, the way in which it is subject to evaluation from listeners is relatively coarse. Listeners only attend to concrete aspects of variation, in this case the overall alternation between presence/absence of a particular segment; evaluation is blind to contextual factors despite their crucial role in conditioning the presence of [ɡ] in speech production, and by extension also blind to the ongoing change taking place in a subset of the environments in which post-nasal [ɡ] occurs. Consequently, the results support theories of sound change that foreground mechanical factors such as density of communication, in which social evaluation plays a more peripheral role.

Acknowledgements

I would like to thank a number of people who have provided insightful feedback during the development of this paper: Ricardo Bermúdez-Otero, Maciej Baranowski, the audiences at the 8th Northern Englishes Workshop at Newcastle University and the 2018 Manchester Forum in Linguistics, and the editors and reviewers at the Journal of Sociolinguistics. This work would also not have been possible without the financial support of the Economic and Social Research Council [NWDTC studentship ES/J500094/1], the subjects who took part in the survey, and of course my (very professional) newsreader. Any remaining errors are my own.
Bibliography


Paper IV

Three’s a crowd: Ternary variation in Northern English (ing)

George Bailey

University of Manchester

Abstract

Despite the attention (ing) has received in the variationist literature, it is comparatively under-studied in the North West of England. This is made all the more surprising by the fact that it holds a particularly unique sociolinguistic profile in this region, most notably in that the variation actually involves three competing forms: a non-coalesced [ɪŋɡ] variant appears alongside the more geographically-widespread [ɪŋ] and [ɪn]. This study uses a corpus of 32 sociolinguistic interviews conducted with speakers from this region to determine whether [ɪŋɡ] replaces [ɪŋ] as the local standard or exists alongside the plain velar nasal variant to fulfil a different sociocultural or functional role. The results suggest that [ɪŋ] is maintained as the standard variant, and that [ɪŋɡ] occupies its own functional space as a prosodic boundary marker and a feature of emphatic or hyper-correct speech. Furthermore, the overall rate of [ɪn] is extremely high, with no increase of velar forms for more nominal grammatical categories despite the strength of this effect in almost all other varieties of English. The three variants are also subject to phonetic analysis, which reveals that [ɪn] involves significantly less articulatory effort than the velar variants, with a laxer vowel and a shorter nasal segment; this suggests that (ing) is in fact a case of lenition, which has implications for predictions regarding its sensitivity to lexical frequency and other factors known to influence reductive phenomena.

Keywords: variation; velar nasal; ing; lenition; phonetics and phonology; dialects of English
1 Introduction

(ing) is widely believed to be one of the most consistent sociolinguistic variables with respect to its patterning along social and internal dimensions and its sociocultural meanings throughout the world’s varieties of English (Labov 1989). The variation between -in [ɪn] and -ing [ɪŋ] as a realisation of the -ing suffix (and of unstressed clusters in monomorphemic words such as during or ceiling) has been shown to be sensitive to a range of factors defined by sociological and linguistic properties, such as age, sex, social class, part of speech, and phonological environment (see Hazen 2006 for an overview).

In this paper, many of these predictors are tested in the North West of England using a corpus of 32 sociolinguistic interviews. Not only does this contribute to our knowledge of how (ing) behaves across different speech communities, it does so by investigating a region that serves as a particularly interesting site for such a case study. There are reasons to believe that (ing) holds something of a unique sociolinguistic profile in this region, as it is known to differ from its behaviour in other varieties of English in a number of ways. Firstly, (ing) is said to be much less variable in the North, with claims that the alveolar -in form is used ‘almost exclusively’ here (Labov 2001:90). Secondly, the North West in particular features a third possible variant -ingga [ɪŋɡ] – termed ‘velar nasal plus’ by Wells (1982b) – that is severely under-studied relative to the more geographically-widespread [ɪn] and [ɪŋ] forms.

The results of this study suggest that previous descriptions of [ɪŋɡ] as a ‘local prestige’ form (see Mathisen 1999; Beal 2004) do not quite capture this variant’s unique functional role as a marker of prosodic boundaries and a feature of emphatic or hyper-correct speech. Furthermore, this study shows that -in is indeed prominent in the North, so much so that this variable shows no sensitivity to grammatical category and exhibits only minimal differentiation along social dimensions.

This study also probes aspects of (ing) variation that have received comparatively less attention, namely the phonetic realisation of its variants and the extent to which they differ in ways not captured by a simply dichotomy between the nasal’s alveolar or velar place of articulation. Specifically, this paper focuses on two aspects of phonetic variation: the vowel quality of /ɪ/ and the duration of the following nasal. The results suggest that the alveolar and velar variants are not only differentiated by the nasal’s place of articulation: the vowel of -in is significantly laxer (higher F1 and lower F2) than that of the velar variants, and the nasal itself is shorter in duration. This points to a decrease in articulatory effort, serving as strong evidence that (ing) is actually a process of lenition, which in turn calls for a more nuanced treatment
of (ing) and its sensitivity to factors such as speech rate and lexical frequency, the latter of which has returned inconsistent results throughout the literature (see Forrest 2017, cf. Abramowicz 2007; Schleef et al. 2011). Furthermore, it highlights the need for future work to pursue dynamic articulatory analysis of this staple sociolinguistic variable.

2 Background

In the following subsections, a number of topics in the study of (ing) variation will be discussed to contextualise the findings presented in this paper.

In Section 2.1, I provide an overview of current theorising about the origin of synchronous (ing) variation, including regional differences in the actuation of this change. Section 2.2 covers the factors that are typically found to influence (ing) in contemporary varieties of English, and in Section 2.3 the focus turns to predictors known to influence [ɡ]-presence specifically. Finally, the sociolinguistic evaluation of this variable is discussed in Section 2.4.

2.1 Origin of (ing) variation

The diachronic account of how (ing) variation began in the history of English is now widely established and discussed in great detail by Visser (1966), who claims that synchronous variation in (ing) stems from the conflation of two distinct grammatical morphemes from Old English: the present participle -inde and verbal noun suffix -ung. The orthographic representation of these forms was subject to variation, particularly between different regional dialects, but crucially they differ consistently in the place of articulation of the nasal+stop cluster. Irwin (1967) claims that the feminine -ung verbal noun suffix was replaced by its masculine counterpart -ing (or -ynge), and that this was completed by the end of the 13th century.

During the Middle English period, a series of phonological reductive processes then took place, reducing the final /e/ to [a], followed by its subsequent deletion, and finally simplification of the consonant cluster resulting in the [ɪn] and [ɪŋ] forms that still exist today. In the North West and West Midlands of England, however, simplification of [ŋɡ] never ran to completion and still exists today as a probabilistic rule in speakers’ grammars: this results in three-way variation between [ɪn], [ɪŋ] and [ɪŋɡ] in words such as walking, during, morning etc. and two-way variation between [ŋ] and [ŋɡ] in words that invariably have [ŋ] in all other varieties, such as hang, wrong and singer (Wells 1982b; Garrett & Blevins 2009; Bermúdez-Otero 2011).

The functional expansion of the velar form into the territory previously occupied
by the apical form resulted in competition between the two for the same grammatical functions. According to Irwin (1967), who studied use of these forms between the 8th and 15th centuries, <ng> had replaced <nd> by the end of this period but the unidirectional nature of this change meant that there were no instances of <ng> being replaced with <nd> orthographically. Although Visser (1966: 1083) points out that it is difficult to pinpoint when the final /ŋ/ began to be pronounced as [n], rhyming pairs provide evidence that substitution of [m] for [ŋ] was common by the 18th century: Hodges (1643), for example, rhymes pairs such as coughing and coffin in his A Special Help to Orthographie, and John Rice also discusses how getting can be mistaken for get in his 1765 work Introduction to the Art of Reading with Energy and Propriety1, which he describes as a “vicious and indistinct Method of Pronunciation”. The overt stigmatisation of the apical form remains to this day, and will be addressed in Section 2.4.

2.1.1 Regional variation

It is important to note that this historical account of (ing) also suggests that the convergence of these two Old English suffixes was not uniform across all of England. As explained by Houston (1985, 1991), an isogloss from Moore et al. (1935) in their study of Middle English dialects indicates that by the middle of the 15th century -ind had been replaced with -ing first in the South of England and had been diffusing northwards, and Rooth (1941) also argues that graphemic replacement of <ind> with <ing> occurred earlier in southern dialects2. The fact that competition between the apical and velar forms began first in the South before diffusing geographically is said to result in patterns of regional variation that still exist today: rates of -in in contemporary British English dialects are supposedly highest in the North of England and in Scotland relative to other regions of the UK because this form existed unopposed as the verbal ending for much longer.

Labov (2001: 90) claims that (ing) shows exceptional behaviour in Southern States English, Northern British English and Scots, where “the /in/ form is used almost exclusively in speech, even of the most formal kind”. As will be shown later in the discussion of stylistic variation in Section 2.2.1, this statement is perhaps too strong, but there is a wealth of evidence to support the claim that rates of -in are remarkably high in these regions. Schleef et al. (2011) compare the behaviour of (ing) among Edinburgh and London adolescents, and find that the apical form is much more frequent

2There are claims that earlier processes involved in the development of (ing) as a sociolinguistic variable, such as the simplification of /nd/ to [n] in the participle ending, also occurred first in the South and the Midlands before the North (Poutsma 1923).

184
in the former (where it used around 75% of the time) than the latter (approximately 25%). Watts (2005) reports very high rates in Colshaw, a working class neighbourhood of Cheshire in the North of England, although not in the nearby community of Wilmslow; high rates have also been reported among young speakers in Manchester (Schleef et al. 2015). Finally, Houston (1985, 1991) shows how the afore-mentioned isogloss from Moore et al. (1935), delimiting the regions that first underwent \(-ind \rightarrow -ing\) replacement, lines up neatly with synchronic rates of the alveolar nasal variant: ‘northern/peripheral’ cities such as Liverpool, Gateshead, Glasgow and Edinburgh all exhibit high rates of \(-in\), relative to ‘southern/internal’ cities such as London, Essex, Bristol and Birmingham.

2.1.2 Vocalic differences

An outstanding question is why exactly this conflation of two previously distinct suffixes occurred first in the South of England. While this could simply reflect yet another example of northern conservatism\(^3\), Houston (1985, 1991) provides an alternative explanation, making reference to regional differences in vowel quality throughout the Middle English period.

It was mentioned earlier in Section 2.1 that both the participle and verbal noun endings exhibited a great deal of orthographic variation in the Old and Middle English periods. Crucially, differences in the pre-nasal vowel reportedly correlated with region, with the participle ending spelled \(-ind(e)\) in the South of England, \(-end(e)\) in the Midlands, and \(-and\) in Northern England (Jespersen 1954b: 378). Houston (1985, 1991) points to acoustic evidence from Habick (1980) and Woods (1979) that indicates how a high, front tense vowel could increase the likelihood of a following alveolar nasal being perceived as velar; assuming that the afore-mentioned orthographic differences were reflected phonetically, it provides a possible explanation as to why \(-ing\) first expanded into verbal contexts in the South, which had a high front vowel in the participle. This explanation falls under Ohala’s (1981, 1989) account of perceptual hypocorrection, which is widely accepted in theories of phonetic innovation as a powerful factor in the initiation of change.

Interestingly, Houston (1985, 1991) provides contemporary data from her British English speakers to suggest that regional differences in (ing) vowel quality still persist: speakers from ‘inside’ the isogloss (i.e. southern/internal regions) use more tense variants, while cities outside this area (i.e. northern/peripheral) use more lax or cen-

\(^3\)Wells (1997) provides discussion of how northern varieties of British English are, broadly speaking, historically conservative compared to those spoken in the South, referring specifically to their resistance to the **foot**-**strut** and **bath**-**trap** splits, the process of **ng**-coalescence, and the diphthongisation of long mid vowels.
Vocalic differences have also been reported in more recent work: there is a strong methodological focus in the study by Yuan & Liberman (2011), who explore the possibility of using computational methods to automatically identify variants of (ing), but an interesting theoretical outcome of this research is that the contrast between -in and -ing in American English is realised more on the preceding vowel than on the nasal segments themselves. They provide evidence in both production and perception. Using Kullback-Leibler divergence to compute phonetic distance between two acoustic models – one for [ɪn] and one for [ɪŋ] – they find more differentiation earlier in the models than later (i.e. the left portion of the vowel+nasal sequence). Perceptually, they also test native English and Mandarin speakers’ performance on coding (ing) variation: given that Mandarin has alveolar and velar nasal codas but no lax /i/ vowel, they hypothesise that these speakers should perform worse in coding (ing) if the contrast is truly realised more on the vowel than the nasal, and this is exactly what they report. On average across the 20 subjects, they find that Mandarin subjects are significantly worse than the native English subjects, and also worse than the automatic identification, in coding for -in and -ing.

Very few studies consider the phonetic quality of the vowel in (ing), but vocalic variation has been attested most notably in Canadian English. Woods (1979, 1999) reports centralisation to [ən] and tensing to [ɪn] in Ottawa, the latter of which Chambers (2009) claims is actually gaining traction in Canadian English. Walker (2016, 2017) also reports covariation between the vowel ([ə, ɪ, i]) and consonant ([ŋ, n]) in Toronto English, treating them as two entirely separate variables. Although Walker’s results indicate a trend towards a tenser vowel in -ing, the effect is not statistically significant: the vowel quality is influenced by social factors and by the preceding segment, but there is no relationship between the vowel and the following nasal. Similarly, Rosen et al. (2016) treat (ing) as a ternary variable in Canadian English, with [ɪn] as a distinct variant with its own social and linguistic profile.

