DISTINGUISHING DIFFERENT LEVELS OF REPRESENTATION IN THE ACOUSTICS:
A CASE STUDY IN SCOTTISH ENGLISH EPENTHESIS

By

Kaylin Marie Smith

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Linguistics—Doctor of Philosophy

2020

ABSTRACT

DISTINGUISHING DIFFERENT LEVELS OF REPRESENTATION IN THE ACOUSTICS:
A CASE STUDY IN SCOTTISH ENGLISH EPENTHESIS

By

Kaylin Marie Smith

Much of the typology for *inserted vowels*—i.e. those that are not present etymologically, but which surface in the acoustics—has come from cross-linguistic patterns identified perceptually by linguists. Crucially, though, inserted vowels may superficially sound the same, despite emerging at different stages of derivation. Inserted vowels may be present in the underlying representation as vowels that have become fossilized over time (i.e. lexicalized vowels), in the surface representation by way of a derived process (i.e. epenthetic vowels), or in the phonetic implementation as a result of gestural alignment (i.e. excrescent vowels). In a modular feedforward framework, in which phonology feeds phonetics, phonological processes can display gradience typically associated with phonetic processes. Determining what level of representation an inserted vowel originates from, then, can be challenging and may result in misclassification.

Misclassifying an inserted vowel can have consequences for phonological theory, since the theory is modeled to generate the typological patterns found in natural language production. A vowel that has been misclassified may appear to be a typological exception, and this exceptional behavior can result in proposals for new categories of inserted vowels which do not fit the phonology-phonetics dichotomy, or revisions to the modular feedforward framework to incorporate some degree of overlap between the phonetics and phonology modules. It is important, then, that the methodology used to classify inserted

vowels as originating in the surface representation, for example, do so by excluding the possibility that the vowel did not originate in the underlying representation or from gestural alignment.

In this dissertation, I utilize an under-documented case of vowel insertion in Scottish English as a case study to distinguish between lexicalized, epenthetic, and excrescent vowels in the acoustic signal. Using data collected in two language production experiments, I assess the inserted vowel's phonological and phonetic properties against two sets of diagnostic criteria—one which I use to establish that the vowel is phonological, and another which I use to establish that the vowel is epenthetic. These diagnostic criteria are applied via a process of elimination, in which I exclude the possibility that the vowel is excrescent to establish that it is phonological using the first set of diagnostic criteria, and subsequently exclude the possibility that the phonological vowel is lexicalized to establish that it is epenthetic using the second set of diagnostic criteria. The language-specific findings, novel diagnostics and exclusion process, and patterns for epenthetic vowels presented in this dissertation serve to supplement Scottish English phonology, improve the methodology available to phonologists investigating the origins of inserted vowels, and contribute to the typology of inserted vowels in a modular feedforward framework.

ACKNOWLEDGMENTS

The first of my sincerest thanks goes to my guidance committee—my co-chairs Karthik Durvasula and Alan Beretta, and committee members Yen-Hwei Lin and Jason Moser.

Karthik, you agreed to be my co-chair later in my degree, when I decided that I wanted to switch my focus from neurolinguistics to phonetics and phonology, and I am so thankful you did. Often times when I felt incredibly challenged and exhausted by this dissertation, you would say “I’m just pushing you” to think about X or reconsider Y. I sometimes felt helpless when you said this, but now I have the hindsight to fully understand why you did that. When I submitted my first draft of the dissertation to you, I was mainly just happy that it was on paper because I was too exhausted to worry about whether it was in the best shape it could be. You helped me move past that short-term complacency and pushed me to want my dissertation to serve as the culmination of this journey—for it to represent just how far I’ve come. You believed in me, and for that reason, when I submitted the final version of this paper to my committee, I didn’t feel relief from being done with it, I just felt incredibly proud of it. I feel so lucky to have been the recipient of your thoughtful and considerate mentoring, and to call you my colleague and friend. Thanks to you, I finally consider myself a phonetician and phonologist.

I would not have ended up at MSU if it weren’t for Alan. In the stressful process of applying for graduate school, you made me feel so welcome to join your lab. While I was at the University of Edinburgh, you checked in many times to see how my M.Sc. program was going and to tell me that you and the other lab members were looking forward to my arrival the following year. Once I arrived at MSU, you taught me everything I needed to know about EEG and psycho- and neuro- linguistics and provided me with so many incredible opportunities to conduct research. The Friday lab beer tradition that you created was such a terrific bonding experience for students and faculty, and I'm hopeful that I can attend lab beers again as an alumni. You were an amazing mentor, advisor, and friend during my five years at MSU, and I will never be able to thank you enough for all that you've done for me. After five years under your wing, I can finally say that my time at MSU is "done and dusted" and that my dissertation is "sorted"!

On my first day at MSU, Yen-Hwei asked me and the other students in her phonology class to write our names in IPA on the chalkboard, and I didn't know how. You offered me assistance outside of class and never made me feel inadequate for not knowing the same amount of material as the other students. I wrote a term paper for your class on vowel insertion in Scottish English and you encouraged me to pursue this research for my second comprehensive paper, which ultimately led me to this dissertation. Without your encouragement, I would not have discovered my love for phonology. Thank you.

Thank you to Jason, who I selected as a committee member because of my appreciation for his research on emotion and his expertise using EEG. Jason, you've provided a fresh perspective on my research that I could not have gotten elsewhere. You've always said to me that you feel like you've learned something new through my various research projects, which really helped build my confidence in my own work. I am so thankful that you helped me along this journey and I am honored to have been your advisee.

I have to extend my heartfelt thanks to Richard Shillcock, who supervised my Masters dissertation at the University of Edinburgh, along with the experiments that constituted the basis of my second comprehensive paper and Ph.D. dissertation. I will never

forget my visits to the Informatics building to grab a cuppa and just have a chat with you. You have always supported my research ideas and this has been so important for my development as an empiricist. And after years of supporting me through my academic journey, you showed your continued support by virtually attending my Ph.D. defense, which meant so much to me.

Thank you, thank you, thank you to my best friends at MSU. To Patrick Kelley, for being my mentor and confidant from the very beginning. To Monica Nesbitt, for being the ultimate role model and friend who pushed me to see in myself what I so effortlessly see in you—a complete and total boss. To Danny Feldscher, for keeping me grounded with your ability to see things from all angles, to never assume the worst, and to care ever so delicately for personal growth. To Scott Nelson, for making me laugh every day, inspiring me to be a better linguist, listening to Built to Spill's *Untethered Moon* with me about 50 times while carpooling, and really believing in me. To Mohammed Ruthan, for keeping me sane in my last year and always knowing exactly what to say when I was worried or upset. To Josh (my cohort!), for sharing this entire experience with me and making me laugh through the lows. I love you guys!

And, of course, I would not have gotten through this program without Saya Uehara, Yan Cong, Alicia Parrish, Alex Mason, Drew Trotter, Joe Jalbert, Chad Hall, Amaresh Joshi, Ye Ma, and Ho-Hsin Huang. You have all been role models to me at different stages of this process and I cannot thank you enough for your friendship, collaboration, and mentorship. I'd like to thank Suzanne Wagner for being such an ally for graduate students and an absolute powerhouse female in academia for us all to look up to. Thanks to Isaac, Nikki, Jillian, and Bethany for annotating massive sound files in Praat. Thanks to the

audience members at GLEAMS, ICPhS, and the Phonetics and Phonology Group, where this research was presented.

Thanks to my extended family in Scotland, the Prentices, for sharing my recruitment advertisements and participating in my pilot study. Thank you to Warren Maguire at the University of Edinburgh for thought-provoking discussions about epenthesis in Scottish English and Irish English. Thank you to my wonderful friends in Scotland, Emma Waterston and Becca Wagstaff, for lovely day trips and hangs at the gin bar both times I was in Edinburgh collecting data. Thank you to my best friends Kelsi Ryan and Sarah Oechsle for caring for me all of these years and being such a huge part of my life.

Making it through this degree would not have been possible without my family. Mom, you've supported me through every single stage of this process by listening to the mundane day-to-day experiences, offering to help me annotate files when you got home from work (even though I would never subject you to that), wanting to drive to Michigan to sit next to me while I give my dissertation defense online during a pandemic, and always knowing how to make me feel better. Thank you for everything you have given me. Dad, you've given me some of the best advice I've ever gotten about how to achieve my goals. The thing I've always appreciated about you is how focused you are on personal growth—whether it be in the form of self-affirmations, taking care of your health through what you eat, and always visualizing who you want to be, what you want in life, and how to get there. You've shared with me so much wisdom in this regard and I am truly thankful for this gift. Nikki, you've sent me inspirational quotes and pictures of cute animals when I was struggling through this process, and you've engaged in meaningful discussions about what

I do by always asking interesting questions about my research. When I was accepted to give a talk at my first phonetics conference in Australia, you came with, listened to me practice my talk at O'Hare, made sure you were let in to watch my talk, and felt so nervous for me that you felt faint. Briana, you've kept me laughing through it all, and you've never hesitated to let me know how proud you are of me. You've counseled me through so many days where I felt not good enough and pushed me harder than anyone to stop apologizing and feeling guilty for focusing on my own needs. I love you all so much and am so lucky to call you my family.

Last but certainly not least, I want to thank my best friend and love of my life, Lucas Elenitsky. We met 16 years ago, and this fall we'll celebrate our 10-year anniversary. We spent 2.5 years long distance and another 2.5 years on different continents, but in Michigan, we bought a home together, went to graduate school together, and got our adorable cat here. We made a wonderful life together in the midst of the stress that comes with being a graduate student. Lucas, you have supported me through this chapter of my life by constantly encouraging me to see in myself what I needed to see in order to get through each phase of this process, and I am eternally grateful for your patience, encouragement, and love. My life is infinitely better than it ever could because you are by my side.

Each of you have pushed me to see that I am capable of so much more than what I let myself believe, and I truly cannot thank you enough.

P.S. In case you were wondering, I *am* crying right now. Typical!

TABLE OF CONTENTS

LIST OF TABLES xiii

LIST OF FIGURES xv

Chapter 1: Introduction 1
1.1 The current dissertation 3
1.2 Theoretical phonological framework 5
1.2.1 Sounds as bundles of features 6
1.2.2 Sounds as gestures 7
1.2.3 Unifying features and gestures 10
1.2.4 Types of inserted vowels in a featural-gestural framework 11
1.3 Three types of inserted vowels 12
1.3.1 Lexicalized vowels: Discrete and phonological 12
1.3.2 Vowel epenthesis: Discrete and phonological 13
1.3.3 Vowel excrescence: Gradient and phonetic 15
1.4 Distinguishing between types of inserted vowels 16
1.4.1 Existing criteria for diagnosing excrescent and epenthetic vowels 16
1.5 The current studies 19
1.6 Organization of this dissertation 20

Chapter 2: Language Background 22
2.1 Scottish English: A linguistic continuum 23
2.1.1 Broad Scots 23
2.1.2 Scottish Standard English 25
2.2 Scottish English Phonology 25
2.2.1 Syllable structure 25
2.2.2 Consonant inventory 26
2.2.3 Scottish English: Vowels 29
2.3 Concluding remarks 32

Chapter 3: Phonological, not excrescent. Epenthetic, not lexicalized. 33
3.1 The current experiment 34
3.2 Diagnostic criteria 36
3.2.1 Criteria for diagnosing the vowel as phonological 36
3.2.2 Criteria for diagnosing the vowel as epenthetic 42
3.3 Organization of the chapter 45
3.4 Methods 45
3.4.1 Participants 45
3.4.2 Stimuli 46
3.4.3 Procedure 48
3.4.4 Acoustic measurements 50
3.4.5 Data analysis 54

3.5 Results ... 54
3.5.1 Diagnostic criteria for the vowel being phonological (not excrescent) ... 55
3.5.2 Diagnostic criteria for the vowel being epenthetic (not lexicalized) ... 60
3.5.3 Exploratory result: Inserted vowel duration and preceding vowel duration ... 64
3.6 Summary of findings and conclusions 3.6.1 The vowel is phonological (not excrescent) ... 66
3.6.2 The vowel is epenthetic (not lexicalized) ... 67
3.6.3 Conclusions ... 68

Chapter 4: The vowel interacts with speech rate. ... 70
4.1 The current experiment ... 71
4.2 Hall's (2006) criterion for excrescent vowels at different speech rates ... 73
4.3 Theoretical modeling and empirical patterns for gestural overlap ... 75
4.3.1 Organization of gestures in a gestural coupling framework ... 75
4.3.2 Gestural timing of consonants across morpheme and word boundaries ... 77
4.3.3 Gestural timing of anti-phase gestures ... 78
4.4 Developing a speech rate diagnostic for phonological and excrescent vowels ... 81
4.5 Methods ... 83
4.5.1 Participants ... 83
4.5.2 Stimuli ... 84
4.5.3 Procedure ... 85
4.5.4 Acoustic measurements ... 87
4.5.5 Data analysis ... 90
4.6 Results ... 91
4.6.1 Vowel duration and speech rate ... 92
4.6.2 Additional diagnostics for the vowel being phonological, not excrescent ... 93
4.7 Summary of findings and conclusions ... 99
4.7.1 The vowel is phonological (not excrescent) ... 99
4.7.2 The vowel is epenthetic (not lexicalized) ... 100
4.7.3 Conclusions ... 101

Chapter 5: The vowel is morpho-phonologically conditioned. ... 103
5.1 The current chapter ... 103
5.2 Organization of the chapter ... 105
5.3 Diagnostic criteria for the vowel being phonological (not excrescent) ... 105
5.3.1 Systematic morpho-phonological distribution ... 105
5.3.2 Interaction with other phonological processes ... 118
5.4 Diagnostic criteria for the vowel being epenthetic (not lexicalized) ... 123
5.4.1 Absence of vowel insertion in words with C_1 deletion, derhoticization, /l/-vocalization, or syllabic C_2s ... 123
5.5 Summary of findings and conclusions ... 125
5.5.1 The vowel is phonological (not excrescent) ... 125
5.5.2 The vowel is epenthetic (not lexicalized) ... 126
5.5.3 Conclusions ... 126

Chapter 6: Discussion and Conclusions ... 128

6.1 Summary of findings 128
6.2 Implications 131
6.2.1 Implications for Scottish English phonology 131
6.2.2 Implications for methodology 132
6.2.3 Implications for the theory 136
6.3 Conclusion 139

APPENDICES 140
APPENDIX A: Recruitment Flyer for Experiment 1 141
APPENDIX B: Experiment 1 Stimuli 142
APPENDIX C: Demographic Questionnaire for Experiment 1 152
APPENDIX D: Speaker Demographics for Experiment 1 153
APPENDIX E: Recruitment Flyer for Experiment 2 154
APPENDIX F: Experiment 2 Stimuli 155
APPENDIX G: Demographic Questionnaire for Experiment 2 156
APPENDIX H: Speaker Demographics for Experiment 2 161