Potential vocalic differences will be explored in this paper for contemporary dialects in the North West of England.

2.1.3 Nominal-verbal continuum

The historical account of (ing) is reflected in its synchronic variation not only with respect to the regional behaviour of this process but also its sensitivity to grammatical category. Widely described as the ‘nominal-verbal continuum’, this is one of the most consistent and robust predictors of (ing) variation in almost all dialects in which it has been studied. The patterning of this effect is such that rates of -in are relatively
higher in more verb-like tokens and relatively lower in more noun-like tokens. As Houston (1985, 1991) explains, this synchronic behaviour is a natural consequence of -in developing from a verbal suffix and -ing from a nominal suffix.

Given that this synchronic effect stems from historical morphology, it may come as no surprise that this contrasting behaviour of nouns and verbs is not a relatively new phenomenon. In fact, it was noticed in scholarly work as early as the late 19th century. Although Ellis (1889) makes no explicit reference to part of speech, it is likely that the difference described below stems from how the first token is a progressive use of (ing) but that the second token is gerundive:

“In the phrase ‘They were dansing and such dansing I never saw,’ note whether the two ings would be pronounced alike; they are sometimes different, and that is very characteristic.” (Ellis 1889: 10)

This effect has been attested in varieties of American English (Huspek 1986; Labov 2001; Abramowicz 2007; Forrest 2017), British English (Houston 1985, 1991; Tagliamonte 2004; Watts 2005), Australian English (Shnukal 1982), and New Zealand English (Bell & Holmes 1992), to varying degrees.

Studies of this grammatical category effect often make different distinctions, making comparisons difficult. For example, in an investigation of the Philadelphia speech community Labov (2001) reports a strong effect of part of speech, distinguishing the following categories: progressives, participles, complements, gerunds, nouns, and adjectives. However, the effect is more categorical than gradient in nature, with the former four categories clustering together (40-50% -in) and the latter three forming their own homogenous group (<10% -in). Huspek (1986) distinguishes only three categories, finding that verbs favour -in more than gerunds, which in turn exhibit higher rates than nouns. In contrast, Shnukal (1982) reports that (ing) in Australian English clusters as a binary opposition between verbs and nouns/adjectives.

In what is perhaps the most thorough investigation of this grammatical conditioning to date, Houston (1985, 1991) distinguishes fourteen different categories, and orders them on a scale of nominal to verbal types. The results indicate an interaction between grammatical category and region: the relative ordering of these categories is consistent across all communities (i.e. they all exhibit similar grammatical conditioning), but the strength of this effect is region-dependent, and intrinsically linked

4Labov (2001: 79) also claims that place names and proper nouns invariably surface with the velar nasal.

5These being progressives, quasi-progressives, VP complements, periphrastic future, appositive (non-finite) participles, pre-nominal adjunct modifiers, Acc-ing, gerundive nominals, action nominals, derived nominals, compounds, monomorphemics, proper names, and prepositions.
with the historical origin of (ing). Crucially, the afore-mentioned regional pattern of -in being used at higher frequencies in the north of England is largely down to higher rates specifically in the verbal categories. Of course verbal tokens still favour -in elsewhere in England but because of the increased use of -ing in verbal categories in the South – due to the encroachment of the velar form into these categories happening earlier here – the nominal-verbal distinction is less extreme than it is in the North.

Indeed, Tagliamonte (2004) reports a very strong effect in the northern community of York: among these speakers, grammatical category is in fact the strongest predictor of (ing) and this effect surfaces as a binary opposition between verbs/gerunds (77% and 71% -in, respectively) and nouns/adjectives (43% and 51% -in). In light of this contrast, Tagliamonte splits the data into verbal tokens and nouns, running separate analyses on each, and finds that they exhibit “entirely separate and unique linguistic and social profiles” (Tagliamonte 2004: 399). There are very few predictors influencing variation in verbal tokens, whereas variation in nominal (ing) on the other hand shows sensitivity to a range of social and linguistic factors; moreover, the results suggest that nominal tokens are still undergoing change in progress towards -in based on the evidence of a monotonic effect of speaker age. This is a particularly interesting result, given that (ing) has been reported to be in stable variation in every other community in which it has been studied quantitatively (Labov 2001: 86).

Tagliamonte (2004) suspects that (ing) will behave similarly in other northern British English varieties, but a review of subsequent research casts doubt on this: Watts (2005) reports only a small effect size for grammatical category in the northern community of Wilmslow⁶, and no effect whatsoever in nearby Colshaw. Furthermore, the grammatical conditioning reported by Schleef et al. (2011) for Edinburgh adolescents does not pattern along the nominal-verbal continuum at all; they also report no significant effect among their London participants, suggesting that this effect may not be as robust as first thought in varieties of British English.

### 2.2 Social and linguistic factors influencing (ing)

Not only is variation in (ing) remarkably widespread throughout the English-speaking world⁷, it is also striking in the degree to which its behaviour is comparable across all of these varieties.

Having addressed the role of grammatical category, in the following subsections

---

⁶Watts (2005) also carries out no statistical tests on this variation, making it difficult to ascertain how robust this effect is; the sample size is fairly large (N=2538), but the percentages only range from 15% -in in verbs to <5% -in in nouns (with adjectives placed in between at 10%).

⁷As pointed out by Hazen (2006: 583), the only native English-speaking population with no reported variation in (ing) is in South Africa (Gordon & Sudbury 2002: 81).
an overview is provided of the many other factors that are typically found to play a role in (ing) variation.

2.2.1 Socioeconomic status and style

Unsurprisingly for a variable involving a standard and a non-standard variant – and one that elicits so much meta-linguistic commentary, as will be discussed later in Section 2.4 – (ing) is often strongly stratified along socioeconomic lines. It also shows parallel stylistic variation, with the non-standard -in variant occurring more frequently in more informal, conversational styles. Comparisons are often drawn between the “gradient social stratification” found in New York City by Labov (1966), where -ing increases linearly with social class, and the “sharp social stratification” found in Norwich by Trudgill (1972, 1974), where the effect patterns as a binary distinction between the three working class groups and the two middle class groups. This is said to reflect sharper divisions of social class in British society (Labov 2001: 81). Trudgill (1972, 1974) also reports an interaction between social class and style, where the upper middle class speakers show the most extreme style-shifting, moving from near-categorical use of -in in conversational styles to near-categorical use of -ing in reading styles; this is attributed to ‘linguistic insecurity’ being highest among these speakers.

Regardless of the community-specific nature of these socioeconomic groupings, the overall pattern remains consistent: -in is favoured by speakers of lower socioeconomic status (see for example Mees 1977 and Coupland 1988 in Cardiff, Reid 1978 in Edinburgh, Petyt 1985 in West Yorkshire, Mathisen 1999 in the West Midlands, Labov 2001 in Philadelphia, and Watts 2005 in Cheshire). It has also been reported to a lesser extent in York, but overall Tagliamonte (2004: 400) observes that internal factors play a much larger role than social factors in this community and based on this claims that there is “very little socio-symbolic value attached to (ing)”.

This stylistic variation is also incredibly robust, having been attested by Fischer (1958) in the earliest quantitative study of (ing)\(^8\), more recently by Schleef et al. (2011) in London and Edinburgh, and in many of the afore-mentioned studies.

2.2.2 Gender and age

Sociolinguistic alternations involving an overtly-prestigious form typically also vary as a function of gender, with male speakers favouring the non-standard form over

\(^8\)Fischer (1958: 485) also reports that “the -ing frequency was higher in the beginning of the interview and later dropped off, presumably as the child became more relaxed and accustomed to the situation”, which also reflects the traditional view of style-shifting as attention paid to speech (Labov 1972a).
female speakers; (ing) is no exception, with this pattern attested throughout the sociolinguistic literature (Labov 1966; Trudgill 1972; Shopen 1978; Shnukal 1982; Mathisen 1999; Watts 2005). Fischer’s (1958) early work in rural New England, in which 24 children between the ages of 3 and 10 were interviewed, also reveals a male preference for -in, although there was inter-speaker variation within the male group. Interestingly, Fischer (1958: 484) notes that the ‘model’ boy who “did his school work well, was popular among his peers, reputed to be thoughtful and considerate” exhibited categorical rates of -ing, but the ‘typical’ boy who was “physically strong, dominating, full of mischief, but disarmingly frank about his transgressions” used much higher rates of -in. This study is often cited as one of the earliest in the variationist paradigm, but this observation in fact echoes much later work in the ‘third wave’ of sociolinguistics, in which the focus moves from macrosociological categories, particularly biologically-defined sex, and onto the social meaning of linguistic forms, the construction of personae, and the performative nature of variation (Eckert & McConnell-Ginet 1992).

With respect to speaker age, Labov (2001: 86) notes that no study has ever uncovered evidence that (ing) is involved in change in progress, although Tagliamonte’s (2004) later work in York reveals a monotonic function of age restricted to nominal tokens, which are exhibiting higher rates of -ing in apparent time. That is not to say that speaker age is irrelevant in the production of (ing) in other communities, however. Diachronically stable variables, at least those involving a form that is overtly evaluated as the ‘standard’ in competition with the ‘non-standard’ variant, typically exhibit an age-graded pattern (Holmes 1992; Sankoff 2005). In age-graded variation, speakers change their rate of use of a particular form throughout their life, but the community as a whole remains stable over time. This is typically reflected by a curvilinear function of age, such that the non-standard form is associated with both younger and older speakers, with a marked decrease for middle-aged speakers; this is said to reflect the strength of peer group pressure to conform to societal norms, and how this is greatest when adults are in the workforce (Eckert 1997).

This has been attested in a real-time study of (ing) in Philadelphia (Wagner 2012c). Although the participants were only interviewed a year apart, Wagner (2012c: 184) notes how this year “represents a time of social and developmental upheaval”: 17–19 year-olds are going to colleges and leaving their existing peer groups behind, and the results indicate that throughout this year most speakers exhibit the kind of sociolectal retrenchment typical of adulthood (Tagliamonte 2012). The minority of subjects who exhibited no such retreat from non-standard -in were those affiliated with a particular

---

9 This is not to be confused with lifespan change, where individuals change throughout their life in the direction of ongoing change in progress within the community (e.g. Sankoff & Blondeau 2007).
neighbourhood or social group where the social meanings indexed by this variant, such as masculinity, lower social class, and informality, are valued highly.

These findings build on earlier work attesting age-grading in (ing) for a younger cohort of African American Vernacular English speakers (Van Hofwegen & Wolfram 2010). It has also been reported for British English in the northern community of Wilmslow, where rates of -in are highest among 12–18 year-olds, although Watts (2005) puts this down to dialect contact with school children from Colshaw who predominantly use -in. Other studies of northern communities, such as Manchester (Schleef et al. 2015) and Edinburgh (Schleef et al. 2011), focus only on adolescents, so it is unclear how strongly age-graded (ing) is in this predominantly -in-favouring region.

2.2.3 Phonological environment

Moving on to internal factors, there has been some debate regarding the sensitivity of (ing) to the immediate phonological environment. Although Labov (2001: 87) claims that there is no such effect, a number of other studies report strong and, crucially, consistent patterns between the rate of -ing and the segments that both precede and follow the [ɪ]+nasal sequence.

Houston (1985: 19) reports two types of phonological conditioning: regressive homorganic assimilation, in which a following velar consonant favours use of -ing and a following alveolar consonant favours -in, and progressive homorganic dissimilation, in which a preceding velar consonant favours -in and a preceding alveolar consonant favours -ing. In other words, the preceding and following segmental material surrounding (ing) pattern in opposite ways to each other, and this has been corroborated in a number of other empirical studies of this variable (e.g. Shuy et al. 1968; Cofer 1972).

This was in fact noted much earlier: Walker (1791: 49), in his Critical Pronouncing Dictionary, states that [ɪn] should be avoided when the velar nasal is also present in the root (e.g. singing, bringing), as such a pronunciation would have “a very bad effect on the ear” due to the repeated [ŋ]. Conversely, he also argues that words with [n] in the root, such as winning, shining etc., should not use -in in the suffix because it results in “the same disgusting repetition”.

There is, however, recent evidence to suggest that this phonological conditioning is not as widespread as first thought. Using a large corpus of speech from Raleigh, North Carolina, Forrest (2017) finds that a following velar consonant does favour -ing at a community-wide level10, but that analysis of individual speakers reveals a great

10Forrest (2017) also finds that words which frequently occur before velar consonants tend to favour
deal of variation, with many speakers actually showing the opposite pattern.

The effect also appears to be somewhat less widespread in varieties of British English. Tagliamonte (2004) reports only a weak effect in York – with a preceding velar consonant promoting use of -in but a following velar favouring -ing – and crucially this effect is present only for nominal tokens, not verbs. Schleef et al. (2011) also report that -ing is more likely before velars and after apicals in London, but this effect is absent in Edinburgh. Finally, although Watts (2005) observes higher rates of -ing in Cheshire when the following consonant is velar, the highest rates of -in are actually found before pauses/vowels, not before homorganic alveolar consonants.

2.2.4 Lexical frequency

In exemplar-theoretic models of grammatical architecture, it has been argued that high frequency words lead sound change due to increased exposure to phonetic bias (Bybee 2002; Hay & Foulkes 2016), although Dinkin (2008) suggests that this is only the case in changes involving lenition.