REFERENCES 162

LIST OF TABLES

Table 1. Scottish English consonant inventory. Recreated from Scobbie et al. (2006, p. 5). 26

Table 2. Expected durational correlations for coupled gestures in in-phase and anti-phase relationships. 38

Table 3. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens. 40

Table 4. Diagnostic predictions for phonological and excrescent vowels. 42

Table 5. Diagnostic predictions for epenthetic and lexicalized vowels. 44

Table 6. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens. 56

Table 7. Linear Mixed Effects model for Consonant-to-Consonant Duration as predicted by Insertion, with Subject and Word as random intercepts. 60

Table 8. Linear Mixed Effects Model for Vowel Duration predicted by Preceding Vowel Duration with Subject and Word as random intercepts. 65

Table 9. Linear Mixed Effects Model for Vowel Duration predicted by Preceding Vowel Duration with Subject and Word as random intercepts. 66

Table 10. Diagnosis of the vowel as phonological, not excrescent. 67

Table 11. Diagnosis of the vowel as epenthetic, not lexicalized. 68

Table 12. Linear Mixed Effects model for Inserted Vowel Duration as predicted by Speech Rate, with Subject and Word as random intercepts. 93

Table 13. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens in the normal speech rate. 95

Table 14. Linear Mixed Effects model for Consonant-to-Consonant Duration as predicted by Insertion, with Subject and Word as random intercepts. 99

Table 15. Diagnosis of the vowel as phonological, not excrescent. 99

Table 16. Diagnosis of the vowel as phonological, not excrescent. 100

Table 17. Summary of experimental stimuli by Word Type and Position of Consonant Sequence. 108

Table 18. Mixed Effects Logistic Regression for Insertion as predicted by Position of Consonant Sequence using Subject and Word as random intercepts. 111

Table 19. Proportion of Insertion by Word Type and Position of Consonant Sequence. 116

Table 20. Mixed Effects Logistic Regression for Insertion as predicted by Position of Consonant Sequence, with Subject and Word as random intercepts. 117

Table 21. Logistic Regression for Insertion as predicted by C_1 Allophone with Subject and Word as random intercepts. 123

Table 22. Absence of vowel insertion in words with C_1 deletion, derhoticization, /l/-vocalization, or syllabic C_2s. 125

Table 23. Diagnosis of the vowel as phonological, not excrescent. 125

Table 24. Diagnosis of the vowel as phonological, not excrescent. 126

Table 25. Phonetic and phonological diagnostics used to classify the vowel as phonological. 129

Table 26. Phonetic and phonological diagnostics used to classify the vowel as epenthetic. 130

LIST OF FIGURES

Figure 1. Features of each segment in the word *form* [fɔɹm]... 6

Figure 2. Diagram of the modular feedforward framework... 7

Figure 3. Example of a gestural score for the word *span*. Recreated from Browman & Goldstein (1992, p. 25). *Note*. The y-axis represents the articulators involved and the x-axis represents time... 8

Figure 4. Temporal landmarks of a gesture. Recreated from Gafos (2002, p. 271). *Note.* The y-axis represents the movement of the articulators and the x-axis represents the time it takes to complete the gesture.. 9

Figure 5. Possible coordination outcomes for pairs of gestures. Recreated from Gafos (2002, p. 271).. 9

Figure 6. Phonetic output. Recreated from Durvasula (2009, p. 184). 11

Figure 7. Diagram of the syllable components in the word *dog* [dɔg]............................ 14

Figure 8. Example of an open transition, in which the vocal tract is open for a brief period. Recreated from Gafos (2002, p. 284). *Note.* "C-center" means center of a consonant gesture.. 16

Figure 9. Scottish English monophthongs. Recreated from (Scobbie et al., 2006, p. 6). *Note.* Height (High-Low) corresponds with F1 (Hertz) values, while frontness (Front-Back) corresponds with F2 (Hertz) values. The shaded area in gray represents lax vowels, while the dotted border surrounds vowels which are rounded. 29

Figure 10. Gestural coordination of in-phase and anti-phase gestures. Based on Shaw and Kawahara (2018). *Note.* C_1=first consonant, V=vowel, C_2=second consonant. A vertical blue line represents the simultaneous timing of gestures. Horizontal blue lines represent the timing between gestures... 37

Figure 11. Example of tapped [ɾ] + inserted vowel (Subject 27, *worm*). *Note.* Dashed lines mark the offset of C_1 and onset of C_2, with vocalic material intervening ("V")...... 51

Figure 12. Example of approximant [ɹ] + no vowel insertion (Subject 18, *worm*)......... 52

Figure 13. Bimodal distribution of counts for Vowel Duration (ms) by Insertion. 58

Figure 14. Average duration in milliseconds (ms) for inserted and underlying vowels. . 63

Figure 15. Linear regression for preceding vowel duration and inserted vowel duration. *Note.* The blue line represents the slope of regression, while gray shading represents a 95% confidence interval. 65

Figure 16. Gestural coupling of complex consonant clusters. Solid lines represent in-phase relationships, while dotted lines indicate anti-phase relationships. Recreated from Marin & Pouplier (2010, p. 382). *Note.* C_1=first consonant in the cluster, C_2=second consonant in the cluster, V=vowel. 76

Figure 17. Acoustic consequences of gestural coordination for onset+vowel+coda sequences and for coda clusters with excrescent vowels intruding between the consonant gestures. Based on Shaw & Kawahara (2018, p. 108) *Note.* C_1=first consonant, V=vowel, C_2=second vowel. Dotted lines indicate different relative gesture lengths. A vertical blue line represents the simultaneous timing of gestures. Horizontal blue lines represent the sequential timing between gestures. 82

Figure 18. Example of [l] + inserted vowel (Subject 3, *culm*, normal speaking rate). *Note.* Dashed lines mark the offset of C_1 and onset of C_2, with vocalic material intervening ("V"). 88

Figure 19. Example of [l] + no vowel insertion (Subject 3, *helm*, normal speaking rate). 89

Figure 20. Average inserted vowel duration (ms) across subjects, by speech rate. *Note.* Box and whisker plots display median values (horizontal line), along with distribution in quartiles. 92

Figure 21. Bimodal distribution of counts for Vowel Duration (ms) by Insertion. 97

Figure 22. Proportion (%) of Vowel Insertion by Suffix. 110

Figure 23. Proportion (%) of Vowel Insertion by Suffix and Subject. 112

Figure 24. Proportion (%) of vowel insertion by Suffix, grouped by Consonant Sequence. 113

Figure 25. Proportion (%) of Vowel Insertion by Position of Consonant Sequence. 115

Figure 26. Proportion (%) of vowel insertion by Suffix, grouped by Position of Consonant Sequence. 118

Figure 27. Proportion of C_1 Allophone used by Insertion. 121

Figure 28. C_1 Allophone usage by Subject and Insertion. 122

Chapter 1: Introduction

Historically, theoretical phonology relied on introspective data provided by native speakers (e.g. well-formedness judgments, perception), since generative linguistics was traditionally concerned with linguistic *competence* (i.e. what speakers understand about the language), rather than linguistic *performance* (i.e. how speakers use the language) (Chomsky & Halle, 1968; Kawahara, 2011). Since then, well-formedness judgments have been supplemented with language production studies which utilize acoustic or articulatory recordings to study the phonetic manifestations of abstract phonological representations (see *Laboratory Phonology*, Pierrehumbert, Beckman, & Ladd, 2000).

One such linguistic phenomenon that may be better investigated using acoustic evidence is *vowel insertion*—audible vocalic or vowel-like material which surfaces in the acoustics but is absent etymologically.[1] Different types of inserted vowels exist: (1) lexicalized vowels (i.e. vowels which previously were not part of the underlying representation, but over time, have become fossilized in the underlying representation), (2) epenthetic vowels (i.e. vowels which are absent from the underlying representation but originate in the surface representation by way of a systematic, derivable process), and (3) excrescent vowels (i.e. vocalic-like material absent from underlying and surface representations that results from gestural alignment).[2] In the acoustics, it is difficult to

[1] I will refer to this acoustic vocalic material as "vowel insertion" throughout this dissertation.
[2] These three inserted vowel types are different from *lexical vowels*, which are present in the underlying representation and also present etymologically (e.g. *forum*, where the lexical vowel is underlined).

distinguish between the three vowel types when relying on perception, alone.[3] Plenty of research on vowel insertion has relied on introspective data and transcriptions from linguists, since identifying phonological patterns in vowel distribution, length, and quality can be approximated in perception. However, it is possible that listeners may fail to perceive vowels which are present in the acoustics (Harms, 1976; Warner & Weber, 2002) and conversely perceive vowels which are not (see *illusory vowels*, Dupoux et al., 1999). Utilizing impressionistic transcriptions in isolation to identify what type of inserted vowel is present acoustically may introduce confounds.

Distinguishing between these three vowel types is important since conflating them can have consequences for phonological theory. Since each vowel type originates at a different level of the theoretical framework, misidentifying one vowel type as another can contribute to the typology of that vowel type and to the place in the framework from which it originated. For example, since excrescent 'vowels' are not actual vowels but a vocalic-like articulatory artefact, misidentifying them as epenthetic vowels could alter the typology of epenthetic vowels and how we expect phonological processes to pattern.

Using an understudied case of vowel insertion in Scottish English as a case study, I generate a multi-step exclusion process by which I identify what type of inserted vowel is present in the acoustic signal. I determine whether a vowel is present in the acoustics using *visual* inspection of the spectrogram and waveform instead of auditory perception. The exclusion process I construct involves assessing the phonetic and phonological properties of the inserted vowel against several diagnostic criteria—those which firstly

[3] Note that excrescent 'vowels' are not phonological segments, since they arise from low-level articulatory constraints. I will refer to them as *vowels*, however, since this terminology is widely used in the literature pertaining to vowel insertion (Hall, 2003, 2006; Levin, 1987).

allow me to classify the vowel as phonological by excluding the possibility that it is excrescent, and subsequently those which allow me to classify the phonological vowel as epenthetic by excluding the possibility that it is lexicalized. I do so in an effort to improve the methodology available to phonologists so as to prevent misidentification which can misinform the typology and theoretical framework.

1.1 The current dissertation

Vowel insertion has been reported in Scottish English liquid+liquid and liquid+nasal consonant sequences (e.g. *farVm, filVm, hurVl*), and has been regarded by most as epenthetic or lexicalized on the basis of impressionistic transcriptions (Dieth, 1932; Murray, 1873; Scobbie et al., 2006; Wettstein, 1942; Zai, 1942).[4] However, it is possible that the inserted vowel in Scottish English is excrescent, as no acoustic data have been analyzed for this phenomenon. In this dissertation, I use Scottish English as a case study to demonstrate how inserted vowels subject to a boundary dispute—i.e. those which can be classified as phonetic (i.e. excrescent) *or* phonological (i.e. lexicalized *or* epenthetic) (Myers, 2000)—can be classified using a process of elimination.

I asked the following questions of the inserted vowel in Scottish English liquid+liquid and liquid+nasal consonant sequences:

(1) Is the vowel phonological or excrescent?

(2) If the vowel is phonological, is it lexicalized or epenthetic?

[4] Where "V" represents an inserted vowel.

To address the questions in (1) and (2), I constructed two sets of diagnostic criteria. The first set of criteria was used to classify the vowel as phonological by excluding the possibility that it is excrescent (3):

(3) Diagnostic criteria for the vowel being phonological (not excrescent)

a. Durational relations of adjacent gestures
b. Bimodal distribution of durations
c. Consonant-to-consonant duration
d. Excrescent vowels and speech rate
e. Morpho-phonological distribution
f. Interaction with other phonological processes

The diagnostics in (3a,d) are novel phonetic diagnostics that I generate using theoretical modeling. The diagnostics in (3b-c, e-f) have primarily been used by researchers to determine whether inserted vowels are epenthetic or excrescent, so using them to discern between phonological and excrescent vowels is a novel usage.

After establishing that the vowel is phonological, I used a second set of diagnostic criteria to classify the vowel as epenthetic by excluding the possibility that it is lexicalized (4):

(4) Diagnostic criteria for the vowel being epenthetic (not lexicalized)[5]

a. Widespread distribution across lexical items
b. Durations of inserted versus underlying vowels
c. Distribution in environments which are otherwise avoided by way of separate phonological processes.

The diagnostics in (4a,c) have been used in previous research but, to my knowledge, have not been used to discern between epenthetic and lexicalized vowels, specifically. The diagnostic in (4b) is a novel diagnostic that I generate based on cross-linguistic empirical

[5] Note that these diagnostics cannot be used in isolation to diagnose a vowel as lexicalized or epenthetic. Only after independently establishing that the vowel is not excrescent can these diagnostics be used to discern between the two phonological vowel types.

patterns and the typology of epenthetic vowels. The diagnostics in (3) and (4) assess the phonetic (3a-d; 4b) and phonological (3e-f; 4a,c) properties of the inserted vowel (i.e. duration and phonological visibility, respectively).

In the next sections, I will provide relevant theoretical background and introduce the two language production experiments which constitute the basis of this dissertation.

1.2 Theoretical phonological framework

There are two phonological frameworks which make different predictions for inserted vowels: (1) one which uses discrete, atemporal representations (Chomsky & Halle, 1968) and (2) one which uses gradient, temporal representations (Browman et al., 1986). (1) invokes the use of categorical, phonological *features*, e.g. a consonant being produced with voicing (i.e. [+voice]) or not (i.e. [-voice]), and is equipped to model processes considered phonological—i.e. those which display categorical behavior. Alternatively, (2) models gradient processes which arise from the articulation of sounds represented as *gestures* (Browman & Goldstein, 1992). A third framework uses *both* featural representations and gestural representations and is modelled to generate categorical, phonological processes *and* gradient, articulatory processes (Zsiga, 1997). I will briefly overview these three frameworks in the following subsections.

1.2.1 Sounds as bundles of features

In traditional generative phonology (Chomsky & Halle, 1968), each phonemic segment is comprised of phonological *features*—each binary and categorical—and segments map onto timing units called *X-slots* (Levin, 1985) (Figure 1):[6]

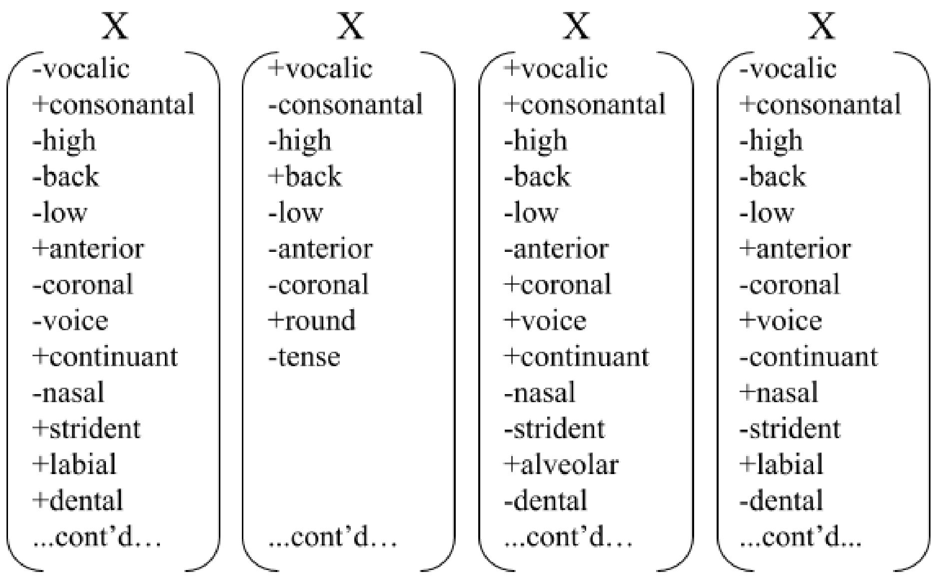

Figure 1. Features of each segment in the word *form* [fɔɹm].