The extent to which [ɪŋ] →[ɪn] can be considered lenition is debatable given that it is not obviously a case of deletion or segment weakening. However, Abramowicz (2007) argues that production of -in does involve less articulatory effort, and the vocalic differences discussed in Section 2.1.1 indicate that -in is supposedly often produced with a laxer, more centralised vowel. Of course, it is also not involved in change in progress, although frequency can still play a role in stable variation: while usage-based theories may account for frequency effects in sound change by making reference to exemplar clouds and phonetically-rich representations, sensitivity to frequency in stable variation may simply reflect hyper-/hypoarticulation (Lindblom 1990), i.e. a tendency for gestural reduction in more predictable, higher frequency words (Hooper 1976; Bybee 2000; Jurafsky et al. 2001; Aylett & Turk 2006).

Lexical frequency effects are not well-explored in the literature on (ing), but there are suggestions that ‘everyday’ words tend to surface more often with -in relative to ‘specialised’ or ‘learned’ words, which could reflect sensitivity to lexical frequency (Wald & Shopen 1981: 225; Tagliamonte 2004: 399). Abramowicz (2007) and Schleef et al. (2011) investigate frequency effects more explicitly, in Philadelphia and Edinburgh/London respectively, but both report null results.

To this author’s knowledge, the only explicit quantitative evidence of (ing) exhibiting sensitivity to lexical frequency is reported by Forrest (2017) in Raleigh, North Carolina. These contradictory reports may simply reflect community-specific be-
haviour of (ing), or may stem from methodological differences: corpus size differs dramatically between the three afore-mentioned studies (573 tokens in Abramowicz 2007, 3389 tokens in Schleef et al. 2011, and 13167 tokens in Forrest 2017), and different measures of frequency are also employed, which has been shown to have a strong influence on the observed effect (Fruehwald 2017a).

2.3 Constraints on [ɡ]-presence

Up to this point the preceding review of literature on (ing) has focused exclusively on the widespread variation between -in and -ing, in part because this reflects the wealth of existing knowledge regarding this alternation. The factors that condition the variable presence of [ɪŋɡ] in the North West and West Midlands regions of England are comparatively under-studied.

The study conducted by Schleef et al. (2015) is largely perceptual in nature, but patterns of variation in production are discussed briefly: based on nine sociolinguistic interviews conducted with 18–20 year-olds in Greater Manchester, they report that -ingg is only present in conversational speech to a marginal extent, with no speaker using this form more than 10% of the time. The frequency with which it appears increases in the more formal reading task, but even then it only occurs about 20% of the time across all speakers (comparable to the rate of -ing in this style). Schleef et al. (2015) argue that this supports the idea that -ingg is a local prestige form, citing an earlier study by Mathisen (1999) who shows that this local variant exhibits parallel stylistic stratification in the West Midlands.

Watts (2005) echoes these observations in Cheshire, where rates of -ingg in conversational styles are as low as 1.6%. It is used slightly more frequently in Wilmslow, the neighbourhood of higher socioeconomic status. Watts (2005) also cites a number of examples in which [ɪŋɡ] is used for emphatic effect, and based on this draws parallels with the claim by Petyt (1985) that in West Yorkshire this variant only surfaces from “extreme carefulness”.

In most cases, the authors conclude that [ɪŋɡ], where present, has local prestige.

2.4 Evaluation of (ing)

In light of the way that (ing) patterns along stylistic and sociodemographic lines, as discussed in Section 2.2, it is fairly uncontroversial to describe this variable as a highly salient ‘stereotype’ that elicits a great deal of overt metalinguistic commentary (Labov 1972a).

Indeed, perceptual experiments that uncover listener evaluations of (ing) variation have revealed exactly this. Matched-guise tasks, in which listeners are played
samples of speech differing only in the realisation of (ing), indicate that use of the -in form in varieties of American English is seen as less professional, less formal, and less educated (Labov et al. 2006, 2011; Campbell-Kibler 2006, 2011b). The attitudes uncovered by these studies are explicit and often reflected in metalinguistic commentary, even dating back as far as the 18th and 19th centuries11.

A later replication of the Labov et al. (2006, 2011) matched-guise studies in British English suggest that this evaluative behaviour is not universal, however. Levon & Fox (2014) find no decrease in perceived professionalism with increased use of -in in formal contexts, despite (ing) being subject to the same degree of prescriptivist discussion as it is in the United States. Based on this, they claim that (ing) lacks the same “social prominence” and that it is associated more with region than social class in the United Kingdom (Levon & Fox 2014: 187).

The indexical field of (ing) is further complicated in northern England by the presence of [ɪŋɡ] which, based on the claims discussed in Section 2.3, would be expected to pattern alongside [ɪŋ] in indexing prestige and formality. However, Schleef et al. (2015) dispute this with direct perceptual evidence: while they find an association with [ɪŋɡ] and prestige, they argue that it is not as socially attractive as the more widespread [ɪŋ] form and that it is associated with an “unenergetic, uptight attitude towards life” (Schleef et al. 2015: 207).

3 Methodology

This study uses a corpus of 32 sociolinguistic interviews to conduct a quantitative analysis into how (ing) patterns in the North West of England. In the following subsections, information is provided on the interviews themselves, the participants, and how the dependent and independent variables were coded.

3.1 Sociolinguistic interviews

The interviews were all carried out between 2015 and 2017, with the exception of two that took place in Manchester in 1971, which are included here to provide extra depth to the population sample with respect to date of birth. Based on methodological guidance from Tagliamonte (2006: 37-49) and from Labov’s (1984) Philadelphia Neighbourhood project, the interviews were conducted one-on-one with the participants, and on average lasted for around an hour. They consisted of spontaneous,

11Historically, the prescriptivist commentary on (ing) is somewhat more complicated than a blanket dismissal of the alveolar -in form: Bailey (1996) claims that -in was used by ‘landed gentry’ as well as the working class even into the 19th century, and that to this date huntin’, shootin’ and fishin’ is stereotypical of traditional aristocratic speech.
unscripted conversation covering topics such as childhood, family life, career, and holidays. Efforts were also made to elicit narratives of personal experience; such an informal style of discourse and unscripted ‘story-telling’ is said to provide the most direct access to a speaker’s vernacular – when the effects of the observer’s paradox are most mitigated – and therefore a more reliable insight into the linguistic variability present in the dialect (Labov 2010b).

The informal conversation was followed up by two elicitation tasks: firstly, a 108-item word list containing tokens of (ing) amidst a number of distractor items, and secondly a short 239-word reading passage, also containing tokens of interest in various phonological and prosodic environments. These are given in Appendix A and Appendix B, respectively.

The interviews were recorded using a Sony PCM-M10 recorder and a lavalier microphone attached to the participant, saved as an uncompressed WAV file at a 44.1KHz sampling rate.

3.2 Participants

These interviews were carried out with 32 speakers (17F; 15M) covering a range of ages from 17 to 83, and with dates of birth spanning almost a century from 1907 to 1998. The participants were all born and raised in the North West of England, and in many cases did not move from the town/city where they spent their childhood years. Although the speakers cover a broad geographic region, a large majority were born and raised in either Greater Manchester or Blackburn and their surrounding areas.

In order to have a balanced sample without introducing a confound of social class, socioeconomic status was controlled for by only interviewing upper working class speakers, with this classification based primarily on occupation (see Baranowski 2017: 303 for a similar operationalisation of social class in Manchester).

3.3 Data annotation and statistical methods

The interviews were orthographically transcribed in breath groups – loosely corresponding to an utterance (i.e. a stretch of speech culminating in the speaker pausing to draw breath) – using ELAN and later force-aligned using the FAVE suite (Rosenfelder et al. 2011) to produce a word- and phone-level TextGrid for each speaker.

The dependent variable was coded manually, based on a combination of auditory and acoustic analysis using Praat (Boersma & Weenink 2017). Each token was coded for whether the surface realisation was [ɪn], [ɪŋ], or [ɪŋɡ]. Without wanting to impose any arbitrary grouping or make any a priori assumptions about how the different segments might pattern, the phonological environment was coded in the
most granular way possible: by annotating each token for the exact segment that precedes and follows the [ɪ]+nasal sequence. In the results section, any groupings will be empirically-grounded and based on the extent to which segments pattern together. The following independent variables are tested:

- **AGE**: the age of the speaker.

- **SEX**: the sex of the speaker.

- **GRAMMATICAL CATEGORY**: coded manually based on a five-way distinction between nouns, pronouns, adjectives, gerunds, and progressive verbs.

- **MORPHOLOGICAL COMPOSITION**: whether the word is monomorphemic (e.g. *ceiling*, *pudding*), polymorphemic (any word with an -ing suffix), or a -thing compound (e.g. *anything*, *something* etc.)

- **PRECEDING SEGMENT**: the segment immediately preceding the unstressed [ɪ] vowel of the (ing) token.

- **FOLLOWING SEGMENT**: the segment immediately following the nasal of the (ing) token; for word-final tokens, this is either the initial segment of the following word or coded as a pause if at least 100ms of silence intervenes between the two words.

- **SPEECH RATE**: calculated for each token using a Praat script that divides the duration of each token’s breath group by the number of words, syllables, and segments it contains\(^\text{12}\).

- **LEXICAL FREQUENCY**: frequency was operationalised as a log-transformed Zipf score on a continuous scale from 1–7, based on frequency counts from the SUBTLEX-UK corpus of UK television subtitles; this has been found to be the most robust measure of lexical frequency in psycholinguistic priming studies (van Heuven et al. 2014). Comparisons between this measure and a traditional measure of frequency per million words are given in Table 1.

- **STYLE**: whether the token was produced in the spontaneous conversational portion of the interview, or in the reading passage or word list elicitation tasks; conversational tokens are further categorised into narrative and non-narrative, based on whether or not the speaker was describing a narrative of personal experience (Labov 2010a).
Table 1: Comparison between Zipf score and frequency per million words, with example (ing) words.

<table>
<thead>
<tr>
<th>Zipf score</th>
<th>Frequency per million words</th>
<th>Example words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.01</td>
<td>stewarding, dislocating</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>misbehaving, rustling</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>printing, draining</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>dining, catching</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>having, seeing</td>
</tr>
<tr>
<td>6</td>
<td>1,000</td>
<td>something, going</td>
</tr>
<tr>
<td>7</td>
<td>10,000</td>
<td>absent in data</td>
</tr>
</tbody>
</table>

In addition to the independent predictors listed above, two phonetic properties of the (ing) variants are investigated: the duration of the alveolar/velar nasal, and the acoustic quality of the unstressed [ɪ] vowel preceding these nasals.

Nasal duration was automatically measured in milliseconds using a Praat script, based on the segment boundaries as determined by FAVE. In order to conduct an investigation into potential vocalic differences in the unstressed [ɪ] vowel, the interviews were also processed using FAVE-extract to perform automatic vowel extraction, resulting in formant frequencies for over 280,000 vowels. While FAVE-extract uses specific measurement points for certain vowels, such as PRICE /aɪ/ and MOUTH /aʊ/, in the case of the vowels of interest here – namely those along the front diagonal, [ɪ], [i], and [ɛ] – the formants are extracted one third of the way through the vowel. The formant measurements were normalised within-speaker using the Lobanov (1971) method to account for inter-speaker differences in absolute formant frequencies, and then transformed back to Hz for ease of interpretation. FAVE-extract automatically extracts the segments immediately preceding/following each vowel, and the syllabify package (Fruehwald 2018) was used for each vowel to determine whether or not it is tautosyllabic with the following consonant.

In total, the interviews contain 5222 tokens of (ing), 4622 of which occur naturally in the informal conversation. All analysis was carried out in R (R Core Team 2016; Wickham 2017). The lme4 package (Bates et al. 2015) was used for fitting mixed-effects logistic regression models to the (ing) dependent variable and mixed-effects linear regression models by Restricted Maximum Likelihood estimation (REML) for the continuous phonetic dependent variables (i.e. nasal duration and vowel quality). In both cases, best-fitting models were reached by fitting a full model including all

---

\[12\] Syllable and segment counts are automatically calculated based on canonical pronunciations from the Carnegie Mellon University pronouncing dictionary, available at http://www.speech.cs.cmu.edu/cgi-bin/cmudict.
relevant predictors followed by a combination of automatic step-wise regression – in which single predictors are dropped at a time – and manual ANOVA comparison between nested models to diagnose significant differences in model fit as measured by AIC. The Satterthwaite approximation was used for estimating degrees of freedom and obtaining \(p\)-values for individual estimates in the linear models (Luke 2017).

4 Results

Given that in the North West of England (ing) variation surfaces as a three-way alternation between \([\text{im}]\), \([\text{in}]\) and \([\text{ing}]\), the results are split into two parts: Section 4.1 addresses the factors that influence the rate of alveolar \([\text{im}]\) relative to the two velar forms, and Section 4.2 excludes these alveolar tokens to focus directly on the factors that influence variable [g]-presence within the velar forms.

Before these two aspects of (ing) variation are investigated in closer detail, the overall distribution of variants for all speakers in this sample is given in Figure 1: this broad overview of the frequencies at which each form appears indicates that -\text{ingg} is particularly marginal in conversational styles and that -\text{in} is extremely frequent.

4.1 Variation between -\text{in} and -\text{ing(g)}

Despite the overall predominance of -\text{in}, there is clearly inter-speaker variation with respect to the relative frequencies of alveolar-velar forms: are these differences ex-
plained by social factors, and does the variation within individual speakers reflect internal constraints on this variable? The best-fitting model for the variation between -in and -ing(g) includes significant main effects of speaker age and gender, preceding and following phonological environment, speech rate, and style; it also includes random intercepts of speaker and word, and by-speaker random slopes of style and following segment based on evidence of inter-speaker variation with respect to the magnitude of style-shifting and how different phonological environments pattern. The model coefficients are given in full in Table 2.

Firstly it can be seen that social factors do influence (ing) variation in the North West, but that they play a very minor role. The probability of -in over -ing(g) is slightly higher for older ($\beta = 0.60$) and younger ($\beta = 0.48$) speakers relative to the middle-aged
Figure 2: The rate of -in by speaker age and gender. In (a), points reflect speaker means with curves fitted to individual tokens using locally-weighted (LOESS) smoothing. In (b), points reflect age group means.

cohort (between 27 and 60 years old), and male speakers are also slightly more likely to use -in than female speakers ($\beta = 0.41$). However, as shown by Figure 2, this gender differentiation is only present for the two older cohorts of speakers and the estimates from Table 2 suggest that both of these external predictors have very little influence on the variation.

There is also evidence of sensitivity to the immediate phonological environment. The effect of the preceding segment is similarly small in magnitude, with a preceding alveolar consonant (e.g. in batting, passing etc.) decreasing the likelihood of -in ($\beta = -0.48$), but there is no evidence that a preceding velar consonant has the opposing effect. That is, there is only partial evidence of the word-internal dissimilatory effects discussed in the previous literature on this variable.

Variation in (ing) shows much greater sensitivity to the following segment, which is in fact the strongest predictor for conversational tokens. This effect is best modelled by making a distinction between three levels of ‘assimilation strength’. A following velar consonant (i.e. [g] or [k]) greatly decreases the probability of -in ($\beta = -3.22$), and there exist other segments that exert a similar influence on (ing) to a smaller but still significant degree: -in is also less likely to appear before [j], [w], and [l], and as such these constitute the ‘weak’ assimilation category ($\beta = -0.99$). This latter effect is only evident for some speakers, however, thus motivating the random slope of assimilation strength by speaker. Figure 3 illustrates this bimodality between the weak and strong assimilation contexts, with average rates of -in between 90% and 72% before [l], [w] and [j], and around 40% when followed by either of the velar stops. Importantly, while
the community-wide behaviour before [l] may point to the absence of an effect, there are a number of speakers for whom [l] exerts one of the strongest influences on the use of -ing(g).

Speech rate is the only other factor to which conversational (ing) is sensitive: -in is more likely to be used in faster speech rates ($\beta = 0.40$), which for this dataset is best modelled by a continuous measure of segments per second\textsuperscript{13}. However, as illustrated by Figure 4, the magnitude of this effect is again fairly small: the use of -in does not drop below approximately 75% even in the slowest speaking rates.

The only other significant factor influencing the use of -in is style; however, this will be discussed in Section 4.3 where the stylistic distribution of all three variants is investigated in more detail.

Among the non-significant factors, it is important to draw attention to grammatical category. Despite the wealth of evidence that this plays an important role – quite often the most important role – in (ing) variation elsewhere throughout the English-speaking world, no such effect is present here. Lexical frequency is also absent from the best-fitting model, although it is only marginally beyond the accepted alpha level ($p = 0.053$) and is in fact highly significant if we adopt a slightly less conservative

\textsuperscript{13}Although operationalising speech rate as segments per second provides the best fit to the variation, the effect is still significant and patterns in the same direction when measured by words or syllables per second instead.

---

**Figure 3:** Rate of -in by (a) preceding segment and (b) following segment. Shaded circles reflect individual speaker means; diamonds reflect overall mean.
approach and exclude the random intercept for word in the model. This suggests that while frequency may play some role, the effect is particularly small and its explanatory power may well be absorbed by the potential for word-specific effects in the model.

4.2 Variation between -ing and -ingg

As has already been shown in this results section, the velar forms [ɪŋ] and [ɪŋɡ] are relatively infrequent in naturally-occurring conversation, and the latter variant – the under-studied form exclusive to the North West and West Midlands of England – is in fact almost entirely absent, occurring in just 32 of 4622 tokens. However, if we assume that the variation stems from two separate processes – one an alternation between -in and -ingg, and the other a deletion of post-nasal [ɡ] to derive -ing – this means we should actually be looking at rates of -ingg relative only to -ing in order to determine the factors that influence this variable [ɡ]-presence. As such, the analysis in this section excludes all tokens of -in, such that the relevant alternation is between [ɪŋ] and [ɪŋɡ].

The best-fitting model for this variation includes main predictors of preceding and following segment, speech rate, and style; the random effect structure includes random intercepts for speaker and word, and a random by-speaker slope for style to reflect inter-speaker variation with respect to the degree of style-shifting. The model coefficients are given in full in Table 3, where [ɡ]-presence is the application value; that is, positive log-odds represent a factor level favouring [ɪŋɡ].

First of all it can be seen that this model is much simpler than that which predicted
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>St. Error</th>
<th>z-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTERCEPT</strong></td>
<td>-4.2591</td>
<td>0.5732</td>
<td>-7.430</td>
<td>&lt;0.001  ***</td>
</tr>
<tr>
<td><strong>PRECEDING SEGMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other consonant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alveolar consonant</td>
<td>-0.1596</td>
<td>0.3685</td>
<td>-0.433</td>
<td>0.6648  *</td>
</tr>
<tr>
<td>velar consonant</td>
<td>-1.1428</td>
<td>0.4912</td>
<td>-2.326</td>
<td>0.0199  *</td>
</tr>
<tr>
<td><strong>FOLLOWING SEGMENT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>consonant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pause</td>
<td>2.6833</td>
<td>0.4956</td>
<td>5.414</td>
<td>&lt;0.001  ***</td>
</tr>
<tr>
<td>vowel</td>
<td>1.9015</td>
<td>0.3935</td>
<td>4.832</td>
<td>&lt;0.001  ***</td>
</tr>
<tr>
<td><strong>SPEECH RATE</strong> (scaled)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>segments per second</td>
<td>-0.8443</td>
<td>0.2299</td>
<td>-3.673</td>
<td>&lt;0.001  ***</td>
</tr>
<tr>
<td><strong>STYLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conversation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reading passage</td>
<td>2.6155</td>
<td>0.6162</td>
<td>4.245</td>
<td>&lt;0.001  ***</td>
</tr>
<tr>
<td>word list</td>
<td>2.7120</td>
<td>0.7197</td>
<td>3.768</td>
<td>&lt;0.001  ***</td>
</tr>
</tbody>
</table>

Table 3: Mixed-effects logistic regression model for velar (ing) tokens; [g]-presence as the application value (968 observations, AIC: 656.3; deviance: 626.3).

The occurrence of -in: the conversational variation is only sensitive to three factors, contrasting with the five that were relevant in the earlier discussion. It is also striking that no social factors are significant. The role of social factors in promoting use of -in was small, but here it is non-existent.

Given that the phonological conditioning of variable [g]-presence does not involve assimilation, the relevant grouping of segments is different than in the earlier regression model. Instead of grouping segments by the strength of their assimilatory influence, a distinction is made between a following pause, vowel, and consonant. As Table 3 indicates, when (ing) surfaces with a velar nasal it is more likely to exhibit [g]-presence when it occurs before a vowel-initial word ($\beta = 1.90$), and even more likely when it occurs before a pause ($\beta = 2.68$). The pre-pausal environment has by far the largest influence on [g]-presence in conversational styles. There is also a weaker effect imposed by the preceding segment, such that the post-nasal [g] is less likely to surface (in fact it is categorically absent) if the segment preceding [ɪ] is also a velar consonant such as in thinking or singing ($\beta = -1.14$). Both of these effects are illustrated in Figure 5.

Variation between -ing and -ingg is also sensitive to prosody: among velar tokens, [g]-presence becomes increasingly unlikely as the speech rate increases ($\beta = -0.84$).
This is illustrated in Figure 6.

As with the variable occurrence of -in, there is no significant effect of grammatical category – although in this case there would be no reason to expect one – and also no effect of word frequency. A further parallel can be drawn in that the strongest overall predictor is once again the formality of discourse. The stylistic stratification of all three forms will now be addressed in the following section.

### 4.3 Style-shifting

The estimates of the two logistic regression models discussed in the previous section indicate that (ing) exhibits a large degree of style-shifting. Figure 7 illustrates the way that these three variants pattern along four different speech styles, distinguishing narrative speech from informal conversation within the spontaneous part of the interviews, and then reading passage and word list styles within the formal elicitation tasks.

It can be seen that -in decreases monotonically along these four styles, which are themselves ordered by increasing level of formality. The only exception to this pattern is the lack of any difference between narrative and non-narrative conversational styles, which pattern together with respect to all three variants. The rate of -ing increases from conversation to reading passage, as one might expect of an overtly
Figure 6: Rate of [g]-presence by speech rate (segments per second); curve fit to individual tokens using locally-weighted (LOESS) smoothing.

Figure 7: The rate of -in, -ing and -ingg by discourse style. Crosses reflect individual speaker means; circles reflect overall means.
Figure 8: The distribution of nasal duration (log-transformed from milliseconds) by (ing) variant.

standard form, but then we see a decrease from the reading passage to the word list; this is because -ing, which is almost entirely absent from the conversation, shows a marked increase in word list style. Of course, it is also more frequent in the reading passage relative to spontaneous conversation, but it is undoubtedly most strongly associated with the word list register of speech.

4.4 Phonetic properties

In this section, attention turns away from the predictors that influence the choice of form and instead I investigate the actual phonetic realisation of the (ing) variants, which is somewhat under-researched. They obviously differ with respect to the place of articulation of the nasal segment, but in addition to this do they differ along other phonetic dimensions? Specifically, this section addresses apparent differences in nasal duration and vowel quality between the alveolar and velar variants.

4.4.1 Nasal duration

As illustrated by Figure 8, there appears to be an observable and statistically-significant difference between the duration of the alvolar nasal in [ɪn] and the velar nasals in [ɪŋ] and [ɪŋɡ]: the former has an average nasal duration of 68ms, compared to the latter’s 110ms ($t = -11.852, p < 0.001$).

However, there is a possible confounding factor involved in this apparent relationship: it has already been shown that -in is more frequent in faster speech rates, which in turn could result in these forms being articulated more often with shorter overall segmental duration. Furthermore, segmental duration is also sensitive to prosody,
with words that appear finally in major prosodic constituents (e.g. the utterance or intonational phrase) undergoing pre-boundary lengthening (Lehiste et al. 1976; Gussenhoven & Rietveld 1992; Wightman et al. 1992). In sum, this contrast in nasal duration may in fact reflect other mechanisms, given that it has already been shown how -in is more likely to occur in faster speech rates and how [ɡ]-presence is boosted phrase-finally.

To tease apart these collinear factors, Figure 9 plots the relationship between nasal duration and speech rate for all tokens of (ing), crucially excluding all tokens appearing in pre-pausal position. There are two things to observe: unsurprisingly, there is a moderately strong inverse correlation between nasal duration and speech rate ($r = -0.348$, $p < 0.001$), but importantly it can be seen that the velar variants consistently surface with longer nasals independent of the rate of speech. The only exception to this pattern is for the very fastest speaking rates, where the contrast in nasal duration appears to be somewhat neutralised.

This phonetic behaviour was modelled using mixed-effects linear regression, with log-transformed nasal duration as the dependent variable; the best-fitting model, reached by following the same methods discussed earlier, includes the following predictors: (ing) variant (i.e. alveolar or velar), speaker gender, following segment, and speech rate. The random effect structure consists of random intercepts for speaker and word, and two by-speaker random slopes for (ing) variant and following segment. The full coefficients table is given in Table 4. As expected, the duration of the (ing) nasal segment is longer for -ing(g) relative to -in ($\beta = 0.40$), and it also increases in duration when it occurs in pre-pausal position ($\beta = 0.49$). It also comes as no surprise...
Table 4: Mixed-effects linear regression model for nasal duration (log-transformed from milliseconds) (4563 observations, AIC: 6442.1; deviance: 6386.1; log likelihood: -3193.1).

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>St. Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>4.110</td>
<td>0.0357</td>
<td>115.247</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>(ING) VARIANT</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>-in</td>
<td>(reference level)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ing(g)</td>
<td>0.4007</td>
<td>0.0274</td>
<td>14.651</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>female</td>
<td>(reference level)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>-0.2367</td>
<td>0.0436</td>
<td>-5.426</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FOLLOWING SEGMENT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>other consonant</td>
<td>(reference level)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alveolar consonant</td>
<td>-0.0046</td>
<td>0.0243</td>
<td>-0.189</td>
<td>0.8505</td>
</tr>
<tr>
<td>velar consonant</td>
<td>-0.0583</td>
<td>0.0368</td>
<td>-1.583</td>
<td>0.1214</td>
</tr>
<tr>
<td>pause</td>
<td>0.4900</td>
<td>0.0442</td>
<td>11.098</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>vowel</td>
<td>-0.0431</td>
<td>0.0249</td>
<td>-1.733</td>
<td>0.0944</td>
</tr>
<tr>
<td>SPEECH RATE (scaled)</td>
<td></td>
<td></td>
<td></td>
<td>***</td>
</tr>
<tr>
<td>segments per second</td>
<td>-0.1467</td>
<td>0.0085</td>
<td>-17.292</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

that nasal duration decreases in faster speech rates (β = -0.15), although the significant effect of speaker gender – with male speakers producing significantly shorter (ing) nasals relative to female speakers (β = -0.24) – is somewhat more difficult to account for.