Under a *modular feedforward* framework of phonology (Pierrehumbert, 2002), phonology and phonetics are distinct modules, since, although features represent *articulatory* settings (e.g. [+/- coronal]), they do not account for gradient aspects of language production, like duration of voicing or pitch. Under such a framework, the phonological module *feeds* the phonetics module in an asymmetric fashion. In the phonological module, discrete underlying representations are transformed into discrete surface representations via the implementation of language-specific phonological rules (see *Generative Phonology*,

[6] Brackets […] represent surface representations (i.e. how the item is produced), while frontslashes /…/ represent underlying representations (i.e. how the item is stored in the lexicon).

Chomsky & Halle, 1968) or ranked constraints (see *Optimality Theory*, Prince & Smolensky, 1993). Subsequently, those surface representations are passed onto the phonetics module, which result in dynamic, phonetic manifestations that unfold in real time (Figure 2).

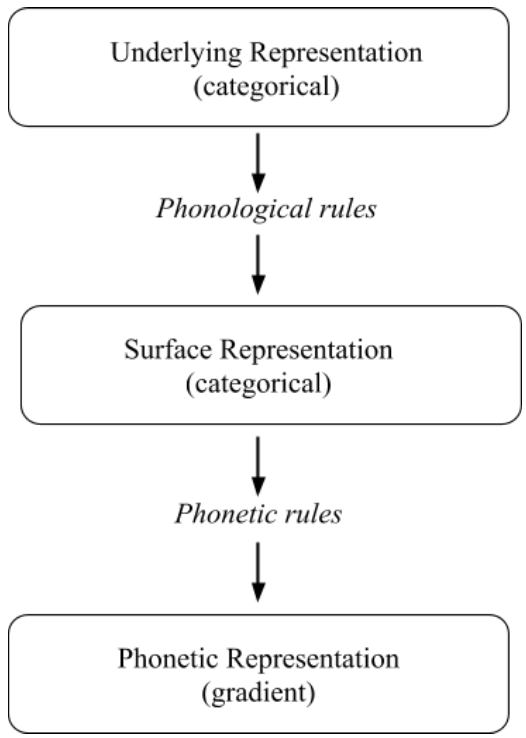

Figure 2. Diagram of the modular feedforward framework.

In this sense, the framework feeds phonology to phonetics in a serial fashion—underlying representations are either transformed into different surface representations via phonological rules or remain the same, and these items pass through the phonetics module, where they are articulated in real time (Pierrehumbert, 2002).

1.2.2 Sounds as gestures

Under an *Articulatory Phonology* framework (Browman & Goldstein, 1990), segments are referred to as *gestures*, i.e. spatiotemporal units specified for location and degree of constriction (Gafos, 2002). Each set of gestures has its own gestural score, which maps the spatial information—the set of articulators needed to produce that vocal tract constriction—and temporal information for that gesture as it is produced in time (Figure 3).

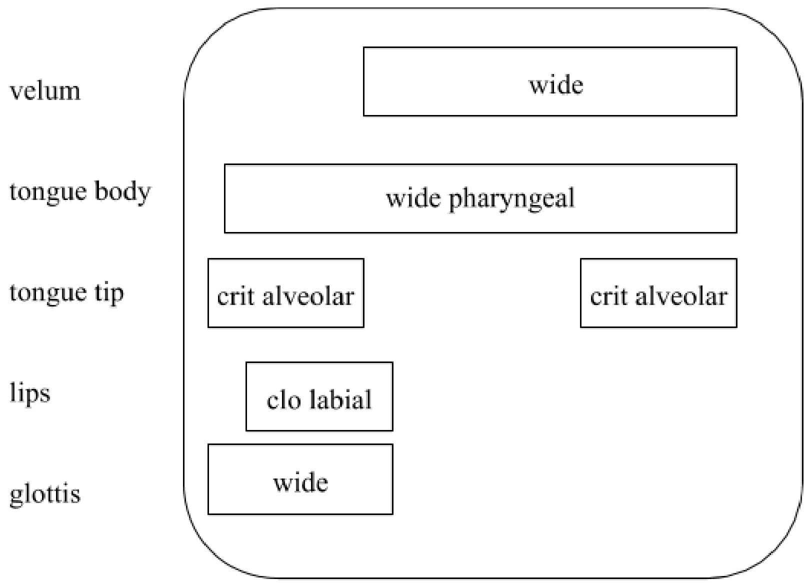

Figure 3. Example of a gestural score for the word *span*. Recreated from Browman & Goldstein (1992, p. 25).
Note. The y-axis represents the articulators involved and the x-axis represents time.

Each relative gesture has its own duration, as well as its own cycle. Gafos (2002) identifies the temporal landmarks of a gestural cycle as: the initiation or *onset*, the goal or *target*, the initiation of *release*, and the actual release of the gesture, called the *offset* (Figure 4).

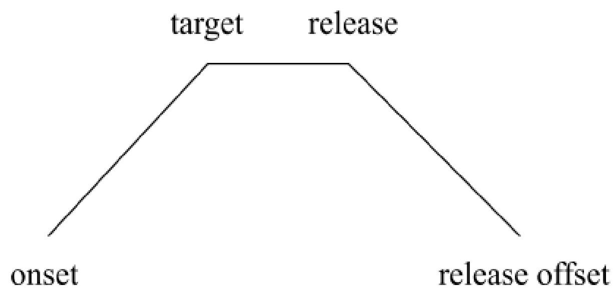

Figure 4. Temporal landmarks of a gesture. Recreated from Gafos (2002, p. 271).
Note. The y-axis represents the movement of the articulators and the x-axis represents the time it takes to complete the gesture.

While the relative duration of an individual gesture can vary, adjacent gestures can be produced with variable relative (i.e. the duration of each gesture) and/or absolute (i.e. the duration of both gestures) duration, and can vary to the degree to which they overlap (i.e. are co-articulated) with one another. Different inter-gestural coordination relationships are modeled in Figure 5.

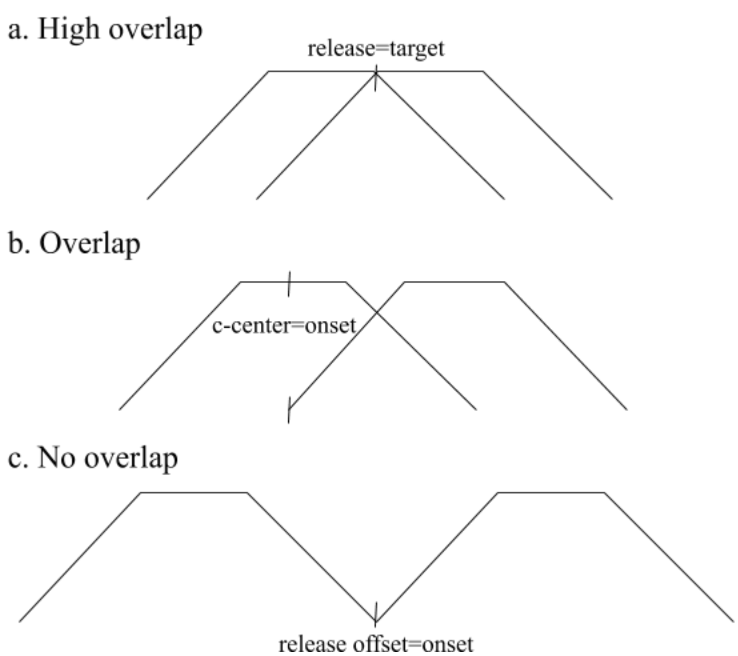

Figure 5. Possible coordination outcomes for pairs of gestures. Recreated from Gafos (2002, p. 271).

In Figure 5a, the *release* of the first gesture is aligned with the *target* of the second, creating a highly overlapped set of gestures. A *low* degree of gestural overlap is depicted in 5b, in which the consonant center, or *c-center*, of the first gesture is aligned with the *onset* of the second gesture. *No* overlap between gestures is depicted in 5c, in which the *release offset* of the first gesture is aligned with the *onset* of the second gesture. Spatiotemporal factors, like which gestures are involved in the coordination (e.g. [r] and [l] vs. [r] and [n]), the speed these gestures are produced at, and/or the relative duration of each gesture can determine the degree of overlap shared by pairs of gestures.

1.2.3 Unifying features and gestures

In an effort to unify discrete and gradient representations within a modular feedforward framework, Zsiga (1997) situates gestural phonology as an implementational *mapping* of featural phonology. Under a *featural-gestural* framework, phonology feeds the phonetics in a serial fashion, but there are two phonological components: featural phonology and gestural phonology (Figure 6).

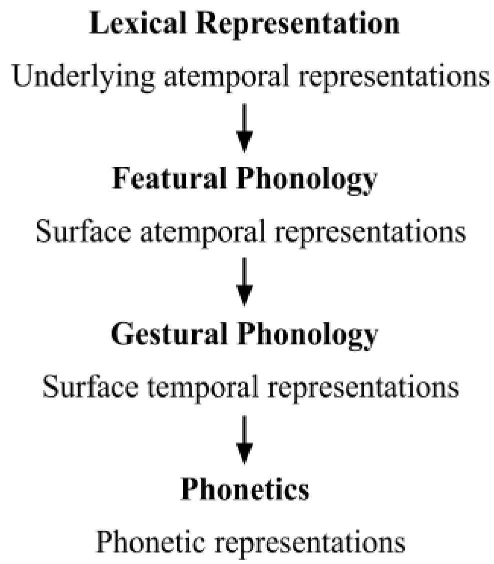

Figure 6. Phonetic output. Recreated from Durvasula (2009, p. 184).

This framework allows for discrete featural representations to be mapped to gradient gestural representations. Each segment has a bundle of features associated with it, and those features have gestural specifications, i.e. the vocal tract settings needed to produce them.

1.2.4 Types of inserted vowels in a featural-gestural framework

While a featural framework can model inserted vowels that are phonological (i.e. lexicalized and epenthetic vowels) but not excrescent vowels, and a gestural framework can model excrescent vowels but not phonological processes in which segments are inserted into a gestural score, a featural-gestural framework can account for both. To determine what stage of representation the inserted vowel in Scottish English originates from (i.e. underlying representation, surface representation, or gestural alignment) and what vowel type it is (i.e. lexicalized, epenthetic, or excrescent), I will operate under a featural-gestural framework throughout this dissertation.[7]

1.3 Three types of inserted vowels

Central to this dissertation is the question of where vowels which are not present etymologically but present acoustically originate in the derivation. I discuss the three types of inserted vowels in detail in the following subsections.

1.3.1 Lexicalized vowels: Discrete and phonological

An inserted vowel which is not present etymologically but has become engrained in the underlying representation over time is referred to as *lexicalized* (Brinton & Traugott, 2005; Pierce, 2007). These vowels should have the same acoustic quality compared to the set of attested vowels in the language, since speakers represent them as part of the vowel inventory of their language.[8] These vowels may have historically originated as non-underlying vowels, but over time became fossilized in the underlying representation for the lexical item that they surface in. In their current form, then, lexicalized vowels are not the result of an active phonological process, but are fossilized in the lexical entries of the words they surface in.[9]

[7] Note that throughout this dissertation, diagnoses of vowel types are constrained by the theoretical framework employed. For example, had I operated under a purely featural framework, I would only be able to diagnose the vowel as either lexicalized or epenthetic, since excrescent vowels are not generated by a purely featural framework.
[8] Unless a new vowel type is introduced, in which case, it would not have a comparable lexical item in the existing inventory.
[9] Note that the modular feedforward framework which I use throughout this dissertation models synchronic processes. Since lexicalization is a diachronic process, I use a modular feedforward framework to determine *where* inserted vowels originate from in the framework, rather than *how* lexicalized vowels become cemented in the underlying representation.

1.3.2 Vowel epenthesis: Discrete and phonological

In any given language or dialect, certain sequences of segments may be considered less preferential or altogether illicit (e.g. not attested in the language). This is regarded as a constraint on the combinatoric possibilities between phonological units, or *phonotactics* (Chomsky & Halle, 1965, 1968; Kenstowicz, 1994; Stanley, 1967). Phonotactic violations can be improved upon or fixed via phonological repair strategies. For example, if two adjacent segments are dis-preferred or illicit, the deletion of one of the two segments or the addition of an intervening segment can repair the sequence.

The insertion of a vowel as a phonological repair strategy is called *vowel epenthesis*. A phonological process like vowel epenthesis would apply via a phonological rule (see *Generative Phonology*, Chomsky & Halle, 1968) or by the selection of an optimal surface representation that violates the least amount of phonological constraints (see *Optimality Theory*, Prince & Smolensky, 1993). Although epenthesis can have many motivations, I will focus on a very common motivation for epenthesis: syllable structure.

The syllable is an abstract phonological representation which contains a mandatory *nucleus*—which a vowel or syllabic consonant can occupy—with an optional *onset* and *coda* (Figure 7).

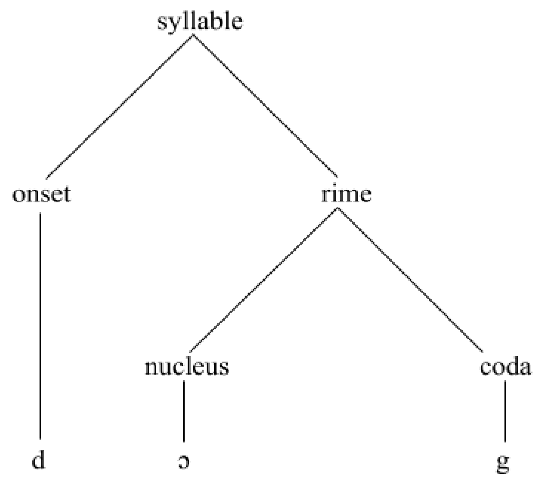

Figure 7. Diagram of the syllable components in the word *dog* [dɔg].

More than one segment may occupy an onset or coda, as part of a *cluster* (e.g. *film*).[10,11] Under the Sonority Sequencing Principle (SSP), a *preferred* syllable structure is one which has the most sonorous segment as the nucleus, with segments in the onset position rising in their sonority values towards the nucleus, and segments in the coda position falling in their sonority values away from the nucleus (Kahn, 1980; Selkirk, 1984; Sievers, 1881). The Sonority Dispersion Principle (SDP) builds off this principle, governing that onset+vowel sequences should maximize the dispersion of sonority, while vowel+coda sequences should minimize sonority dispersion (Clements, 1992). Clements (1990) provides a universal sonority scale for the natural classes of speech sounds (5).

(5) *Most sonorous* *Least sonorous*

vowels > glides > liquids > nasals > obstruents

[10] The relevant cluster is underlined.
[11] Although some argue that these consonants are word-final appendices (Borowsky, 1986; Vaux & Wolfe, 2009), I will regard them as coda clusters, since prior reports of vowel insertion in Scottish English have regarded them as such.

For consonants broken up by a syllable boundary, the Syllable Contact Law states that consonants preceding a syllable boundary should have a higher sonority than consonants following a boundary, so that sonority maximally falls across the boundary (Clements, 1990; Vennemann, 1988). For example, *metro* is syllabified as *me.tro,* since the fall from the first vowel to the obstruent consonant /t/ involves a maximal fall, and not *met.ro*, since this creates a *reverse* sonority cline in which /t/ to /ɹ/ has a rising sonority (Vennemann, 1988).[12,13]

Epenthesis can serve to repair violations of syllable structure, like reverse sonority clines and sonority plateaus. Reverse sonority clines are those which have the *opposite* sonority sequencing outlined by the SSP, while a *sonority plateau* refers to segments which have identical sonority values (e.g. liquid+liquid sequences).[14,15] Epenthetic vowels can break up consonant clusters with dis-preferred sonority sequencing by offering a viable syllable nucleus to alter the syllable position of each consonant (e.g. *hurl* [hə.ɹVl]).[16,17]

1.3.3 Vowel excrescence: Gradient and phonetic

Excrescent 'vowels' are a vocalic transitional sound that intrudes between gestures which overlap to a low degree (Gafos, 2002; Hall, 2003; Levin, 1987). This quasi-vowel is an unintended consequence of low-level articulatory constraints. For example, consonant

12 Where "." indicates a syllable boundary.
13 Frontslashes /…/ surrounding individual speech sounds represent phonemic uses (i.e. underlying representations).
14 Of the two violations, plateaus are considered less atypical, or *marked*, and are therefore more preferential than reverse sonority clines (Clements, 1990; Kreitman, 2006).
15 I will return to the discussion of epenthesis driven by reverse sonority clines and sonority plateaus in Chapter 5.
16 Brackets […] surrounding individual speech sounds represent allophonic uses (i.e. surface representations).
17 Where "V" is used to represent an epenthetic vowel, and "." indicates a syllable boundary. In this case, the inserted vowel resyllabifies the sequence so [ɹ] is no longer in the coda, but the onset of a new syllable.

clusters produced with an *open* transition, rather than a *close* transition, can result in an acoustic release because their alignment creates a period in which the vocal tract is left open (Gafos, 2002) (Figure 8).

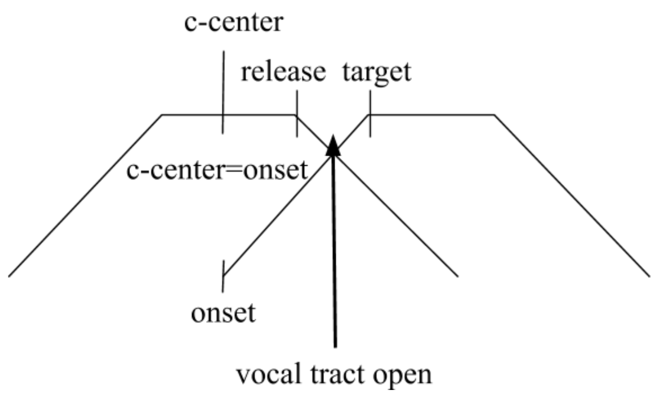

Figure 8. Example of an open transition, in which the vocal tract is open for a brief period. Recreated from Gafos (2002, p. 284).
Note. "C-center" means center of a consonant gesture.

This unintended consequence of how adjacent gestures are produced, then, is starkly different from a deliberate process of vowel insertion as a phonological repair strategy (i.e. epenthetic vowels) or a true vowel which is present in the lexicon (i.e. lexicalized or lexical vowels). Crucially, since excrescent vowels do not have an associated gesture (Figure 8) or segment, confusing them with phonological vowels can alter the typology of phonological vowels.

1.4 Distinguishing between types of inserted vowels

1.4.1 Existing criteria for diagnosing excrescent and epenthetic vowels

Most of the prior research on inserted vowels has centered around determining whether an inserted vowel is epenthetic or excrescent. The most comprehensive diagnostic criteria comes from Hall (2003, 2006), who compiled patterns of each vowel type using data from a multitude of cross-linguistic studies. Hall's (2006) criteria for phonologically invisible excrescent vowels and for phonologically visible epenthetic vowels are provided in (6) and (7), respectively (p. 391).

(6) Properties of excrescent vowels

a. The vowel's quality is either a schwa, a copy of a nearby vowel or influenced by the place of the surrounding consonants.
b. If the vowel copies the quality of another vowel over an intervening consonant, that consonant is a sonorant or guttural.
c. The vowel generally occurs in heterorganic clusters.
d. The vowel is likely to be optional, have a highly variable duration or disappear at fast speech rates.
e. The vowel does not seem to have the function of repairing illicit structures. The consonant clusters in which the vowel occurs may be less marked, in terms of sonority sequencing, than clusters which surface without vowel insertion in the same language.

(7) Properties of epenthetic vowels

a. The vowel's quality may be fixed or copied from a neighboring vowel. A fixed-quality epenthetic vowel does not have to be schwa.
b. If the vowel's quality is copied, there are no restrictions as to which consonants may be copied over.
c. The vowel's presence is not dependent on speech rate.
d. The vowel repairs a structure that is marked, in the sense of being cross-linguistically rare. The same structure is also likely to be avoided by means of other processes within the same language.

While most of the epenthetic and excrescent vowels surveyed by Hall (2003, 2006) exhibit the properties listed in (6) and (7), Hall (2003, 2006) identifies three languages which have 'exceptional' excrescent vowels: Scots Gaelic, Dutch, and Hocank.[18] In the next section, I will discuss how these typological 'exceptions' may be the result of methodological

[18] Which are later regarded as epenthetic in (Hall, 2011).

limitations and discuss why updating the diagnostic criteria used to identify epenthetic and excrescent vowels is necessary.[19]

1.4.2 Why distinguishing between inserted vowel types is important

While misidentifying inserted vowels can have implications for the typology of each vowel type, classifying an inserted vowel as excrescence when it is actually epenthetic can have implications for the modular feedforward framework. Excrescent vowels should be invisible to the phonology (Hall, 2003, 2006) for two primary reasons: (1) they originate outside of the phonology module, and (2) the phonetics module cannot access the phonology module, since the phonology module *feeds* the phonetics module in an asymmetric fashion. A modular feedforward framework, then, is not equipped to generate an excrescent vowel which displays partial phonological visibility. While it is possible that the exceptional excrescent vowels identified by Hall (2003, 2006) reflect true typological exceptions, it is also possible that their seemingly exceptional behavior is merely the result of misclassification induced by methodological limitations. This is possible because the inserted vowels in Hocank, Dutch, and Scots Gaelic have been found to display phonologically visibility insofar as they interact with other phonological processes (for Scots Gaelic, see Bosch & Jong, 1997; Hammond et al., 2014; for Dutch, see Warner et al., 2001; for Hocank, see Stanton & Zukoff, 2018).[20,21] Treating these vowels as

19 In this dissertation, I use different criteria from those provided in (6) and (7) to expand upon the diagnostics available to researchers investigating inserted vowels.

20 In a series of studies on syllabification, Hammond et al. (2014) found that the inserted vowel syllabified to the left and also to the right. It is possible, then, that the inserted vowel is optionally resyllabified or that the preceding consonant is ambisyllabic (i.e. both the coda of the preceding syllable and onset of the following syllable) in tokens containing an inserted vowel.

21 Note that Stanton & Zukoff (2018) argue that the vowel is epenthetic and that its resistance to interacting with a separate process of reduplication is because the vowel is not associated with a full syllable mora.

excrescent but considering them to be typological ‘exceptions’ may inadvertently lead phonologists to develop new categories of vowels which do not fit the phonology-phonetics dichotomy (Blevins & Pawley, 2010) or conclude that the modular feedforward framework need be revised to allow for some degree of overlap at the phonetics-phonology interface (Hammond et al., 2014). Although these are worthwhile lines of inquiry in their own right, pursuing these possibilities to account for ‘exceptional’ typological behavior which arises from methodological limitations is problematic. If we treat ‘exceptional’ excrescent vowels which display phonological visibility as epenthetic vowels that have been misidentified, there would be no need to change the typology of excrescent and epenthetic vowels or to try to adjust the framework to ensure that it can generate these exceptional behaviors, rather, we would merely need to update our methodology. For this reason, I use this dissertation as an opportunity to further develop the methodology used by phonologists to classify inserted vowels as lexicalized, epenthetic, or excrescent so as to ensure that inserted vowels which are subject to a boundary dispute (Myers, 2000) are not misidentified.

1.5 The current studies

In this dissertation, I present two language production experiments and generate a multi-step exclusion process which assesses the phonetic and phonological properties of the inserted vowel against two sets of diagnostic criteria. I do so with the goal of determining where the inserted vowel in Scottish English liquid+liquid and liquid+nasal sequences originates from—the underlying representation (i.e. lexicalization), the surface representation (i.e. epenthesis), or gestural alignment (i.e. excrescence). The exclusion process, along with three diagnostics, are novel methodological outputs of this dissertation. Additionally, the diagnostics which I use that have been used in previous research are implemented in a novel way (i.e. by using them to distinguish between phonological and excrescent vowels and subsequently between epenthetic and lexicalized vowels, which are both phonological). In three separate chapters (3, 4, and 5), I treat the vowel as "inserted" and use the multi-step exclusion process to eliminate the possibility that the vowel is excrescent, and subsequently that it is lexicalized, in order to conclude that vowel insertion in Scottish English is a process of phonological vowel epenthesis.[22]

1.6 Organization of this dissertation

This dissertation will be organized as follows. In Chapter 2, I will provide a language background for Scottish English as it pertains to inserted vowels in liquid+liquid and liquid+nasal consonant sequences. In Chapter 3, I will present the first language production study conducted on vowel insertion in Scottish English, in which I use five diagnostics to classify the vowel as epenthetic. In Chapter 4, I present another language production experiment in which I investigate whether the vowel interacts with speech rate adjustments in the same way that phonological segments should. This diagnostic, along with the diagnostics used in Chapter 3, are used to confirm the classification made in the previous chapter. In Chapter 5, I use three diagnostics that assess the vowel's morpho-phonological distribution, interaction with other phonological processes, and distribution in environments which are avoided by way of other phonological processes to confirm the diagnoses made in Chapters 3 and 4. In Chapter 6, I summarize the major findings of this

[22] The results of Chapter 4 confirm the diagnosis made in Chapter 3, and the results of Chapter 5 confirm the diagnoses made in Chapters 3 *and* 4.

dissertation and discuss the language-specific, methodological, and theoretical implications of this work. I will conclude by discussing how the diagnostics employed in this dissertation can help aid in the classification of inserted vowels as lexicalized, epenthetic, or excrescent, and how determining where inserted vowels originate in the derivation—whether it be in the underlying representation, surface representation, or in gestural alignment—help us better understand the modular feedforward framework.

Chapter 2:
Language Background

Since this dissertation investigates vowel insertion in Scottish English, I will provide a brief overview of the language in this chapter. Scottish English is spoken in Scotland, which is a part of the United Kingdom, along with Wales, England, and Northern Ireland. The 2011 national census reported the population of Scotland to be 5,295,403 (National Records of Scotland, 2011), with the majority of the population living in the Central Belt region of Scotland, which includes the major cities Glasgow and Edinburgh. The three primary languages spoken in Scotland are Scottish Standard English (SSE), Broad Scots (Scots), and Scots Gaelic, but for the remainder of this dissertation I will focus on the first two, which form the poles of a bipolar linguistic continuum referred to as "Scottish English" (Aitken, 1984, 1979; Stuart-Smith, 2008).[23]

In Section 2.1, I will provide an overview of the usage of each pole of the bipolar linguistic continuum (i.e. Scottish Standard English and Broad Scots). In Section 2.2, I will provide a brief phonology of Scottish English.[24] In this section, I will pay special attention to processes which affect rhotics, laterals, and vowel length, as they are directly involved in the phenomenon that this dissertation is concerned with—vowel insertion in liquid+liquid and liquid+nasal consonant sequences. I will then overview previous references which have been made to vowel insertion, which I will directly investigate in Chapters 3, 4, and 5.[25] Lastly, I will conclude in Section 2.3.

[23] To be discussed in detail in Section 2.1.
[24] Note that "Scottish English" is not synonymous with Scottish Standard English. The phonology presented here represents a bipolar linguistic continuum which includes *both* SSE and Scots.
[25] I will use the term "inserted" throughout the dissertation to take a neutral stance on the function of the vowel, as the purpose of this dissertation will be to classify the vowel as lexicalized, epenthetic, or excrescent.

2.1 Scottish English: A linguistic continuum

The bipolar linguistic continuum referred to as "Scottish English" is the result of language change and contact (Stuart-Smith, 2008), in which speakers vary with respect to where they lie on the continuum. Aitken (1979) proposes that discrete style-*switching*, in which speakers categorically switch between each pole of the continuum, occurs more frequently in rural varieties of Scottish English, while style-*drifting*, in which speakers gradiently drift along the continuum, occurs more frequently in urban varieties, particularly those spoken in Glasgow and Edinburgh (Aitken, 1979, 1984; Stuart-Smith, 2008). As this dissertation is concerned with Central Belt Scottish English—particularly as spoken in Glasgow and Edinburgh—I will provide a brief overview of the usage of Broad Scots (2.1.1) and Scottish Standard English (2.1.2), which represent each pole of the Scottish English continuum.

2.1.1 Broad Scots

Broad Scots (Scots) has four primary dialects: Mid/Central Scots, Southern/Border Scots, Northern/Doric Scots, and Insular Scots, with the central variety—spoken in the Central Belt region of Scotland—often being referred to as 'Urban Scots' (Stuart-Smith, 2003, 2008). The results of the 2011 census indicate that 3,188,779 citizens reported having no proficiency in using or understanding the Scots language (National Records of Scotland, 2011). Of the citizens who reported proficiency with the Scots language (around 30% of the population), 267,412 people were not actually speakers of Scots themselves but reported being able to understand it; 1,225,622 reported being able to speak, read, and write Scots; 179,295 reported being able to speak but not read or write Scots; 132,709 reported being able to speak and read but not write Scots; 107,025 reported being able to read but not speak or write Scots; and 17,381 reported having other combinations of Scots proficiency (National Records of Scotland, 2011).