Due to the logarithmic transformation applied to the dependent variable, it is not straightforward to interpret the estimates in absolute terms; however, it can clearly be seen that with respect to effect size, the increase in nasal duration by virtue of being a velar (ing) form is almost as large as the pre-boundary lengthening effect registered in the following segment predictor.

4.4.2 Vowel quality

Having established a difference in nasal duration, attention now turns to the pre-nasal vowel. This is broadly transcribed as unstressed [ɪ], but the discussion in Section 2.1.2 highlighted reports of variation in the exact phonetic quality of this vowel.

Is there any systematic relationship between the vowel of the (ing) variant and its following nasal? The vowel formants for all conversational (ing) tokens, plotted in
Figure 10: Normalised vowel plot for all conversational (ing) tokens, colour-coded by type of variant. Shaded circles reflect mean F1 and F2, and ellipses reflect 95% confidence intervals. Mean formant values given for /iː/ as a baseline for comparison.

Figure 10 seems to suggest so: the velar (ing) variants clearly surface with a tenser, more [iː]-like vowel, although there is still a great degree of overlap between the two clouds of tokens and the average formant values for -ing(g) are clearly still distinct from the fleece vowel itself. The quality of this vowel is more accurately transcribed as [ɪ̟], given that this difference is mainly registered in F2 (a difference of 205Hz) as opposed to F1 (19Hz), though crucially the difference is significant along both dimensions (for F2: $t = -13.075, p < 0.001$; for F1: $t = 4.8346, p < 0.001$). Importantly, this effect does not arise due to differences in vowel duration; while the tenser vowels found in -ing(g) mirror the nasal segment in being longer than their counterparts in -in, this is only a difference of 4ms on average and is therefore unlikely to be enough to account for the tenser realisation. That is, the vowel in -in is not laxer because it is for some reason inherently shorter than the vowel in -ing(g).

Mixed-effects linear regression also indicates that this difference in vowel quality is indeed robust. A model was fit with a dependent variable of front diagonal, a composite measure of movement along the front diagonal in both F1 and F2. Following Labov et al. (2013: 40), this parameter was calculated as $F2 - (2 \times F1)$, such that higher values equate to higher position along the front diagonal (i.e. a lower F1 and a higher F2) and lower values to a lower position (i.e. higher F1 and lower F2). The best-fitting model contains main effects of (ing) variant (i.e. whether or not the token was transcribed as -in or -ing(g)), preceding segment (velar vs. other), vowel duration, and speech rate. Random intercepts of speaker and word, and a random
Table 5: Mixed-effects linear regression model for front diagonal of \( /i/ \) in (ing) (4079 observations, AIC: 58457.5; deviance: 58437.5; log likelihood: -29218.8).

by-speaker slope of (ing) variant, were also included.

The model summary is given in Table 5. The strongest predictor is in fact (ING) VARIANT, with an increase of 173Hz on the front diagonal when the token is velar instead of alveolar; a preceding velar consonant imposes a similar effect, although in this case the increase is smaller at 115Hz. There are also two continuous predictors of this variation: as speech rate increases, and vowel duration decreases, the vowel is lower on the front diagonal.

At this point it is important to determine whether this difference is inherently tied to the different types of (ing) variants (i.e. alveolar or velar) or if it actually reflects a more general coarticulatory effect of velar consonants on vowel quality in all types of VC sequences. Certainly, the fact that the preceding segment is a significant predictor of variation along the front diagonal, with a velar consonant having the same effect as the post-vocalic velar nasal, seems to suggest that this is a more general phenomenon. This kind of effect is also widely attested and referred to as the ‘velar pinch’, which sees an increase in F2 and a decrease in F3 (Olive et al. 1993; Harrington & Cassidy 1999; Thomas 2010; Ladefoged & Johnson 2011).

Figure 11a plots the F2 trajectories throughout the vocalic period for three vowels along the front diagonal – /ɛ/, /i/ and /iː/ – in stressed position and before either alveolar or velar coda consonants (e.g. led, leg; wit, wick; bead, beak). In all cases it is clear to see that F2 increases towards the end of the vowel, and that this is largely
restricted to the pre-velar tokens, suggesting some coarticulatory influence from the velar place of articulation of the following consonant. Is this driving the apparent vocalic difference between the -in and -ing(g) variants? Dynamic formant analysis of these (ing) tokens, plotted in Figure 11b, suggests not. In this case, the F2 trajectory is completely flat for the pre-velar -ing(g) tokens; in other words, the difference in vowel quality between the alveolar and velar variants is consistent throughout the entire vocalic period, and there is no evidence of this difference increasing in magnitude towards the vowel-consonant transition, which one might expect if this was truly coarticulation. Furthermore, it is not the case that F3 is lower for -ing(g), which would be expected if this difference in F2 does in fact reflect the ‘velar pinch’.

However, it is possible that the fronter vowel in -ing(g) still reflects velar coarticulation, but that the effect is present throughout the entire vowel because it is unstressed and therefore shorter in duration than the stressed vowels plotted in Figure 11a. Unfortunately, a much larger dataset is required to investigate the effect of following velars on unstressed front vowels outside of the (ing) variable, which would provide the most direct evidence bearing on this debate.

5 Discussion

In this section I will address the two major components of this study’s results: in Section 5.1, what the conditioning factors of (ing) suggest about the nature of this variation and its sociosymbolic role in the North West of England; in Section 5.2, what
the phonetic realisation of (ing)’s variants tells us about its articulatory mechanisms and the implications this has for theories of the variable’s origin.

5.1 Social and internal factors on (ing)

Much like Tagliamonte (2004) finds in York, this study reveals that social/external predictors play only a small role in (ing) variation in the North of England. There are significant effects of age and sex, which pattern in the expected direction with a male preference for -in and a decrease in its use throughout the middle-aged life stage both reflecting the overtly non-standard nature of this variant and the societal pressure of linguistic norms. However, the marginal nature of these effects is reflected by the fact that even for the middle-aged women of this population sample – those for whom the prestigious velar forms should be most frequent – -in is still the majority variant used 84% of the time.

The rates of -in are so high across the board for these speakers that even grammatical category plays no role in its probabilistic occurrence, despite the frequency with which this effect is reported elsewhere in the literature (e.g. Houston 1985; Huspek 1986; Labov 2001). In the North West of England, the synchronic patterning of (ing) variation is not a reflection of historical morphology.

(ing) appears to be much more constrained by phonological factors rather than those of a morphological nature: the choice of variant is not determined at all by whether the (ing) token is polymorphemic or monomorphemic, despite the divergent behaviour of these categories in other varieties (see e.g. Tamminga 2016). In comparison, it is interesting to note that the immediate phonological environment exerts by far the strongest influence on conversational (ing) in the North West of England. This sensitivity to the segments preceding and following the vowel+nasal cluster patterns in ways predicted by previous reports (e.g. Shuy et al. 1968; Cofer 1972; Houston 1985; Tagliamonte 2004) with evidence of both assimilatory and dissimilatory effects, although the behaviour of the former differs somewhat from expectations. If (ing) is followed by a velar consonant, it is much more likely to surface with the velar nasal, reflecting a clear articulatory anticipation of the tongue dorsum gesture for the following word-initial stop. However, a similar (though weaker) effect is present before /w/ and /j/, which has not been reported previously. These segments can be considered as belonging to the same natural class, all sharing a [+DORSAL] feature (Hayes 2009: 95-97), so the same mechanisms appear to be motivating the use of [ɪŋ] before

---

14The only caveat to this conclusion that social factors play a peripheral role in North West (ing) variation is that this study only investigated the upper working class vernacular. It is possible, given the discussion of previous findings in Section 2.2.1, that the use of -in shows a marked decrease among speakers of a higher socioeconomic status.
all of these segments.

Additionally, for some speakers /l/ exhibits a comparable effect. While this may on the surface appear to lack the same mechanistic motivation, it should be noted that for many varieties spoken in the North Western county of Lancashire /l/ surfaces as a dark velarised allophone [l̴] in all environments, rather than being positionally restricted to the coda like in most other varieties of English (Kelly & Local 1986; Carter 2002; Beal 2004; Hughes et al. 2012; Cruttenden 2014; Turton 2014, 2017). Although we would need direct articulatory evidence to be sure that this is the case, it at least provides a likely explanation. In line with the results of Forrest (2017), there is a great deal of inter-speaker variation with respect to how these different contexts pattern, suggesting that while the aggregate effect is well-established across varieties of English, the ranking of these environments is somewhat less clear at the level of the individual.

It is rather more difficult to argue that the progressive dissimilation effect – by which an alveolar consonant preceding the (ing) suffix leads to higher rates of the velar variant(s) – has the same articulatory motivation. Previous literature suggests that this effect may be perceptually-grounded: as discussed in Section 2.2.3, Walker (1791: 49) prescribes not to use -in if the root also ends in the alveolar nasal (e.g. winning) because of the “disgusting repetition”. However, such an explanation would predict that the effect should only be present – or should at least be strongest – for stem-final [n] rather than all alveolar consonants, but this is not supported by the data. It is of course possible that at an earlier stage of the dialect this effect was only present for [n] before being generalised to all consonants sharing this coronal place of articulation.

(ing) also shows very strong style-shifting, with a marked decrease of -in in both the reading passage and word list elicitation tasks, which initially seems at odds with the absence of strong social stratification. If speakers are so sensitive to the non-standard status of -in with respect to stylistic constraints, should this not be reflected by strong differences between social groups? This is made even more surprising by the fact that Labov’s (1972a) marker–indicator–stereotype typology suggests an implicational hierarchy such that variables exhibiting stylistic stratification should also be differentiated along social dimensions. (ing) is undoubtedly a highly salient variable subject to overt meta-linguistic commentary, and speakers are showing awareness of the prescriptivist commentary regarding (ing) variation – or ‘g-dropping’ as it is often described by non-linguists – by using the overtly ‘correct’ variant in more conscious reading styles. However, the sociosymbolic meaning of (ing) variation is so weak in these northern communities that speakers are under no societal pressure to
reduce their use of -in regardless of gender or age group. As described by Levon & Fox (2014), in British English (ing) is associated more strongly with region than social status, with the use of -in so ingrained in the Northern vernacular.

In sum, (ing) clearly has much less sociocultural significance in the North West of England relative to other (particularly non-northern) communities. Variation still persists between -in and -ing(g), but this is predominantly influenced by mechanical factors such as speaking rate (though not frequency) and assimilation with adjacent segments, rather than by external factors driven by sociosymbolic meaning.

5.1.1 Role of [ɪŋɡ]
Having established the sociolinguistic profile – or lack thereof – of the alveolar -in variant in North West Englishes, attention now turns to the variable presence of post-nasal [ɡ]: what function or sociocultural role does [ɪŋɡ] play in this region?

The role of social factors in determining the variable use of -in was fairly peripheral, but it is entirely non-existent in the case of [ɪŋɡ]. There are further similarities in that its appearance is also most strongly predicted by factors relating to the phonological environment, although the mechanisms underpinning these effects are different. The nature of the following segment effect with respect to use of -in reflects assimilation to place of articulation, while in the case of [ɡ]-presence the favouring effect of a following vowel likely reflects phrase-level resyllabification bleeding a coda-targeting process of post-nasal /ɡ/-deletion. A similar interpretation has been applied to the same effect present in /t,d/-deletion (Guy 1980; Tamminga 2018).

As the alternation between [ɪŋ] and [ɪŋɡ] is much more clearly a case of lenition, it is not surprising to find a significant effect of speech rate, with the likelihood of [ɡ]-presence decreasing as speech rate increases15. Style is also a strong predictor, but in this case the observed stylistic stratification is less likely to reflect overt prestige and metalinguistic commentary. While the (ing) variable is generally quite salient, independent evidence suggests that the dialectal feature of post-nasal [ɡ]-presence is relatively below the radar compared to other variable phenomena (see Paper III of this

---

15Although the variation in [ɡ]-presence may be a clearer case of lenition, this analysis is still not uncontroversial. It may be the case that surface [ɡ]-presence occurs not from non-application of deletion to an underlying segment, but rather because it is inserted during the phonological derivation; such an explanation finds support from independent evidence that a pre-pausal insertion process might exist in the stressed (ng) environment, i.e. affecting words such as king, wrong etc. (see Papers I–II of this thesis). In either case, assumptions about the underlying representation of this surface three-way variation do not change dramatically: either -ingg is derived from -ing through a process of [ɡ]-insertion after word-final velar nasals, or -ing is derived from underlying -ingg through a process of coda /ɡ/-deletion in post-nasal contexts. Neither does it change the interpretation of this speech rate effect: in either case, we are dealing with a surface alternation between segment presence/absence, therefore it is not surprising that there might be a hyper-articulation effect with segment presence favoured in slower speech rates.
thesis). It is more probable that the increase in [g]-presence in word list elicitations reflects prosodic sensitivity and the way that word list items are produced with clear pauses and intonational breaks between them. This finds support from the fact that [g]-presence is also favoured before pause in conversational styles. Furthermore, the rate of [g]-presence does not show the same drastic increase in the similarly-formal reading passage, most likely due to its similarity with the conversational style with respect to prosodic factors such as speech rate and intonational phrasing.