2.1.1.1 Sociolinguistic variation in Scots

Scots is often spoken by working class speakers (Stuart-Smith, 2003), with many older speakers retaining a Scots lexis in addition to a Scottish Standard English lexis (Scobbie et al., 2006), while Scottish Standard English is often spoken by middle class speakers (Stuart-Smith, 2003).[26]

The 2011 census was the first to explicitly ask about individuals' proficiency in using the Scots language. Prior to the census, those attempting to survey the number of Scots speakers in Scotland noted that self-assessment provided a challenge to obtaining an accurate count, since some speakers have difficulty self-identifying as Scots speakers. Stuart-Smith (2008) has attributed this difficulty to a few sociolinguistic factors: (1) Some native speakers and linguists, alike, regard Scots as a dialect of Scottish Standard English, rather than its own language (Macafee, 1997; Scobbie et al., 2006; Stuart-Smith, 2008), which has been regarded as offensive by some speakers of Scots (e.g. in Ulster-Scots, see Craith, 2001); (2) mixed attitudes toward Scots—while some regard it as "traditional", others consider it "degenerate" (Stuart-Smith, 2008, p. 49); (3) varying levels of competence in Scots; and (4) "dialect levelling towards English throughout Scotland" (Stuart-Smith, 2008, p. 50). The individuals identified as most likely to self-identify as Scots speakers are the individuals who use it the most—older populations of working-class speakers (Maté, 1996; Murdoch, 1995).

[26] "Lexis" refers to the set of words available in the language.

2.1.2 Scottish Standard English

Scottish Standard English (SSE) refers to "Standard English spoken with a Scottish accent" (Stuart-Smith, 2008, p. 48). Only 8,615 people who took the 2011 census reported having no conversational proficiency in English and not being able to understand it, and 98,320 reported being able to understand but not use English (National Records of Scotland, 2011). Stuart-Smith (2008) reports that there is not much regional variation for Scottish Standard English, but that its usage varies with several sociolinguistic variables.

2.2 Scottish English Phonology[27]

In the previous section, I briefly described the usage of each pole of the bipolar linguistic continuum referred to as "Scottish English". In this section, I will present an overview of Scottish English phonology, which represents the full continuum.

2.2.1 Syllable structure

The minimal syllable size for Scottish English is a single vowel. Complex onset and coda clusters are permitted, with the largest permitted onset cluster being CCCV—restricted to /s/-initial tri-consonantal clusters (e.g. /str-/ *string*, /spr-/ *spring*)—and the largest coda cluster being VCCCC (i.e. *strengths* /-ŋkθs/). However, some have argued that the entire consonant sequence in VCCCC sequences produced in English dialects does not necessarily form a coda, rather, some consonants may be word-final appendices (Borowsky, 1986; Vaux & Wolfe, 2009).

[27] Note that "Scottish English" is not synonymous with Scottish Standard English. The phonology I present here is for Scottish English, the bipolar linguistic continuum including *both* SSE and Scots.

2.2.2 Consonant inventory

Scottish English has 27 consonants. These include the 25 consonants shared by other dialects of English, plus the velar fricative /χ/ (e.g. *loch*) and labial velar fricative /ʍ/ (Table 1).[28,29]

Table 1. Scottish English consonant inventory. Recreated from Scobbie et al. (2006, p. 5).

	Bilabial	Labiodental	Dental	Alveolar	Post alveolar	Palatal	Velar	Glottal
Plosive	pʰ b̥				tʰ d̥		kʰ g̊	ʔ
Nasal	m				n		ŋ	
Fricative	ʍ	f v	θ ð	s z	ʃ ʒ		χ	h
Approx	w				ɹ	j		
Lateral approximant				ɫ				
Affricate					t͡ʃ d͡ʒ			

Note. In this table, /ɹ/ is used to represent all rhotic allophones.

In the subsections that follow, I will describe the attested allophones for rhotics and laterals, since the phenomenon investigated in this dissertation is vowel insertion in liquid+nasal and liquid+liquid consonant sequences.

28 Frontslashes /…/ represent phonemic uses. For example, the phoneme /ɹ/ is used in Table 1 but has many different allophones, which I will indicate as allophonic usages with brackets […].

29 Note that I include /ʍ/ in the consonant chart, following (Wells, 1982), but the status of this consonant in the phonemic inventory has been the subject of debate (Scobbie & Stuart-Smith, 2008). Given that this dissertation is not concerned with /ʍ/, I have kept it in the inventory to remain neutral.

2.2.2.1 Rhotics

Scottish English has the following rhotic allophones: alveolar trilled [r], alveolar tapped [ɾ], postalveolar approximant [ɹ], and post-alveolar and retroflex approximant [ɹ]. In word-initial positions, the rhotic in /tɹ/ and /dɹ/ onset clusters is a rhotacized post-alveolar affricate (Scobbie et al., 2006). Word-medially, tapped [ɾ] usage is frequent in trochaic (e.g. *carry*), pre-vocalic, and postvocalic positions (Johnston, 1997; Marshall, 2004). Post-alveolar and retroflex approximant [ɹ] is also found in post-vocalic positions (Johnston, 1997). A preference for using tapped [ɾ] in intervocalic position has been found in Dundee SSE, where rhotics in intervocalic position are almost categorically tapped (Jauriberry et al., 2015).[30] Word-finally, trilled [r] is found in working-class speech—albeit rarely (Stuart-Smith, 2007).

2.2.2.1.1 Sociolinguistic use of rhotics: Derhoticization

Although Scottish English is considered a rhotic language (i.e. one that does not categorically delete coda /ɹ/ which does not come before a vowel-initial word), for a period of time in the 19th and 20th century, some middle-class speakers became non-rhotic (Johnston, 1985; Wells, 1982). Now, although middle-class Scottish English is primarily rhotic, working-class speech displays weakened rhoticity (Lawson, Scobbie, Stuart-Smith, 2014; Stuart-Smith et al., 2014). Central Belt Scottish English speakers may weaken or altogether delete coda /ɹ/ in unstressed syllables which are word-final (Stuart-Smith et al., 2014) in a socially stratified way (Lawson et al., 2014).[31] This change has been recorded

[30] In Chapter 5, I will show that vowel insertion co-occurs with rhotic allophony, in which speakers switch between rhotic allophones in words with and without vowel insertion.
[31] Except for coda /ɹ/ followed by a vowel-initial word.

in the lowland Central Belt region for the past several decades (Lawson et al., 2008). This process—referred to as *derhoticization*—is a diachronic, gradient shift towards deletion of trilled [r] or a synchronic change in which /ɹ/ is acoustically weakened (Stuart-Smith et al., 2014, p. 2). For example, for working-class speakers who use approximant [ɹ], the acoustic quality may become pharyngealized, velarized, or retracted if the preceding vowel is low (Johnston & Speitel, 1983; Scobbie et al., 2006; Stuart-Smith, 2003, 2007). Conversely, middle-class speakers—particularly females—retain an acoustically strong approximant [ɹ] that is occasionally hyperarticulated (Lawson, Scobbie, Stuart-Smith, 2014; Lawson et al., 2013).

2.2.2.2 Laterals

Most dialects of English exhibit /l/ allophony—a categorical switch between light [l] in syllable onsets (e.g. *lip*) and dark [ɫ] in syllable codas (Boersma & Hayes, 2001). The difference between the two lateral allophones is that light [l] is produced with less tongue retraction than dark [ɫ] (Gartenberg, 1984). Scottish English, however, has dark [ɫ] in all syllable positions (Johnston, 1997; Wells, 1982).

2.2.2.2.1 Sociolinguistic use of laterals: L-vocalization

/l/-vocalization, a process by which coda /l/ loses contact with the alveolar ridge and diachronically vocalizes, has been recorded in Scots and analyzed as a marker of sociolinguistic identity (Scobbie & Wrench, 2003; Stuart-Smith et al., 2006). In Glasgow, a new type of /l/-vocalization has been noted, in which /l/ is vocalized to a high back vowel (i.e. *people* [pipo]) (Stuart-Smith et al., 2006). This type of /l/-vocalization has been regarded as distinctly separate from the traditional Scots /l/-vocalization in that it occurs post-consonantally (e.g. *people*) and pre-consonantally (e.g. *shelf, milk*), i.e. environments that otherwise do not condition traditional /l/-vocalization (Stuart-Smith et al., 2006).

2.2.3 Scottish English: Vowels

There are 14 vowels in Scottish English—10 monophthongs and four diphthongs (i.e. /ae/ /əi/, /ʌʉ/, /oe/) (Stuart-Smith, 2004). The monophthongs which can appear in both open and closed syllables are: /ɔ/, /ʉ/, /i/, /e/, /o/ (Scobbie et al., 2006). A vowel space chart for the attested monophthongs in Scottish English is provided in Figure 9.

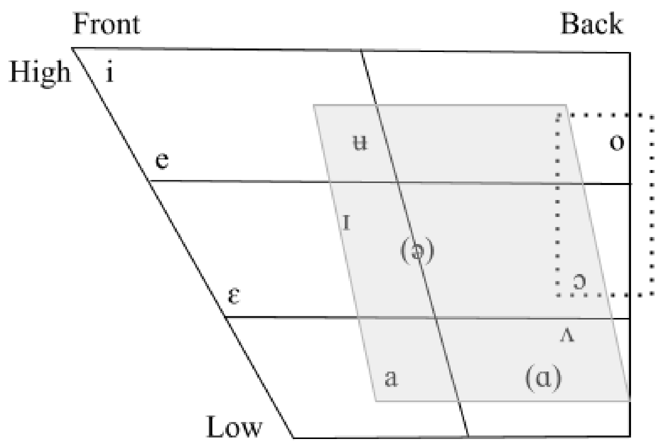

Figure 9. Scottish English monophthongs. Recreated from (Scobbie et al., 2006, p. 6). *Note.* Height (High-Low) corresponds with F1 (Hertz) values, while frontness (Front-Back) corresponds with F2 (Hertz) values. The shaded area in gray represents lax vowels, while the dotted border surrounds vowels which are rounded.

Schwa (i.e. /ə/) occurs in unstressed positions, as indicated by the parentheses in Figure 9. The parentheses around the low back vowel /ɑ/ are used to indicate that the phoneme is produced by a subset of speakers, i.e. those that distinguish between *pam-palm*, which, for most speakers, are homophonous, like the minimal pairs *cot-caught* and *pull-pool* (Scobbie & Stuart-Smith, 2008).

2.2.3.1 Scottish Vowel Length Rule

Scottish English has a "quasi-phonemic contrast" between the vowels used in minimal pairs like those listed in (8) (Scobbie & Stuart-Smith, 2008, p. 9).

(8) *Minimal Pair* *Phonetic Transcription*

a. need-kneed [nid]-[niːd][32]

b. rude-rued [ɹʉd]-[ɹʉːd]

c. side-sighed [said]-[saiːd]

In (8a-c), the first member of each pair contains a short vowel, while the second member contains a long vowel, even though both members of each pair use the same vowel—/i/, /ʉ/, and /ai/, respectively. This process is referred to as the *Scottish Vowel Length Rule* (SVLR) or *Aitken's Law* (Aitken, 1981)—a rule which creates additional phonemes by distinguishing between short and long variants of /i/, /ʉ/, and /ai/. Long variants are conditioned by post-vocalic voiced fricatives and /ɹ/, open syllables (e.g. *brew* [bɹʉː]), and open syllables that have undergone suffixation (e.g. *brewed* [bɹʉː+d]) (Scobbie & Stuart-Smith, 2008).[33] Although this process has become less active for some middle-class speakers in Edinburgh (Jones, 2002), the SVLR is prevalent across Scotland (Stuart-Smith, 2004).

2.2.3.2 Vowel insertion

The focal point of this dissertation is a process by which audible vocalic material variably appears in liquid+liquid and liquid+nasal consonant sequences—primarily coda clusters of

[32] "ː" represents long vowel variants.
[33] Where "+" indicates a morpheme boundary.

monosyllabic, monomorphemic words. This process has been widely regarded as phonological vowel *epenthesis*, by which a vowel is (optionally) deliberately inserted by the speaker to repair dis-preferred phonological structure. Since prior reports of this process do not justify why they classify the vowel as epenthetic, and alternatives exist, I will use the label "inserted" throughout this dissertation descriptively to identify the relevant vocalic material. A review of the published reports of vowel insertion in Scottish Standard English (9) and Scots (10) is provided below.[34]

(9) Vowel insertion reported in Scottish Standard English

Sequence	*Word(s)*	*Dialect*	*Source*
/ɹm/	farm	--	Scobbie et al. (2006)
/ɹl/	world	--	Scobbie et al. (2006)
/lm/	film	--	Scobbie et al. (2006)

(10) Vowel insertion reported in Scots

Sequence	*Word(s)*	*Dialect*	*Source*
/ɹm/	arm, farm, storm, term, warm(ed)	Berwickshire	Wettstein (1942)
	barm, warm, worm	Morebattle	Zai (1942)
	farm	Buchan	Dieth (1932)
	arm, harm, worm, warm, term	Lowland	Murray (1873)
	firm, form, term	--	Maguire (2017)
	arm, harm, warm, worm	--	Wilson (1915)
/lm/	elm, film, helm	Berwickshire	Wettstein (1942)
	elm	Morebattle	Zai (1942)
	elm, helm, film	Lowland	Murray (1873)
	film	--	Maguire (2017)
/ɹl/	curl, pearl	Berwickshire	Wettstein (1942)
	arles, carl, dirl	Morebattle	Zai (1942)
	hurl, pirl, skirl, dirl, arles, carl	Buchan	Dieth (1932)
	curl, dirl, world, earl	Lowland	Murray (1873)
	earl, curl	--	Maguire (2017)
/ɹn/	barn, bairn, turn	Berwickshire	Wettstein (1942)
	shorn, torn	Morebattle	Zai (1942)

[34] To my knowledge, this list is exhaustive.

burn, horn, corn	Buchan	Dieth (1932)
barn, turn	Lowland	Murray (1873)
turn, born	--	Maguire (2017)

2.3 Concluding remarks

In this chapter, I have provided an overview of the phonology of Scottish English, a bipolar linguistic continuum including Scottish Standard English on one end and Scots on the other. In this overview, I focused specifically on rhotics, laterals, vowels, and syllable structure, since they are relevant to the phenomenon which this dissertation is concerned with—vowel insertion in liquid+liquid and liquid+nasal consonant sequences.[35]

I have provided a brief overview of the published accounts of vowel insertion in Scottish English in this chapter. In Chapters 3, 4, and 5, I will discuss how the published accounts summarized in this chapter do not argue *why* they diagnose the vowel as epenthetic. In each of the following chapters, I will explain that the inserted vowel can be analyzed in many different ways—as lexicalized, epenthetic, or excrescent—since these vowels can superficially sound similar in the acoustics. I will argue that the only way to adequately classify the vowel as epenthetic is by assessing the phonetic and phonological properties of the vowel against diagnostic criteria which exclude the possibility that the vowel is lexicalized or excrescent.

[35] For a more comprehensive overview of the phonology of Scottish English, I recommend the reader consult Stuart-Smith (2004).

Chapter 3:
Phonological, not excrescent. Epenthetic, not lexicalized.

In an overview of the acquisition of Scottish English phonology, Scobbie et al. (2006) report that Scottish English "avoids /rl/, /rm/, /lm/ clusters" (p. 10), which is considered "a well-known feature" of the phonology (Maguire, 2017, p. 1). However, reports of vowel insertion in liquid+liquid and liquid+nasal clusters in Scottish English are sparse, and rely exclusively on non-acoustic data.[36] Without assessing the vowel against diagnostic criteria, it is difficult to determine whether the inserted vowel is truly epenthetic, since there are alternative explanations for why a vowel may be heard in the acoustics but absent etymologically.

One such explanation is that inserted vowels may not be phonologically inserted, but an articulatory artefact, called vowel *excrescence*. An excrescent vowel is a period of voicing in the articulation which intrudes between low-overlapping gestures by way of the vocal tract being open for a brief period of time (Hall, 2006). However, if an inserted vowel is phonological, not excrescent, it may be *either* epenthetic or lexicalized. A lexicalized vowel is one that is not originally part of an underlying representation, but becomes fossilized in a speaker's underlying (or lexical) representation over time. What distinguishes these two types of phonological vowels is that epenthetic vowels arise in the surface representation by way of an active phonological process, while lexicalized vowels arise in the surface representation because they are present underlyingly. Using acoustic

[36] I will refer to the vowel as "inserted", since the goal of this dissertation is to diagnose the vowel as epenthetic only after excluding the possibility that the vowel is *not* lexicalized or excrescent.

data that results from language production experiments can allow one to be certain that the vowel they are hearing is actually visibly present in the spectrogram and waveform.

3.1 The current experiment

In this chapter, I will present the results of a language production experiment that was conducted with the aim of using acoustic data to inform my classification of the inserted vowel as lexicalized, epenthetic, or excrescent. In this chapter, I use a multi-step exclusion process which assesses the phonetic and phonological properties of the vowel against two sets of diagnostic criteria. The first set of criteria will be used to diagnose the vowel as phonological while excluding the possibility that it is excrescent (11).

(11) Diagnostic criteria for the vowel being phonological (not excrescent)

a. Durational relations of adjacent gestures

(i) The inserted vowel will have the same acoustic consequences as underlying gestures predicted by a gestural coupling model. Specifically, the durations of the vowel and the flanking consonants will overlap more or less (i.e. have a negative or positive correlation), based on their syllable position (i.e. onset+nucleus, nucleus+coda).

b. Bimodal distribution of durations

(i) When plotting the distribution of the duration of the vowel in tokens with vowel insertion against tokens without vowel insertion (coded as having a duration of zero milliseconds), their distributions will be bimodal, not unimodal.

c. Consonant-to-consonant duration

(i) The total consonant-to-consonant duration (i.e. the duration from the acoustic onset of the first consonant to the acoustic offset of the second consonant) will be significantly larger in tokens with vowel insertion than in those without vowel insertion.

After establishing that the vowel is phonological using the diagnostics in (11), I will then use the second set of criteria to exclude the possibility that the phonological vowel is lexicalized, and conclude that it must be epenthetic. In this sense, only *after* I exclude the possibility that the vowel is excrescent can I use the second set of diagnostics to diagnose the vowel epenthetic, rather than lexicalized (12).

(12) Diagnostic criteria for the vowel being epenthetic (not lexicalized)[37]

a. Widespread distribution across lexical items.

(i) The inserted vowel will be systematically present across a large range of lexical items.

b. Durations of inserted versus underlying vowels

(i) The duration of the inserted vowel will be significantly different from the duration of underlying vowels.

The criteria in (11a) and (12b) are novel diagnostics which have not previously been used to diagnose inserted vowels. The diagnostic criteria in (11b-c) have been used in previous research to diagnose inserted vowels as epenthetic or excrescent, and the diagnostic criterion (12a) has been previously used to determine whether vowels are lexicalized or not. In this chapter, I use (11b-c) to determine whether the vowel is phonological or excrescent and (12a) to determine whether the vowel is lexicalized or epenthetic, specifically. In this sense, each of the five diagnostics used in this chapter are used in a novel way.

[37] Note that this set of diagnostics *cannot* be used in isolation to classify an inserted vowel as epenthetic. Only after establishing that the vowel is phonological can this set of diagnostics be used to exclude the possibility that the vowel is lexicalized, in which case one can conclude that the vowel must be epenthetic because the possibility that the non-underlying vowel is excrescent has already been eliminated.

In this chapter, I use the diagnostic criteria in (11) and (12) in a multi-step exclusion process, whereby I diagnose the vowel as epenthetic not solely because it displays the properties of an epenthetic vowel, but because it also does *not* display the properties of an excrescent or lexicalized vowel. Doing so will allow me to determine that the inserted vowel originates in the surface representation, rather than in the underlying representation or from gestural alignment. The exclusion process and diagnostic criteria used in this chapter will be discussed in detail in the next section.

3.2 Diagnostic criteria

3.2.1 Criteria for diagnosing the vowel as phonological

3.2.1.1 Durational relations of adjacent gestures

A model for how gestures are coordinated with one another comes from a gestural coupling framework, in which relative gestural timing can be established by the coupling of *individual* oscillators that become *relative* oscillator phases (Goldstein et al., 2009; Nam & Saltzman, 2003; Saltzman & Byrd, 2000). Gestures can have one of two target specifications: in-phase or anti-phase. An in-phase target specification is one in which both individual oscillators occur simultaneously, while an anti-phase target specification is one in which individual oscillators are triggered in a stepwise fashion (Figure 10).

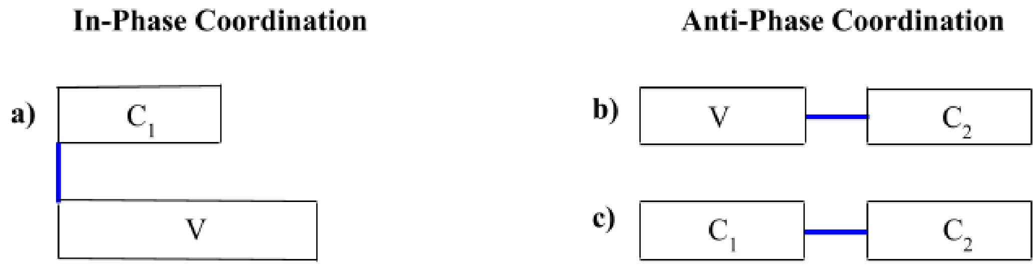

Figure 10. Gestural coordination of in-phase and anti-phase gestures. Based on Shaw and Kawahara (2018).
Note. C_1=first consonant, V=vowel, C_2=second consonant. A vertical blue line represents the simultaneous timing of gestures. Horizontal blue lines represent the timing between gestures.

Figure 10(a) depicts a singleton onset consonant and the following vowel as an example of how in-phase gestures should coordinate. The coordination of gestures in an anti-phase relationship are modelled for a vowel plus a singleton coda consonant in (b) and for complex coda clusters in (c).

Different acoustic consequences are predicted for gestures with in-phase and anti-phase relationships. Since in-phase gestures are produced simultaneously, the gestures should have a negative correlation when comparing their durations, since the duration of the vowel gesture is overlapped by the consonant gesture (Figure 10a) (Shaw & Kawahara, 2018). Conversely, since anti-phase gestures are triggered in a stepwise fashion, the second gesture should be produced sequentially after the first gesture has been completed. The gestures should maintain the same degree of timing between them, resulting in a positive correlation when comparing their relative durations (Shaw et al., 2011; Shaw & Kawahara, 2018).

This framework generates predictions for how gestures in varying syllable positions should coordinate.[38] Singleton onset consonants should have an in-phase relationship with the following vowel (C_1-V), resulting in a negative correlation, while singleton coda consonants should have an anti-phase relationship with the preceding vowel (V-C_2), resulting in a positive correlation. Lastly, the consonants in a coda cluster should have an anti-phase relationship with one another (C_1-C_2#), resulting in a positive correlation.[39,40] A summary of the predicted correlations for the durations of each pair of gestures is presented in Table 2, below.

Table 2. Expected durational correlations for coupled gestures in in-phase and anti-phase relationships.

	Pair of Gestures		
	C_1-V	**V- C_2**	**C_1- C_2#**
Coupling	In-phase	Anti-phase	Anti-phase
Gestural overlap	Triggered simultaneously. Increased gestural overlap due to coupling.	Triggered one after another. No increased overlap due to coupling.	Triggered one after another. No increased overlap due to coupling.
Correlation between durations	Negative	Positive	Positive

Note. C_1=first consonant, V=vowel, C_2=second consonant.

38 When referring to singleton consonants (i.e. those not part of a "cluster"), "C_1" will be used to indicate a singleton onset and "C_2" a singleton coda. In the context of complex consonant clusters, "C_1" will refer to the first consonant in the cluster, and "C_2" the second consonant in the cluster.

39 "#" is used to indicate a word boundary. In this example, C_1- C_2# indicates a word-final complex coda cluster.

40 The coordination of complex onset clusters will be discussed and used as a relevant comparison in Chapter 4 for a separate diagnostic, but not in Chapter 3, since the environment where vowel insertion in Scottish English is found most often is in coda clusters.

The acoustic consequences (i.e. correlations) between the relative durations of the gestures involved in a coupling relationship have been used by Shaw and Kawahara (2018) to assess whether a case of Japanese vowel devoicing is the result of increased gestural overlap of the flanking consonants or a categorical rule of vowel deletion. Shaw and Kawahara (2018) found that the durations of C_1 and V were negatively correlated, supporting an in-phase relation, and concluded that the vowel may be categorically deleted, since, when present, it displays the coupling relationship expected for C_1 and V, where V has its own gesture.[41] Following Shaw and Kawahara (2018), I will use the acoustic consequences of gestural coupling to assess whether the vowel behaves as a phonologically inserted segment or an artefact of the gestural overlap of the flanking consonants (i.e. excrescence). This experiment will be the first to use the acoustic consequences of gestural coupling to distinguish between different types of inserted vowels—specifically phonological and excrescent vowels.

A phonologically inserted vowel should bear the same coupling with the surrounding consonants as do underlying vowels in similar environments, as they have their own gesture in the gestural score. Alternatively, the consonant clusters which excrescent vowels intrude upon should have the same gestural coupling as vowel-less tokens, since excrescent vowels do not have a spot in the gestural score and are not phonological segments. The acoustic consequences of a gestural coupling framework for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens are summarized in Table 3.

[41] Where "C_1" indicates the first consonant and "C_2" the second, and "V" represents the vowel flanked by two consonants.

Table 3. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens.

	Predicted correlation between durations		
Vowel Type between CC	**C_1-V**	**V-C_2**	**C_1-C_2**
Underlying Vowel (e.g. *forum*)	Negative	Positive	*Not* positive[42]
Phonologically Inserted Vowel (e.g. *forVm*)[43]	Negative	Positive	*Not* positive
Excrescent Vowel (e.g. *for^{V}m*)[44]	--	--	Positive
No Vowel (e.g. *form*)	--	--	Positive

Note. CC=consonant-consonant sequence. C_1=first consonant, V=inserted vowel, C_2=second consonant. Vowels of interest are underlined.

In the current experiment, I will show that the inserted vowel behaves as a phonological vowel since it displays the acoustic consequences predicted for vowels which have their own gesture associated with them.

3.2.1.2 Bimodal distribution

The second diagnostic that I will use to diagnose the vowel as phonological is one that has been used in recent acoustic experiments (Bellik, 2018; Plug et al., 2019).[45] This diagnostic involves plotting the distribution of the vowel's duration in tokens with and without vowel

[42] Note that the consonants flanking a phonological vowel can have a significantly negative correlation *or* no correlation.
[43] A capitalized "V" is used to indicate a phonologically inserted vowel, to distinguish it from the underlying vowel that is present etymologically (e.g. *forum*).
[44] A superscript "V" is used to indicate a phonetic, excrescent vowel, since it is not a true segment.
[45] Note that Plug et al. (2019) used bimodality to establish that two sets of vowels in the same language were excrescent and epenthetic.

insertion. This involves marking words *without* vowel insertion as having a vowel duration of zero milliseconds.

The expectation for excrescent vowels is that their duration should merge with the duration of vowel-less tokens and together form a unimodal distribution. This is because excrescent vowels appear only between gestures which overlap to a low degree, and the degree to which the period of voicing between low-overlapping gestures varies should be very small. Conversely, phonologically inserted vowels are true gestures that are variably inserted and structurally different in their surface representation from tokens without vowel insertion. For this reason, the duration of phonological vowels should form a separate distribution from the duration of vowel-less tokens in the form of a bimodal distribution. In the current experiment, I will show that the duration of the inserted vowel forms a separate distribution from the zero-duration distribution of vowel-less tokens, suggesting that it is phonological.

3.2.1.3 Consonant-to-consonant duration

Previous research has found that excrescent vowels do not significantly add length to the duration of the consonant clusters that they intrude upon (Ridouane & Fougeron, 2011). I will use consonant-to-consonant duration, then, to exclude the possibility that the vowel is excrescent. To calculate consonant-to-consonant duration, I will use the total duration of the first consonant, vowel (if any), and second consonant in tokens with and without vowel insertion. I will show that a significant difference exists between the total consonant-to-consonant duration in tokens with and without vowel insertion, and that this difference excludes the possibility that the vowel is excrescent.

3.2.1.4 Summary of diagnostic predictions

The diagnostic predictions discussed in the previous subsections are summarized in Table 4.

Table 4. Diagnostic predictions for phonological and excrescent vowels.

Diagnostic	**Prediction for phonological vowels**	**Prediction for excrescent vowels**
Durational relations of adjacent gestures	Significantly negative correlation for C_1 and V. Significantly positive correlation for V and C_2. Significantly negative correlation for C_1 and C_2.	No significantly negative correlation for C_1 and V. No significantly positive correlation for V and C_2. Significantly positive correlation for C_1 and C_2.
Bimodal distribution of durations	Bimodal.	Unimodal.
Consonant-to-consonant duration	Duration of consonant sequence significantly lengthened in words with vowel insertion, compared to those without.	Duration of consonant sequence not significantly lengthened in words with vowel insertion, compared to those without.

3.2.2 Criteria for diagnosing the vowel as epenthetic

Following my diagnosis of the vowel as phonological using the diagnostics in the previous section, I will eliminate the possibility that the vowel is lexicalized using the criteria summarized in the following subsections, concluding that the phonological vowel in Scottish English is epenthetic.