Taken together, these results suggest that -ingg is not necessarily a local prestige form in these North Western varieties, contrary to earlier claims (e.g. Mathisen 1999; Beal 2004). It is more accurately described as a kind of ‘clear speech’ variant that is almost exclusively restricted to word lists or emphatic speech, possibly influenced by the fact that it more closely resembles the orthography and might therefore be deemed a hyper-articulated form.

In sum, -ingg seems to hold more of a functional than a sociosymbolic role: it could be argued that [g]-presence in (ing) serves as some kind of prosodic boundary marker, or that it has some discourse-pragmatic function in signalling a temporal juncture in discourse and therefore acts as a turn-taking device, but regardless of its function it is certainly not a ‘prestige’ form in the same way that the plain velar nasal -ing form is perceived to be in most varieties. Parallels can be drawn with the social meaning of -ingg: Schleef et al. (2015) argue that -ingg occupies its own indexical field in Manchester English rather than taking on the indexical role of -ing, and the production results from this study point to a similar situation in that -ingg occupies its own functional space as a hyper-articulated clear speech variant.

5.2 Phonetic correlates of variation

In the second part of this analysis the direction of causality was inverted and the surface(ing) variant was instead tested as a predictor of two phonetic properties, with the aim of investigating the extent to which the surface realisation of these variants differs in more than just nasal place of articulation. In doing so, this study finds that the alveolar -in variant differs from the velar -ing(g) variants with respect to both vowel quality and nasal duration.

-in surfaces with a significantly laxer [ɪ] vowel relative to -ing(g), with this difference registered primarily in F2. This is particularly interesting when contextualised with the theory proposed by Houston (1985, 1991) for the historical origin of (ing) variation. If the -ing(g) and -in variants in PDE truly originate from the verbal noun and present participle suffixes in OE, and if the latter of these was truly realised with a laxer vowel (-and) in northern varieties of Old English, then this might result in the
vowel quality differences that persist between (ing) variants in contemporary northern dialects.

However, it is also possible that these vocalic differences only arise due to a more general coarticulatory effect with surrounding velar consonants, i.e. the ‘velar pinch’. This would account for why the vowels of -in and -ing(g) differ mostly in F2, although the fact that F3 is not lower for -ing(g) makes this explanation somewhat less likely. It is also interesting to note that there is a tendency for /æ/ to undergo pre-velar tensing in certain varieties of American English, such that bag is realised like [bɛɪɡ] (Zeller 1997; Labov et al. 2005; Baker et al. 2008; Mielke et al. 2017), and that this is more widespread before [ŋ] than it is before velar stops. Baker et al. (2008) suggest that this is because [ŋ] is ‘more velar’ than [ɡ] due to its double articulation (with lowering of the velum in addition to the raised tongue body), and that this effect simply reflects an exaggeration of the velar coarticulation. This points to a wider phonetic motivation for pre-velar tensing, suggesting that the vocalic differences in (ing) reported in this paper might not be specific to the suffixes themselves and might not therefore result from the historical vowel quality of their OE predecessors.

To a large extent the source of this effect is not important, as in either case the mere fact that this difference exists in contemporary northern varieties has implications for the theory of (ing)’s origins. The argument first proposed by Houston (1985, 1991) is that the vowels of the present participle (originally -ind(e)) and verbal noun suffix (originally -ynge) were more similar in the South of England relative to the rest of the country – where the present participle had a much lower vowel (-end(e) or -and) – and that this drove the conflation of these two forms to occur first in southern varieties. This paper provides further evidence that front vowels – regardless of whether or not they partake in (ing) variation – have a phonetic tendency to undergo tensing before velars; if this tendency can lead to listeners misperceiving an alveolar nasal as velar when preceded by a tenser vowel, then the proposed explanation for the actuation of this change in southern England becomes increasingly convincing.

Of course, a perception experiment to directly test this would provide even stronger evidence, but ideally this would be conducted with speakers from the Middle English period when this conflation occurred. Needless to say, this is now impossible. Future work could, however, shed more light on the extent to which the laxer vowel in -in stems from the historical -and form in the North of England by comparing contemporary phonetic realisations with southern varieties, where -in supposedly derives from -ind(e) with a tenser vowel.

---

16 Wells (2009) also reports that the same effect is present for /i/, such that king is realised more like [kɪŋ] by Californians in particular.
This study also reveals a difference in nasal duration, such that -\textit{in} surfaces with a shorter nasal relative to -\textit{ing(g)}; crucially, this is independent of speech rate and does not therefore reflect the fact that the velar variants are more likely to occur in slower speaking rates. How and why does this difference – which to my knowledge has not previously been reported – arise in naturally-occurring speech? I propose that it reflects the fact that the lingual gesture for the velar nasal takes longer to articulate than the lingual gesture for the alveolar nasal, due to the former involving movement of the tongue body, which is larger in mass than the tongue tip involved in the alveolar place of articulation.

Taken together, these two phonetic correlates of (\textit{ing}) variation speak to open questions regarding its status as a lenition process. This is not often discussed in the vast literature on (\textit{ing}); Abramowicz (2007: 29) notes that (\textit{ing}) is \textit{not} a case of lenition, but does point out that the alveolar nasal may involve “less articulatory effort” than the velar nasal, and more recently Forrest (2017: 151) claims that lenition “plays a partial role in the realization of (\textit{ING})” based on observed frequency effects (discussed in Section 2.2.4). The results of this paper, specifically the observation that -\textit{in} is realised with a laxer vowel and a shorter nasal, lend strong support to the notion that variation between [\textit{in}] and [\textit{in}] can be construed as a lenition process. Consequently, \textit{(ing)} might be expected to show sensitivity to lexical frequency, but this is not well-supported in the literature: significant results are found by Forrest (2017), but not by Abramowicz (2007) or Schleef et al. (2011), and in this study any such effect of token frequency is marginal at best. This might suggest that some of these negative results stem from methodological issues, such as corpus size or choice of frequency measure.

6 Conclusion

(\textit{ing}) variation is remarkably well-studied throughout the English-speaking world, and behaves in largely the same way in all the varieties in which it has been attested. However, the behaviour of this variable in the North West of England departs from most other dialects in a number of important ways: it does not vary greatly along social dimensions, it shows no sensitivity to grammatical category (despite this being one of the strongest predictors of (\textit{ing}) in most other communities), and perhaps even more strikingly the variation in this region is between \textit{three} forms instead of two. This third variant [\textit{ing}] exists not as a replacement of [\textit{in}] to index prestige or formality, but rather occupies its own functional space either as a marker of prosodic boundaries, a turn-taking device, or as a feature of emphatic, hyper-correct speech.

It might be tempting to somehow attribute the lack of grammatical category effect
to the presence of the third variant -ingg, since in all the communities in which post-
nasal [ɡ] has been studied extensively no grammatical conditioning has been found. 
However, there is no obvious causal link between these two facts. The third variant
only arises through purely incidental means, resulting from the fact that a completely
separate phonological process of post-nasal /ɡ/-deletion did not run to completion in
the North West of England, and that the environment of this process just so happens
to encompass the unstressed (ing) suffixes (alongside stressed /ŋɡ/ clusters such as in
king and wrong etc.). Instead, it is likely that the lack of grammatical conditioning
simply stems from the fact that these speakers are so predisposed towards use of
-in, regardless of linguistic or social context, that factors constraining this variation
in most other varieties of English play only a small role here, or are in fact absent
total.

In addition to investigating the predictors of North Western (ing) variation, this
paper explored two phonetic properties of (ing) variation, and in doing so reveals a
relationship between (ing) variant and nasal duration that has not yet been reported
in the vast literature on this staple variable of sociolinguistic study. The tendency for
[n] to be shorter in duration than [ŋ], combined with the alveolar form’s laxer vowel,
provides strong evidence to argue that [m] involves less articulatory effort than [ŋ]
and that (ing) should therefore be construed as a lenition process.
# Appendix A: Word list

Tokens of (ing) are highlighted in **bold**; tokens of (ng) are highlighted in *italics*.

<table>
<thead>
<tr>
<th>heard</th>
<th>bought</th>
<th>eight</th>
<th>tool kit</th>
</tr>
</thead>
<tbody>
<tr>
<td>boat</td>
<td>had</td>
<td>batty</td>
<td>really</td>
</tr>
<tr>
<td>happy</td>
<td>balm</td>
<td>wing span</td>
<td>longest</td>
</tr>
<tr>
<td>caught</td>
<td>H</td>
<td>boy</td>
<td>shoe</td>
</tr>
<tr>
<td>heed</td>
<td>running</td>
<td>paw</td>
<td>know</td>
</tr>
<tr>
<td>pour</td>
<td>book</td>
<td>peel</td>
<td>thing</td>
</tr>
<tr>
<td>weight</td>
<td><em>slang talk</em></td>
<td>nerd</td>
<td>throw</td>
</tr>
<tr>
<td><em>ring up</em></td>
<td>spa</td>
<td>city</td>
<td>moist</td>
</tr>
<tr>
<td>hoot</td>
<td>pal</td>
<td>tear</td>
<td>base</td>
</tr>
<tr>
<td>void</td>
<td>bomb</td>
<td>turn</td>
<td><em>wrong price</em></td>
</tr>
<tr>
<td>ceiling</td>
<td>purr</td>
<td>singer</td>
<td>price</td>
</tr>
<tr>
<td>Hannah</td>
<td>hid</td>
<td>share</td>
<td>cool</td>
</tr>
<tr>
<td>Bill</td>
<td>batter</td>
<td>toast</td>
<td><em>sing songs</em></td>
</tr>
<tr>
<td>out</td>
<td>fool</td>
<td>pear</td>
<td>face</td>
</tr>
<tr>
<td>comma</td>
<td>kick it</td>
<td>fur</td>
<td>phone</td>
</tr>
<tr>
<td>sang</td>
<td>wait</td>
<td>too</td>
<td>young</td>
</tr>
<tr>
<td>Blackburn</td>
<td>how</td>
<td>dare</td>
<td>dude</td>
</tr>
<tr>
<td>code</td>
<td><em>hanger</em></td>
<td>lie</td>
<td>fair</td>
</tr>
<tr>
<td>head</td>
<td>cold</td>
<td>waiting</td>
<td><em>stronger</em></td>
</tr>
<tr>
<td>ate</td>
<td>hot</td>
<td>worse</td>
<td>nurse</td>
</tr>
<tr>
<td>beer</td>
<td>hide away</td>
<td>chair</td>
<td>anything</td>
</tr>
<tr>
<td>born</td>
<td>love</td>
<td>do</td>
<td>crowd</td>
</tr>
<tr>
<td>think</td>
<td>hate</td>
<td><em>bring it</em></td>
<td><em>something</em></td>
</tr>
<tr>
<td><em>sing it</em></td>
<td><em>bang it</em></td>
<td>annoy</td>
<td>why</td>
</tr>
<tr>
<td>bear</td>
<td>coat</td>
<td>care</td>
<td><em>thinking</em></td>
</tr>
<tr>
<td>nowt</td>
<td>pore</td>
<td>king</td>
<td><em>nothing</em></td>
</tr>
<tr>
<td>poor</td>
<td>hood</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B: Reading passage

Tokens of (ing) are highlighted in **bold**; tokens of (ng) are highlighted in *italics*.

Dave and Jane were **taking** a *long* boat ride across the sea with a *gang* of friends. They were **enjoying** the sun and *singing* *along* to music when they heard a loud scream that caught them by surprise.

“Over there! That *young* boy is **drowning**! What should we do?” Dave said, noticing a commotion in the water where the *screaming* came from. The boy was *complaining* about being *stung* by jellyfish, but things soon took a turn for the worse as a *ring* of sharks began *circling* him.

Dave was **removing** his shoes, ready to dive into the cold water, when Jane stopped him: “Wait! They’ll just kill you too!”

“You’re wrong... I can save him, and if we wait any *longer* he’s *going* to be eaten!”

They couldn’t stop **worrying** while **waiting** for help. The boy was *hanging* onto the float for dear life, but his grip was *slipping* and he didn’t look very *strong*...

“*Hang* on!” Dave shouted, **looking** for the coast guard’s number to *ring*...

He pointed to a life jacket near Jane, “*bring* it here!” he said. Luckily, Dave had taken *swimming* lessons when he was *younger* so he was a confident swimmer.

Suddenly, the shadow of a whale approached, *scaring* the sharks away.

“What an *exciting* holiday! This is much better than *going* to *boring* school!” Jane said, as the radio began to play their favourite *song*...

“We’ve had a whale of a time!”


Fruehwald, Josef. 2017. Choice of word frequency norms can dramatically affect inference. Poster presented at UKLVC11, Cardiff University, 30th August.


R Core Team. 2016. *R: A Language and Environment for Statistical Computing*.


Wickham, Hadley. 2017. tidyverse: Easily install and load the ‘Tidyverse’. Available at: <https://cran.r-project.org/package=tidyverse>.


CONCLUSION
HAVING presented the results of these four manuscripts, the conclusion of this thesis follows two major lines of discussion: firstly, in Section 6.1 I provide a summary of the results from all four papers, making explicit how they inform one another and how they speak to the over-arching questions outlined at the start of this work. Secondly, in Section 6.2 I return to the status of /ŋɡ/ in German, discussed briefly in Chapter 2, to further contextualise the results of this dissertation and to provide a reanalysis in terms of the life cycle of phonological processes. Finally, in Section 6.3 I offer some final thoughts and outline possible avenues for future work.