3.2.2.1 Widespread distribution across lexical items

The first diagnostic that I will use to diagnose the vowel as epenthetic is its widespread distribution across a variety of lexical items. A process of vowel epenthesis should present as a systematic (even if variably applied) rule in which a vowel is optionally inserted to repair dis-preferred phonological structure. A lexicalized vowel, however, may evolve from an audible low-level phonetic manifestation that became phonologized, or from a prior phonological rule that became fossilized over time, with the vowel becoming part of the underlying representation. In its present form, a vowel that has undergone lexicalization should not bear the same productivity as do phonological vowels that arise by way of a synchronic process of epenthesis. For this reason, a widespread distribution across lexical items should surface for epenthetic vowels, while a narrow distribution—possibly only in a handful of lexical items in the language—should surface for lexicalized vowels. I will show that the vowel in Scottish English is not lexicalized to a handful of items, but present across a wide range of items.

3.2.2.2 Durations of inserted versus underlying vowels

Hall (2006) has claimed that excrescent vowels should have a significantly shorter duration than vowels which are underlying, however, many cross-linguistic studies also note this tendency in epenthetic vowels (Coleman, 1999; Davidson, 2006; Gouskova & Hall, 2009), which can be shorter than underlying vowels by as much as 50% (Tashlhiyt Berber, Coleman, 1999; Bruges Dutch, Sebregts, 2015).

Since epenthetic and excrescent vowels may *both* be shorter than their lexical counterparts, a better use of this diagnostic would be to use it to distinguish between vowels that originate in the underlying representation (i.e. lexicalized) and vowels that do not (i.e. epenthetic *or* excrescent vowels). A significant difference between the duration of the inserted vowel (e.g. *farVm*) and an underlying vowel in a similar phonological environment (e.g. *forum*) would suggest that the vowel is not lexicalized.[46] Furthermore, if the vowel is shorter than lexical vowels and if one can independently show that the vowel is phonological, then one can infer that the vowel is *epenthetic*, i.e. non-underlying and phonological. In the current experiment, I will show that the vowel has a significantly shorter duration than underlying vowels in similar environments, suggesting that the vowel is not underlyingly present (i.e. lexicalized), but epenthetic.

3.2.2.3 Summary of diagnostic predictions

The diagnostic predictions discussed in the previous subsections are summarized in Table 5.[47]

Table 5. Diagnostic predictions for epenthetic and lexicalized vowels.

Diagnostic	Prediction for epenthetic vowels	Prediction for lexicalized vowels
Widespread distribution across lexical items	Widespread distribution across lexical items, including novel contexts.	Narrow distribution across lexical items, occurring primarily in a handful of high frequency items.
Durations of inserted versus underlying vowels	Significantly different.	Not significantly different.

Note. These diagnostics can only distinguish between epenthetic and lexicalized vowels *after* one has eliminated the possibility that the vowel is excrescent.

46 Where "V" indicates an inserted vowel, and the relevant vowels are underlined.
47 Note that these diagnostics, in isolation, cannot allow one to distinguish whether an inserted vowel is lexicalized or epenthetic—only that a vowel is underlying or not. In this dissertation, all diagnostics used to determine whether the vowel is lexicalized or epenthetic have been used *after* having established that the vowel is phonological, not excrescent.

3.3 Organization of the chapter

In the remainder of this chapter, I will overview the methodology used in the production experiment in Section 3.4, the results of each diagnostic in Section 3.5, and my discussion of the results and conclusion that the vowel is epenthetic in Section 3.6.

3.4 Methods

3.4.1 Participants

27 participants (M=10, F=17, Mean age= 57.18, Range=21-93, SD=18.53) were recruited at The University of Edinburgh to participate in the current study using the Volunteer Panel provided by the School of Philosophy, Psychology, and Language Sciences.[48,49] Participants were native Scottish English speakers born in the Central Belt region of Scotland, who had lived there for the past 10 years. The Central Belt is home to more than half of the population of Scotland (69.5%, Stuart-Smith, 2004, p. 7), extending from Dundee, Dumbarton, Ayr, and Dunbar, encompassing Edinburgh—the capital of Scotland, Glasgow, and Stirling. Participants had normal hearing and normal-to-corrected vision, were older than 18 years of age, and were native English speakers. Participants were compensated £7 for their involvement.[50]

48 See APPENDIX A for the recruitment form used.

49 I'd like to express my gratitude for access to the Volunteer Panel.

50 I'd like to express my gratitude to the Michigan State University College of Arts and Letters for providing this funding in the form of a Research Enhancement Award.

3.4.2 Stimuli

Stimuli were sectioned into two blocks: one containing words from a Scottish Standard English lexis and one containing words specific to a Scots lexis, along with Scottish place names and surnames.

3.4.2.1 Block 1: Scottish Standard English words

38 five-word sets—e.g. *farm, farms, farmed, farmer, farming*—were constructed (APPENDIX B). Each item was repeated twice in the experiment, and all items were randomized. The experimental stimuli consisted of words containing one of the following Consonant Sequences: /ɹl/ (e.g. *twirl*), /ɹm/ (e.g. *farm*), /ɹn/ (e.g. *warn*), or /lm/ (e.g. *film*). Control items included 29 five-word sets containing other attested onset and coda clusters. Control items were included to ensure that participants did not guess the target region of each word, i.e. consonants in coda clusters (e.g. *fa<u>rm</u>*) or consonants situated across a syllable boundary (e.g. *fa<u>r.m</u>ing*).[51]

Each consonant sequence occurred in a mono- or multi- morphemic word. Multimorphemic words contained a bare root and a suffix. The suffixes used were: past tense /d/ (i.e. -ed), plural /z/ (i.e. -s), agentive /əɹ/ (i.e. -er), comparative /əɹ/ (i.e. -er), and /ɪŋ/ (i.e. -ing).[52] Example stimuli are provided in (13):

[51] Where "." indicates a syllable boundary.
[52] The comparative morpheme /əɹ/ was used in only a handful of stimuli (e.g. *warmer*, *firmer*, *calmer*) since most of the items were verbs and could only be combined with agentive /əɹ/ morphemes (e.g. *filmer, farmer*).

(13)

Root (+grammatical morpheme)	*Word*	*Phonetic Transcription*
root	*farm*	[fɑɹm]
root + /z/	*farm*s	[fɑɹm+z][53]
root + /d/	*farmed*	[fɑɹm+d]
root + /ɪŋ/	*farming*	[fɑɹ.m+ɪŋ][54]
root + /əɹ/ (agentive)	*farmer*	[fɑɹ.m+əɹ]
root + /əɹ/ (comparative)	*warmer*	[wɔɹ.m+əɹ]

In addition to the list in (13), the first block of the experiment included 7 words with an underlying vowel situated between liquid+liquid and liquid+nasal sequences, which had a CVCVC (e.g. *forum*) or CVCCCVC structure (e.g. *fulcrum*) (APPENDIX B). These words were included to compare the duration of the inserted vowel to the duration of lexical, underlying vowels in the same environment.[55]

3.4.2.2 Block 2: Scottish place names, surnames, and Scots words

The second block of the study was included to increase the number of relevant words and encourage speakers to utilize both ends of the bipolar linguistic continuum, since vowel insertion has been reported in Scottish Standard English and Scots. Given the uncertainty in both the linguistic and native speaker communities regarding whether Scots is a language or dialect (Macafee, 1997; Scobbie et al., 2006) discussed in Chapter 2, in addition to providing a survey question about whether each speaker self-identifies as a speaker of Scottish Standard English and Scots, the second block of the experiment contained Scots

[53] Where "+" indicates a morpheme boundary.
[54] Where "." indicates a syllable boundary.
[55] Note that words with underlying vowels contained the schwa lexical vowel, which was selected as the shortest vowel in the language (besides the high front vowel [ɪ], which was not present in comparable words).

vernacular to encourage and facilitate the usage of Scots, regardless of whether speakers consider themselves to be Scots speakers.

Stimuli included 127 experimental items and 70 control items (APPENDIX B). Words included Scottish place names, surnames, and Scots words collected from *Scotland's Places* (scotlandsplaces.gov.uk), *Scotland's People* (scotlandspeople.gov.uk), and the *Dictionary of the Scots Language* (dsl.ac.uk), respectively. Examples are provided in (14).

(14)

Word Type	*Position*	*Word*	*Translation*	*Phonetic Transcription*
Scots word	Coda	skirl	'to shriek'	[skʌɾɫ]
Scots word	Non-Coda	contermashious	'obstinate'	[kɔn.tər.mɑ(:).ʃəs]
Place name	Coda	Leckmelm		[ɫɛk.mɛɫm][56]
Surname	Non-Coda	Kilmer		[kIɫ.mər]

3.4.3 Procedure

Participants were seated in a soundproof booth with the experimenter in the adjacent room.[57] Participants were instructed to read aloud from a demographic questionnaire (APPENDIX C) and answer freely, in however many sentences they like, while being recorded with a Samson Go Mic Portable USB Condenser Microphone in Audacity 2.1.2 (Audacity Team, 2016).[58] Audio was recorded using the Cardoid 10db setting, with a sampling rate of 44100 Hz on a mono channel. Since the experimenter is a native speaker of General American English, it was crucial that participants ask and answer the

[56] Note that the word-initial lateral is transcribed with a dark [ɫ] because Scottish English has dark [ɫ] in all syllable positions, unlike other dialects of English (Johnston, 1997; Wells, 1982).
[57] I'd like to express my gratitude to the School of Informatics at The University of Edinburgh for providing me with access to a soundproof booth. Access was made possible via a Postgraduate Researcher opportunity in the School of Philosophy, Psychology, and Language Sciences at The University of Edinburgh.
[58] I refer the reader to APPENDIX D for a summary of participant demographics.

demographic questions in isolation to avoid speech accommodation—a process involving the convergence of a speaker's dialect and their interlocutor's dialect (Gallois et al., 2005; Giles et al., 1973).

Word lists were distributed after the completion of the demographic questionnaire and an informed consent. The experimenter instructed participants to read aloud from a word list presented on a 2008 Macbook Pro laptop as naturally as possible while being recorded with a desktop microphone.

Words were presented visually in PsychoPy 1.84.2 (Peirce, 2007) using a Rapid Serial Visual Presentation (RSVP) paradigm, which displays stimuli at timed intervals, to ensure that participants provided a pause between each word in the wordlist. This was done because, in a pilot study, it was noted that participants rushed at the end of word lists which displayed words all at once, consequently producing a string of words without pauses, which affected the acoustic quality of the target region of monosyllabic words (i.e. coda).

Instructions were displayed visually for eight seconds at the start of the experiment, detailing that each word would appear for a fixed amount of time and to "speak as you would with your family or close friends". These instructions were followed by a sentence instructing participants to "Begin with a few practice items". After participants read aloud ten practice words, a prompt on the computer screen notified them that the practice portion of the experiment had ended and to "please ask the researcher if you have any remaining questions before beginning the experiment". Words were presented one at a time in the center of the screen for 1.5 seconds and were replaced by a fixation cross for an inter-stimulus interval of .25 seconds. The two blocks of the experiment were counterbalanced and the presentation of each word was randomized across participants. In between blocks, participants were provided an automated break of two minutes and encouraged to drink water and clear their throats at this time. Sessions lasted about 50 minutes in total.

3.4.4 Acoustic measurements

For each file, maximum formant values for males were set to 5000 Hz and 5500 Hz for females. Files were resampled to 8000 Hz. The automatic detection of word and phoneme boundaries—forced alignment—was implemented using the Penn Phonetics Lab Forced Aligner (P2FA, Yuan & Liberman, 2008). A custom Scottish English pronunciation dictionary was created, since the default dictionary for P2FA contains phonetic transcriptions for General American English. Word and segment boundaries were hand-corrected using the acoustic software Praat (Boersma, 2001) following forced alignment by myself and three paid assistants. Inserted and underlying vowels were identified by visual inspection of the spectrogram and waveform.[59] The phonetic variable measured in this study was segment duration. After all of the *.wav* files were segmented, a Praat script was written which extracted the start time and end time of the segment boundaries. Segment duration was calculated by subtracting the start time from the end time. In the following subsections, I will detail how each segment was identified and measured.

3.4.4.1 Vowel measurements

The presence/absence of inserted and underlying vowels and the relevant phoneme boundaries were determined manually by inspection of the spectrogram and waveform. The following visual cues were used to determine the presence of a vowel: U-shaped

[59] Rather than with forced alignment.

curvature in the waveform, clear dark formants, and vertical striations in the spectrogram (Figure 11). The absence of an inserted vowel is presented in Figure 12 for comparison.

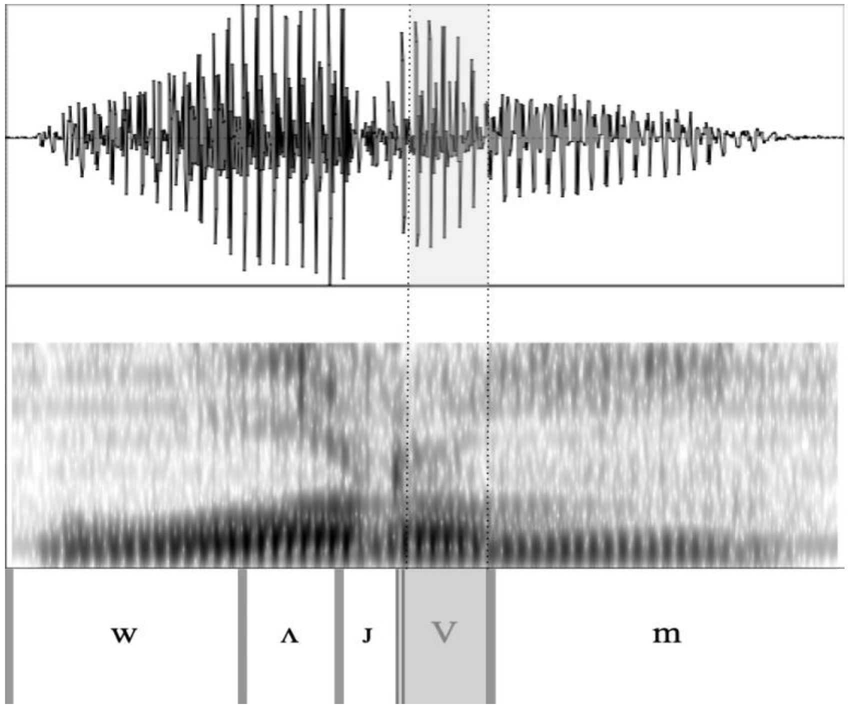

Figure 11. Example of tapped [ɾ] + inserted vowel (Subject 27, *worm*).
Note. Dashed lines mark the offset of C_1 and onset of C_2, with vocalic material intervening ("V").