6.1 Summary of results

The four original papers that constitute this dissertation have all contributed to our knowledge of synchronic (ng) variation, an understudied variable in the sociophonetic literature. However, the implications of these results have far-reaching consequences beyond this specific variable.

Paper I shows how synchronic variation in post-nasal /ɡ/-deletion is strongly conditioned by morphosyntactic factors: the probability of deletion is lowest in the word-medial pre-vocalic environment (e.g. singer, hanging etc.), followed by the word-final pre-vocalic environment (e.g. sing aloud, hang on), and finally we find the most deletion in word-final pre-consonantal tokens (e.g. sing tunes, hang portraits).

The variable rates of [ɡ]-presence across these environments mirrors the diachronic progression of /ɡ/-deletion as reported by Bermúdez-Otero & Trousdale (2012), and in turn provides strong evidence in favour of a cyclic analysis. That is, variable [ɡ]-presence is best modelled in a modular architecture of grammar that separates stem-, word-, and phrase-level processes, with deletion applying variably in all three domains. The differential behaviour of /ŋɡ/ between these contexts stems from the fact that in words like singer the rule has only one chance to apply (i.e. at the stem level, before /ɡ/ moves to the onset after word-level suffixation of -er), but that
in word-final pre-vocalic contexts such as *sing aloud* the /ɡ/ does not move into the onset of the following word until the phrase level, and is thus eligible for two rounds of deletion. Pre-consonantally, e.g. *sing tunes*, the /ɡ/ is in the coda in all three cycles and is therefore subject to three rounds of deletion. A crucial empirical observation is that /ɡ/-deletion does not apply with equal probability in each of these three morphosyntactic domains: stem-level deletion applies at a lower rate than word-level deletion, which in turn applies at a lower rate than the phrase-level process. This provides further support to the theory of the life cycle of phonological processes, which predicts that sound changes become more embedded in morphosyntactic structure over time and that as a result of this, rules that apply in wider domains are much older and should apply with higher probability. In other words, there is a clear relationship between the depth of the morphosyntactic domain, the age of the process, and the rate at which it applies; this has been formulated as the Variation Collorary of the Russian Doll Theorem, which also finds empirical support from quantitative investigations of British English /l/-darkening (Turton 2016).

It should be noted that the main findings presented in this thesis are not incompatible with, and do not therefore provide direct evidence against, non-modular phonological architectures such as the hybrid exemplar models discussed in Section 4.1.1. However, such models make no direct predictions about how (ng) variation should pattern across these different morphophonological environments. In other words, the stratified modular architecture described in Chapter 4 and in Paper I requires no arbitrary stipulation of how these environments should be ordered with respect to variable [ɡ]-presence: it falls out naturally from the cyclic manner in which phonological processes apply to underlying representations. Furthermore, usage-based accounts of phonological representation emphasise the role of token frequency, particularly how high frequency words should register the effects of ongoing reductive change at a faster rate (see e.g. Bybee 1998, 2002), but there is no evidence that frequency plays any such role in (ng) variation.

The results of Paper I also reveal that while the patterning of pre-vocalic and pre-consonantal environments is largely consistent across all speakers in this population sample, the pre-pausal environment is a source of widespread inter-speaker variation. For some speakers, this environment is invariably [ŋ], and for others it is invariably [ŋɡ]. For many, this context is still subject to variation, but its influence on (ng) relative to the other environments follows no consistent pattern. It is also shown in Paper IV how, despite the overall scarcity of [ŋɡ] as a realisation of unstressed (ing) clusters (e.g. *walking* [wɔːkɪŋɡ]), it has a similar tendency to appear in phrase-final position for many speakers.
This raises interesting questions regarding why the effect of a following pause should be so variable between speakers, and also highlights a tension between the homogeneity of pre-pausal environments both within and across communities. On the one hand there are reports that the effects of pausal boundaries on other external sandhi processes are variable between dialects: for example how English /t,d/-deletion is inhibited pre-pausally in some varieties (see Guy 1980; Santa Ana 1996; Tagliamonte & Temple 2005), but applies at much higher rates in others (Fasold 1972; Bayley 1994; Hazen 2011; Baranowski & Turton forthcoming), and how /s/-debuccalisation is blocked before pause in Argentinian Spanish (Kaisse 1996) but not in other varieties (Harris 1983). On the other hand, contrary to the results of Paper I, there are reports that the effect of pause within a single community are remarkably consistent: in a large-scale investigation of (ing) variation in Raleigh, North Carolina, Forrest (2017: 151) reports how speakers were more consistent in the realisation of (ing) pre-pausally relative to pre-alveolar and pre-velar environments, pointing out how “from an articulatory standpoint, we might expect more consistency in articulation for pre-pausal environments, since no coarticulatory factors play a role”.

However, a crucial property of this inter-speaker variation in the North West of England is that the variable behaviour of (ng) in pre-pausal environments correlates with speaker age. That is, this environment has changed from one that favours /ɡ/-deletion to one that inhibits it, so much so that for younger speakers – of which a majority have a complete ban on phrase-final [ŋ] without [ɡ] – I propose that a new process has been innovated. This would take the form of a synchronic rule that actually inserts (or rather re-inserts) a post-nasal [ɡ] when in domain-final position at the phrase-level.

Discovering what this exact domain is formed the basis of the investigation in Paper II. This study took the form of a carefully-controlled elicitation task, designed to probe the collinearity between various phenomena that co-occur at the right-edge of major linguistic constituents. Boundaries can be categorised with respect to their place in the prosodic hierarchy (Selkirk 1984; Nespor & Vogel 1986; Hayes 1989), are marked by an increase in segmental duration through pre-boundary lengthening (Lehiste et al. 1976; Gussenhoven & Rietveld 1992; Wightman et al. 1992), and are often followed by a pause of varying duration, which is said to reflect production planning (Wagner 2012a). As such, on the surface this innovation appears to be conditioned by a number of factors (i.e. prosodic boundary strength, segmental duration, and presence/duration of pause), which must be teased apart in order to determine the exact mechanisms that condition this change (see Kilbourn-Ceron 2017; Tamminga 2018 for similar approaches to this problem).
The results of Paper II suggest that segmental duration – specifically the lengthening that the preceding nasal undergoes in pre-boundary position – does not influence probabilistic [ŋ]-presence; any apparent correlation on the surface simply stems from the collinearity between segmental duration and other boundary phenomena. It is relatively more difficult to tease apart the roles of pause and prosodic phrasing; in almost all cases, an intonational phrase juncture is marked by pause, and likewise a pause is almost always present as a cue to an intonational phrase boundary. However, the rare cases where this collinearity breaks down – that is, when a speaker pauses but then continues with the same intonational phrase – are all consistent in exhibiting [ŋ]-presence, suggesting that this process may not make direct reference to prosodic constituency.

In Paper II, I also posed the question of why this innovation has taken place and proposed two possible explanations for what this change might represent. On the one hand, it could be externally motivated, with intergenerational incrementation driven by overt sociolinguistic evaluation. Alternatively, it might reflect the outcome of a more general boundary-strengthening device that speakers can employ in conversation to signal turn-taking and other discourse-pragmatic functions, adding to existing phonetic cues to boundary presence such as segmental lengthening (Turk 2012), intonational boundary tones (Pierrehumbert & Hirschberg 1990), and non-modal phonation (Cutler & Pearson 1985).

The possibility of this being evaluation-driven change motivated the perception study in Paper III, which provides the first direct evidence of how post-nasal [ŋ] is evaluated in stressed (ng) tokens using a matched-guise paradigm. The results reveal two interesting diachronic patterns regarding the extent to which [ŋɡ] is: (a) characterised as a feature of northern dialects, and (b) subject to overt sociolinguistic evaluation. The salience and indexical strength of post-nasal [ŋ]-presence is increasing over time: older members of North Western speech communities do not perceive [ŋɡ] as more northern than [ŋ], and as a result do not evaluate them any differently. On the other hand, younger speakers are more aware of the dialectal nature of [ŋɡ], exhibiting more agreement on its northern status. However, this increase in awareness does not translate to uniform evaluation: there is inter-speaker variation with respect to norm alignment, with some adopting the local [ŋɡ] variant as the prestige form and others aligning with the regionally-unmarked standard [ŋ].

This raises interesting theoretical questions regarding the concept of the speech community and specifically how it should be defined, suggesting that Labov’s (2001) Principle of Uniform Evaluation – the notion that members of a speech community are united in their adherence to a set of shared sociolinguistic norms – may not
necessarily apply in the early stages of social meaning. Furthermore, the lack of a strong, community-level evaluation of (ng), coupled with the fact that these diachronic changes in its social meaning are independent of morphophonological environment, suggest that it cannot be invoked as an explanation of the pre-pausal change in production. Alternatively, it should be noted that the heterogenous evaluations presented in Paper III may to some extent reflect the fact that the notion of community can be conceptualised in different ways, and the complex hierarchical manner in which they are ideologically constructed, as described in Chapter 3. Further investigation is needed to explore the social meaning of this form across regions of the North West in a more fine-grained manner, and it would be particularly illuminating to complement these results with a more qualitative, ethnographic approach to this topic.

In relation to the second proposal regarding the motivations behind this pre-pausal innovation, namely that it might reflect a more general process of phrase-final strengthening or boundary marking, we can turn to a number of independent observations that might support this claim.

There is suggestive evidence that ejectivisation of /k/ is increasing in apparent time in Glasgow English (discussed in McCarthy & Stuart-Smith 2013, who draw comparisons with earlier data reported in Stuart-Smith 1999); not only is this more frequent in turn-final position, its probability is also increased when the /k/ is preceded by a nasal (McCarthy & Stuart-Smith 2013: 292). Given that ejectivised stops are more phonetically salient than their pulmonic counterparts – due to the increased supralaryngeal air pressure leading to a higher amplitude burst (see Ladefoged & Maddieson 1996; Gallagher 2010) – and that this phenomenon is favoured in the same segmental and prosodic environment, it is tempting to analyse this parallel strengthening of (ng) and (nk) as part of a wider phrase-final velar fortition process. While this proposal calls for a more detailed investigation of ejectivisation patterns in northern Englishes, we can actually find further empirical support by turning the lens of inquiry away from English.

### 6.2 Cross-linguistic evidence

To what extent does this synchronic behaviour of post-nasal /g/-deletion, if truly reflecting its diachronic pathway of change, lend support to the theory of the life cycle? Furthermore, if this pre-pausal innovation is not motivated by external factors relating to attitude, evaluation and indexicality, then what is driving this change? Why are these younger speakers so marked in their near-categorical use of [ŋɡ] before
pause? As summarised in the previous section, the results of this dissertation provide some answers to these questions, but they can be further illuminated by turning our attention to the cross-linguistic behaviour of /ŋɡ/, namely its status in German.

### 6.2.1 Rule generalisation in the life cycle of /ɡ/-deletion

In Chapter 2 I provided a brief overview of the status of /ŋɡ/ in German. Recall that the distribution of [ŋ]~[ŋɡ] is such that /ɡ/-deletion applies across the board in word-final position, but in word-medial position its application is dependent on the following vowel. If /ŋɡ/ is followed by schwa, coalescence takes place, but if it is followed by a full, unreduced vowel, the [ɡ] is present on the surface. Crucially, this is the case regardless of morphological structure and whether or not the following vowel is stressed: all that matters is whether or not it is schwa. Contrast this with English, where word-medial deletion of /ɡ/ is sensitive to morphological composition: it deletes intervocally if the following vowel belongs to a stem-attaching suffix (e.g. sing-er [sɪŋə]), but otherwise the /ɡ/ is present on the surface (e.g. finger [fɪŋɡə]). The explanation proposed by Hall (1989) and Wiese (2000) is that in German the schwa occupies an empty ⊗ slot on the skeletal tier and does not become syllabified into prosodic structure until after post-nasal /ɡ/-deletion has applied; consequently, /ɡ/ remains in the coda at the relevant stage of the derivation, tautosyllabic with [ŋ], and is thus eligible for deletion in words like Finger [fɪŋə].

However, there is an alternative analysis that falls out naturally from the life cycle of phonological processes. A crucial diachronic prediction of the life cycle is that a process can not only undergo domain narrowing to climb from lower to higher morphosyntactically-defined strata (i.e. from the phrase level to the word and stem levels), but can also undergo rule generalisation. This involves an expansion of the environments targeted by a process, and is most often defined with respect to prosodic domain. In other words, the specification of a phonological process, at least in cases of lenition, involves two properties: what morphosyntactic domain it applies in (the stratal choice) and what segmental or prosodic environment it targets (the prosodic choice) (Bermúdez-Otero 2013: §3.1).

Take English /l/-darkening as an example, the life cycle of which has been discussed in detail by Turton (2014, 2017). This is a syllable-based process in most varieties of British English – that is, it darkens /l/ in the coda – but has advanced to a foot-based process in American English. This is evidenced by the fact that speakers of such varieties darken /l/s that do not appear in a syllable coda in any of the three morphosyntactic domains, for example in monomorphemic words such as yellow or helix (see Olive et al. 1993: 366; Hayes 2000: 95–96). Under a cyclic architecture, these
tokens of /l/ can only be darkened by a process that targets weak (i.e. non-initial) position in the prosodic foot, rather than weak position in the syllable (the coda). These inter-dialectal differences are not random deviations from the same starting point: they reflect different ordered stages along a diachronic trajectory laid out by the life cycle.

A crucial point is that rule generalisation does not necessarily involve an expansion along a scale of prosodic constituents, but can also involve a phonetically-grounded markedness hierarchy (Bermúdez-Otero 2013). This is illustrated by a number of changes, such as word-final velarisation in Caraqueño Spanish (see Ramsammy 2015), and the Old High German consonant shift (see Davis 2008: 212, discussed in Bermúdez-Otero 2013); in both cases rule generalisation involves different environments ranked in order of phonetic favourability, with more advanced stages corresponding to less phonetically-favourable environments.

We can apply the same theoretical approach to /ŋɡ/-coalescence in German. Throughout the existing literature and the discussion in Chapter 2 it has been assumed that the process of post-nasal /ɡ/-deletion targets this segment in coda position. However, it is entirely possible that rule generalisation in German has led to an expansion of the target environment, such that deletion now applies if the /ɡ/ is in the onset of a schwa syllable in addition to instances of coda /ŋɡ/. If we think of these environments in terms of phonetic cues to the presence and place of articulation of stops, which rely heavily on adequate formant transitions with surrounding vowels\(^1\), then this is the next natural step from the kind of coda-targeting rule we find in English. Although being pre-vocalic in a syllable onset provides better cues than being in the coda, a following schwa is undoubtedly the weakest of all vowels given its reduced and unspecified nature (Recasens 1991; Browman & Goldstein 1992; van Bergem 1994; Bates 1995).

Crucially, there is independent evidence to suggest that non-coronal consonants, being a marked place of articulation, rely more heavily on these cues than unmarked coronal stops (Burzio 2007)\(^2\). In sum, the ordering of these environments is not arbitrary but rooted in perception with a clear phonetic grounding. The coda environment provides the weakest cues, followed by the onset of a schwa syllable. More robust

---

\(^1\)There is a wealth of evidence that formant transitions play a crucial role as a perceptual cue to the place of articulation of stops: e.g. Dorman et al. (1977); Stevens & Blumstein (1978); Kewley-Port (1983); Nearey & Shammas (1987); Fowler (1994); Sussman & Shore (1996); Fruchter & Sussman (1997); Alwan et al. (2011).

\(^2\)See also the ‘Ross stress’ rule in English, named after Ross (1972) who observed that final syllables checked by a non-coronal obstruent tend to bear secondary stress, e.g. \textit{kidnap} [k\textipa{d}n\textipa{p}], \textit{shindig} [\textipa{ʃ}\textipa{n}d\textipa{ɪ}d\textipa{ɪ}]. The fact that final syllables checked by coronal obstruents show no such behaviour (cf. \textit{carpet} [\textipa{ka}\textipa{p}t\textipa{ɪ}], not *[\textipa{ka}\textipa{ˈp\textipa{ɪ}t}\textipa{ɪ}]) is said to reflect the fact that coronals, being an unmarked place of articulation, rely less on place cues and therefore permit reduction of the preceding vowel.
phonetic cues are provided in the onset of a full vowel syllable, particularly if this bears stress and is therefore foot-initial with domain-initial strengthening, and this is reflected by the fact that [ɡ] is invariably present in such contexts. Based on this, we can construct the following markedness hierarchy, which likely reflects different diachronic stages of /ɡ/-deletion’s target environment:

- **Stage 1**: \( \sigma \)
  - in the syllable coda, e.g. *Ding* [ˈdɪŋɡ] ‘thing’

- **Stage 2**: \( \sigma \)
  - in the onset of a schwa syllable, e.g. *Finger* [ˈfɪŋɡə] ‘finger’

- **Stage 3**: \( V_f \)
  - in the onset of a full vowel syllable, e.g. *Tango* [ˈtaŋɡo] ‘tango’

- **Stage 4**: \( V_sp \)
  - in the onset of a stressed syllable, e.g. *tangieren* [tɑŋˈɡiːrən] ‘to affect’

In English, post-nasal /ɡ/ is forbidden in the coda but permitted in a syllable onset – even if that onset leads into a schwa nucleus – and is therefore still stuck at the first stage. In German, however, the process has undergone rule generalisation and now targets heterosyllabic [ŋɡ] clusters if that [ɡ] would otherwise surface before [ə] (i.e. stage 2). The next logical step in the life cycle of post-nasal /ɡ/-deletion would involve rule generalisation to all unstressed syllable onsets (stage 3), and possibly after that even stressed syllable onsets, but these are thus far unattested.

This reanalysis of German /ŋɡ/ serves to illustrate the predictive power of the life cycle of phonological processes: this model of language change allows us to account for these different ‘systems’ of ng-coalescence as dialectal microtypologies at different stages along the same diachronic pathway of change. The distribution of [ŋ]~[ŋɡ] in contemporary dialects of German is simply registering the effects of a more innovative avatar of post-nasal /ɡ/-deletion relative to the process that applies in English.

### 6.2.2 [ŋk] in Northern German

Investigating /ŋɡ/ in German not only provides suggestive evidence for rule generalisation in the life cycle of post-nasal /ɡ/-deletion, it also provides an interesting parallel to the phrase-final innovation that this dissertation has attested in northern Englishes.
Table 6.1: Example derivations illustrating how different rule ordering can lead to word-final [ŋ] in Standard German and [ŋk] in northern German varieties. Based on Table 9.5 from Lass (1984: 207) (NPA = nasal place assimilation).

<table>
<thead>
<tr>
<th>[ŋk]-dialect</th>
<th>Standard German</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ding</strong></td>
<td><strong>Dinge</strong></td>
</tr>
<tr>
<td>Underlying</td>
<td>/dɪŋɡ/</td>
</tr>
<tr>
<td>NPA</td>
<td>/dɪŋɡ/</td>
</tr>
<tr>
<td>Final devoicing</td>
<td>/dɪŋk/</td>
</tr>
<tr>
<td>/ɡ/-deletion</td>
<td>/dɪŋɡ/</td>
</tr>
</tbody>
</table>

Although, as has been noted in various parts of this dissertation, word-final /ŋɡ/ is reduced to [ŋ] in German, there are northern German varieties in which this cluster surfaces as [ŋk], e.g. *Ding* [dɪŋk] ‘thing.sg’ (cf. *Dinge* [dɪŋɡə] ‘thing.pl’). That is, in some varieties word-final /ŋɡ/ surfaces with the velar stop actually present on the surface, though in voiceless form. This is typically taken to be the underlying post-nasal /ɡ/, simply subject to the word-final obstruent devoicing process that is so widespread and well-reported in German (Rubach 1990; Brockhaus 1995; Wiese 2000; Piroth & Janker 2004; Kleber et al. 2010; Roettger et al. 2014). Referring to this regional phenomenon as ‘intrusive [k]’, Lass (1984: §9) points out that it simply requires extrinsic rule ordering such that final devoicing occurs before post-nasal /ɡ/-deletion in northern varieties of German (i.e. a bleeding order) but after deletion in standard varieties (i.e. a counter-bleeding order). Example derivations are given in Table 6.1.

This analysis accounts not only for the presence of [ŋk] word-finally but also its absence word-medially where final devoicing does not take place and can not therefore prevent the deletion process from applying. Lass makes no note of the extent to which this word-final [ŋk]-presence is variable, but a later quantitative study by Féry et al. (2009) reveals that, for speakers in the Berlin-Brandenbourg region at least, [ŋk] is present just 30% of the time averaged across all subjects. There is significant inter-speaker variation with respect to the relative frequencies of [ŋk]-[ŋ], but all speakers had a preference for [ŋ]; that is, for no single speaker was [ŋk] more frequent than [ŋ].

Féry et al. (2009) also observe a number of contextual factors that influence the presence of this post-nasal stop, with the effect of prosodic boundaries being the most pertinent to this dissertation. The rate of stop presence increases from 19% when final in the phonological word, to 46% when final in the intonational phrase. In other words, the variation between word-final [ŋ]-[ŋk] in northern German mirrors the variation between [ŋ]-[ŋɡ] in northern English in that both non-coalesced variants are favoured phrase-finally. As the study does not take into account any social factors,
it is unclear whether or not this phrase-final effect is increasing in magnitude over time like [ŋɡ] is in North West England. However, these independent observations in two different languages do suggest that there might be something universal in favouring non-coalescence – what might be construed as a ‘clear speech’ variant – in phrase-final position. Féry et al. (2009: 192–193) speculate that it might be used as an “additional cue to finality of a prosodic domain”, adding it to a list of phenomena that cue prosodic boundaries, such as durational lengthening and boundary tones. As discussed in Paper II, I propose a similar functional explanation to the pre-pausal behaviour of (ng) in northern Englishes.

6.3 Final thoughts and future work

At the beginning of this dissertation I introduced Northern English [ŋɡ] as a prototypical example of this region’s so-called ‘conservatism’ with respect to language change. While this is still undoubtedly the case – after all, we are dealing with the retention of a historical form that has been lost in almost every other variety of English – the results of this dissertation suggest that [ŋɡ] is far more than a shibboleth of dialects spoken in the North West. In contemporary dialects, this sociophonetic variable exhibits patterns of synchronic variation that offer a valuable insight into centuries of linguistic change. This in turn contributes to ongoing debates concerning the course of sound change and how linguistic knowledge is represented and computed in speakers’ grammars. Furthermore, this dissertation has revealed that (ng), regardless of its ‘traditional’ status, is still subject to ongoing change not only in production but also perception. This work has captured it at a crucial stage of incipient social meaning with no community-wide consensus on its indexical properties, and simultaneously towards the tail-end of a change in production that now situates this non-coalesced variant as a pre-pausal boundary marker for younger speakers.

Throughout this dissertation the goal has been to provide answers to a range of questions regarding the synchrony and diachrony of (ng) in Northern English. In answering these questions it has also raised a number of issues that warrant further study. This work has treated the North West of England as a homogenous community with respect to this variable, and while there is no evidence to suggest that this is not the case, it calls for a more detailed investigation into the extent to which (ng) shows consistency in its social and internal constraints across varieties spoken in this linguistically-diverse region.

This dissertation has also uncovered a hitherto-unreported change in pre-pausal [ɡ]-presence. Given that northern dialects of German exhibit a similar resistance
to phrase-final nasal+stop coalescence, this raises questions regarding the extent to which this reflects a universal phonetic tendency. It may not even be necessary to turn to the cross-linguistic status of /ŋɡ/ to illuminate this question: outside the North West and West Midlands of England, data from the 1950s Survey of English Dialects attests [ŋɡ] in a small pocket around Kent in the south of England. Does (ŋɡ) variation exhibit the same social and linguistic profile in these two disparate regions?

Furthermore, does this change reflect a more general trend of younger speakers avoiding lenition and instead favouring the use of ‘clear speech’ variants in phrase-final position? This calls for a more nuanced approach to intra-speaker variation, which pays due regard to the patterns of covariation between various probabilistic lenition processes within single speakers.

Whilst these questions must remain unanswered for now, the landscape of synchronic (ŋɡ) variation has begun to take shape. The life cycle of phonological processes – specifically its synthesis of diachronic and synchronic explanation – provides a complete account of (ŋɡ): its historical evolution, its status in contemporary dialects, and the possible pathways of its future change.
Bibliography


Fruehwald, Josef. 2014. Forced alignment and vowel extraction for sociophonetics. Paper presented to the Glasgow University Laboratory of Phonetics (GULP), University of Glasgow, 13th March 2014.

Fruehwald, Josef. 2017a. Choice of word frequency norms can dramatically affect inference. Poster presented at UKLVC11, Cardiff University, 30th August.


Fruehwald, Josef. 2018. syllabifyr: Syllabifier for CMU dictionary transcriptions. R package version 0.1.0.


265


Nguyen, Noël, Sophie Wauquier & Betty Tuller. 2009. The dynamical approach to speech perception: From fine phonetic detail to abstract phonological categories.


271


Strycharczuk, Patrycja & James M. Scobbie. 2017. Whence the fuzziness? Morpho-
logical effects in interacting sound changes in Southern British English. Laboratory
Phonology: Journal of the Association for Laboratory Phonology 8(1), 1–21.

Docherty (eds.), Urban voices: Accent studies in the British Isles, 203–222. London:
Arnold.

Sundara, Megha, Katherine Demuth & Patricia K. Kuhl. 2011. Sentence-position ef-
effects on children’s perception and production of English third person singular –s.

Sussman, Harvey M. & Jadine Shore. 1996. Locus equations as phonetic descriptors

Swerts, Marc. 1997. Prosodic features at discourse boundaries of different strength.

In Britt-Louise Gunnarson, Lena Bergström, Gerd Eklund, Staffan Fridell, Lise H.
Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren & Mats Thelander
(eds.), Language Variation in Europe: Papers from the Second International Conference

University Press.

Tagliamonte, Sali. 2012. Variationist sociolinguistics: Change, observation, interpreta-

Tagliamonte, Sali & Rosalind Temple. 2005. New perspectives on an ol’ variable: (t,d)

Tamminga, Meredith. 2016. Persistence in phonological and morphological variation.
Language Variation and Change 28(3), 335–356.

Tamminga, Meredith. 2017. Matched guise effects can be robust to speech style. The

Tamminga, Meredith. 2018. Modulation of the following segment effect on English
coronal stop deletion by syntactic boundaries. Glossa: A Journal of General Linguis-
tics 3(1), 1–27.


Wickham, Hadley. 2017. tidyverse: Easily install and load the ‘Tidyverse’. Available at: <https://cran.r-project.org/package=tidyverse>.