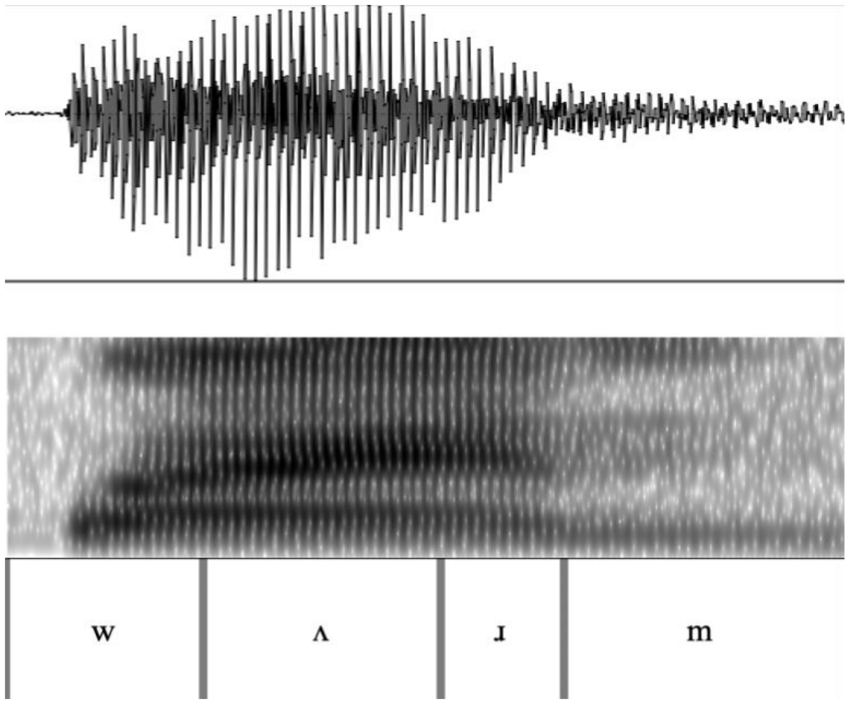

Figure 12. Example of approximant [ɹ] + no vowel insertion (Subject 18, *worm*).

Boundaries were marked at the start and end of the vowel, at the point at which there was a clear change in formant structure in the spectrogram and a change in amplitude in the waveform from the preceding and following segments.

3.4.4.2 Consonant measurements

In order to ensure that the analysis provides the most accurate description of the phonological environment which conditions vowel insertion, we did not code all rhotic-initial consonant sequences (i.e., /ɹl/, /ɹm/, /ɹn/) as [ɹ]—rather, we visually inspected the quality of C_1 to determine which rhotic allophone was used in each item.[60] We coded for five possible rhotic realizations: alveolar approximant [ɹ], trilled [r], tapped [ɾ],

[60] "We" refers to myself and three trained annotators.

derhoticized (i.e. acoustically-weakened) alveolar approximant /ɹ/, and fricativized approximant /ɹ/, along with /ɹ/-deletion.[61]

A key visual cue used to identify rhotics, especially approximant [ɹ], was the lowering of the third formant. As for tapped [ɾ], we looked for brief reduction in energy for F1 and F2 (Stuart-Smith & Lawson, 2017), as well as a reduction in amplitude and pitch across all formats, slight frication after the tapped portion, and a duration of approximately 25 milliseconds. For trilled [r], we visually inspected the spectrogram for multiple reductions in amplitude, which slightly resemble the dark vertical striations seen in vocalic segments. For derhoticized alveolar approximant /ɹ/, we looked for a flat or rising F3 (Stuart-Smith et al., 2014). For fricativized approximant /ɹ/, we searched for noisy waveform, dark, unstructured formants. Lastly, for /ɹ/-deletion, we looked for an absence of a rhotic consonant.

Syllabic C_2s were identified as consonants with high amplitude, intensity, and longer duration than non-syllabic C_2s. /l/-initial consonant sequences displaying /l/-vocalization or /l/-deletion were identified by acoustically weak formants, and an acoustic absence of a lateral consonant, respectively.

Note that syllabic C_2s and cases where C_1 was deleted or weakened (i.e. rhotic deletion, derhoticization, /l/-deletion, /l/-vocalization) were excluded from analyses since I was only concerned with vowel insertion, however, their distribution will be analyzed in Chapter 5.

[61] See Jauriberry et al. (2015) for a discussion of fricativized approximant /ɹ/ in Scottish English.

3.4.5 Data analysis

Data from 18 speakers were analyzed (M=8, F=10, mean age=56, Range=21-93, SD=20.16). Nine participants were excluded from the analysis for not meeting the eligibility criteria (i.e. by living in England for an extended period of time, being born in Scotland but not in the Central Belt region, or being born in England).

Durations were subject to normalization, in which the average segment duration by Subject and by Word were calculated and this value was divided from the corresponding segment duration from each individual token.[62,63]

After the Praat script was run for each speaker, the output was saved as a *.txt* file. Each *.txt* file was imported into R 3.3.1 (R Development Core Team, 2016) and merged into the same data frame to be visually inspected via histograms, bar plots, and violin plots. *Tidyverse* packages (Wickham, 2017) were used to clean and reshape the data, and to create data visualizations. The *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages were used for linear mixed effects modeling.

3.5 Results

In Section 3.5.1 I present the results for the diagnostics used to classify the vowel as phonological. In Section 3.5.2 I present the results for the diagnostics used to classify the vowel as epenthetic. In Section 3.5.3 I present an exploratory finding that suggests that the

[62] For example, for the token *farVm*, where "V" represents an inserted vowel, the average duration of all inserted vowels produced by that speaker for the repeated item *farm* was calculated (e.g. 50 ms), and then the average duration was divided from each utterance of the word *farm* (e.g. utterance 1: 65ms/50ms, utterance 2: 60ms/50ms, etc.).

[63] The only places where normalization was not used are in plots of average duration, so as to not lose the unit of measurement. Where normalized durations were not used, a footnote will be included to indicate this.

vowel is sensitive to minor adjustments in speaking rate, which will be directly investigated in Chapter 4 via a speech rate experiment.

3.5.1 Diagnostic criteria for the vowel being phonological (not excrescent)

3.5.1.1 Durational relations of adjacent gestures

Under a gestural coordination framework, consonants in onset positions bear an in-phase relationship with the following vowel (i.e. C_1-V) (Löfqvist & Gracco, 1999), while consonants in coda positions bear an anti-phase relationship with the preceding vowel (i.e. V-C_2) (Goldstein et al., 2009).[64] A negative correlation between C_1 Duration and Inserted Vowel Duration is expected if the C_1 and V gestures are in-phase with one another (Shaw et al., 2011). A positive correlation between V and C_2 is expected if the gestures have an anti-phase relationship (Goldstein et al., 2009). Both of these correlations would suggest that the vowel is phonological, as it constitutes its own gesture and breaks up the consonants that would otherwise form a coda cluster. Alternatively, a positive correlation between the two consonants (C_1 and C_2) is expected if they are coordinated with an anti-phase relationship, with an excrescent 'vowel' intervening. This is because excrescent vowels are not gestures, but an unintended consequence of the vocal tract being opening when articulating successive consonants that share a small amount of gestural overlap. Although excrescent vowels are present in the acoustic signal, they should not serve to break up the consonants in a coda cluster as do phonological vowels.

[64] "C_1-V" represents an onset+vowel coordination, "V-C_2" represents a vowel+coda coordination, and "C_1-C_2" a consonant+consonant coordination.

Pearson's correlations were conducted for normalized C_1-V, V-C_2, and C_1-C_2 durations in words with an underlying vowel between C_1 and C_2 (e.g. *forum*), an inserted vowel between C_1 and C_2 (e.g. *farVm*), and no vowel between C_1 and C_2 (e.g. *farm*).[65] To ensure that I used the most controlled comparison, I only included monosyllabic items where C_2 would be part of the coda (regardless of whether the intervening vowel was phonological or excrescent). Tokens included consonant sequences that were word-final codas (e.g. *farVm*, *forum*, *farm*) and consonant sequences that were not word-final but a part of a coda cluster (e.g. *world*, *film+s*, *film+ed*). For each subset of the data (Underlying Vowel, Inserted Vowel, No Vowel), outliers greater than 3 standard deviations above the mean were removed. A summary table is provided below (Table 6).

Table 6. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens.

	Correlation between durations		
Vowel Type between CC	**C_1-V**	**V-C_2**	**C_1-C_2**
Underlying Vowel (e.g. *forum*) N=152	-0.20(.0155*)	0.25(.0020*)	0.00(.9885)
Inserted Vowel (e.g. *farVm*) N=646	-0.03(.2658)	0.02(.5376)	-0.14(.0001*)
No Vowel (e.g. *farm*) N=2230	--	--	0.01(0.707)

Note. Table displays Pearson's r, followed by the p-value in parentheses. CC=consonant-consonant sequence. C_1=first consonant, V=inserted vowel, C_2=second consonant. N=number of utterances. Vowels of interest are underlined.

In words with an inserted vowel, C_1 and C_2 were not significantly positively correlated, in line with the prediction for phonological vowels. In words with an underlying vowel, C_1

[65] Where the inserted vowel is underlined.

and C_2 had a non-significant, weak positive correlation. In words without a vowel, C_1 and C_2 had a non-significant positive correlation. Although C_1 and C_2 had an unexpected positive correlation in tokens containing an underlying vowel, this correlation is non-significant. Crucially, the positive correlation for words with an underlying vowel was close to zero and smaller than the positive correlation obtained for vowel-less tokens.

A significantly negative correlation between C_1-V was obtained for tokens with underlying vowels and a non-significant negative correlation between C_1-V was found for tokens with vowel insertion, as expected for gestures with an in-phase relationship. A significantly positive correlation was obtained for V-C_2 in words with an underlying vowel and a non-significant positive correlation was obtained for V-C_2 in words with vowel insertion, as was expected for vowel+consonant sequences (i.e. those with an anti-phase relationship) in which the vowel is phonological.

Taken together, words with vowel insertion pattern with words containing an underlying vowel situated in a similar environment. Most importantly, the significantly negative correlation between C_1 and C_2 in words with vowel insertion excludes the possibility that the vowel is excrescent.

3.5.1.2 Bimodal distribution

A metric that has been used for determining whether inserted vowels are epenthetic or excrescent is visualizing the distribution of vowel duration by plotting counts for raw vowel duration against the counts of zero-duration for tokens *without* an inserted vowel. If the duration of tokens containing a vowel and the zero-duration tokens form a unimodal distribution when plotted together, this would suggest that the vowel is excrescent. That is because the duration of excrescent vowels should be closer to zero than the duration of phonological vowels and their range of possible durations should reflect the constrained set of articulatory settings needed to produce an excrescent vowel (i.e. low-overlapping gestures). Conversely, a bimodal distribution would support a classification of the vowel as phonological, in which a vowel is categorically (optionally) inserted and has a durational target (Bellik, 2018; Bürki et al., 2007). A histogram of raw vowel duration (in milliseconds), separated by Insertion (No Vowel Insertion vs. Vowel Insertion) is provided in Figure 13, which displays a bimodal distribution.[66]

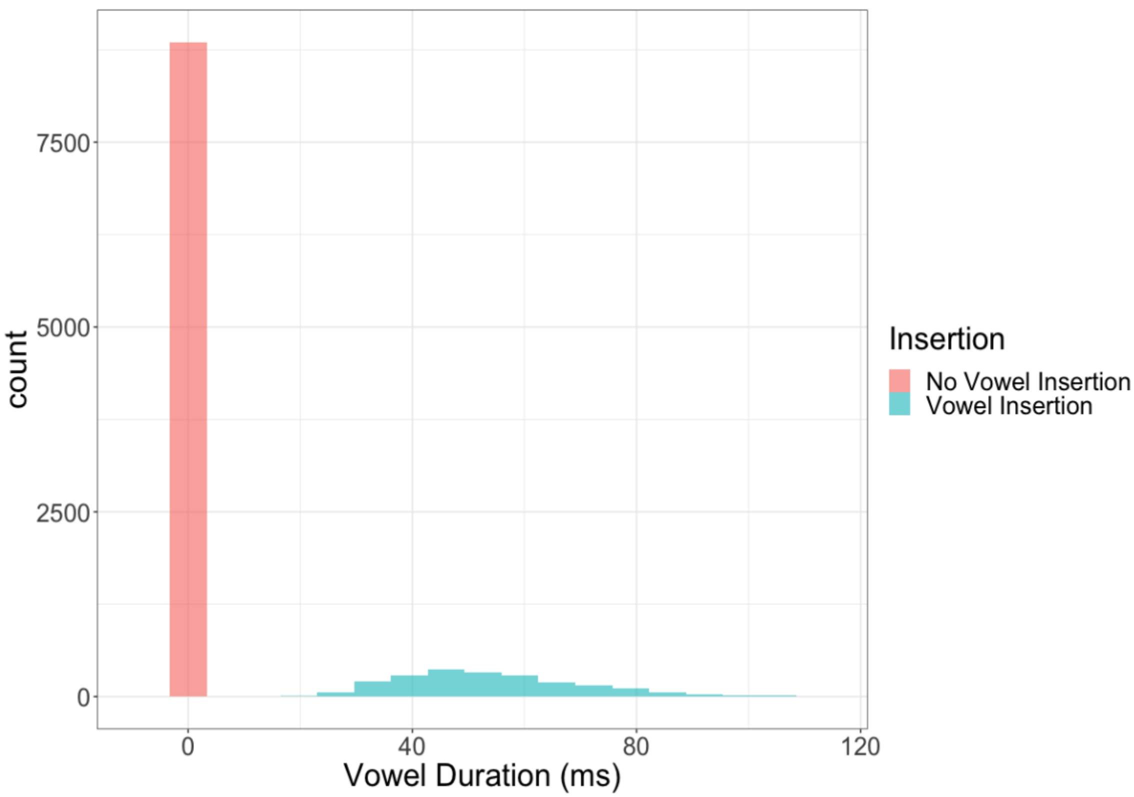

Figure 13. Bimodal distribution of counts for Vowel Duration (ms) by Insertion.

[66] Raw, non-normalized, values were used, following the methodology employed by Bellik (2018).

Figure 13 shows a bimodal distribution—a sharp change in the distribution of vowel-less tokens and the distribution of tokens with vowel insertion, suggesting that the vowel is phonological, not excrescent.

3.5.1.3 Consonant-to-consonant duration

The last diagnostic used to exclude the possibility that the vowel is excrescent was determining whether the vowel adds significant length to the duration of consonant sequence that it intervenes, since empirical evidence suggests that some putative cases of excrescent vowels do not (Ridouane & Fougeron, 2011). Since the duration of excrescent vowels should be highly constrained by the degree of overlap shared by the flanking consonants and do not add a gesture to the sequence, the total consonant-to-consonant duration of words with vowel insertion and those without should be very similar, if the vowel is excrescent. If the vowel is phonological, it should add a significant amount of length to the total consonant-to-consonant duration since the vowel adds a segment to the sequence.

The total consonant-to-consonant duration of words containing an inserted vowel was calculated by adding normalized C_1, V, and C_2 durations together. In words without vowel insertion, normalized C_1 and C_2 were added together. A linear mixed effects model was fit to the data to determine whether the Consonant-to-Consonant Duration was significantly different in tokens containing an inserted vowel from tokens without a vowel between C_1 and C_2 (Table 7). The continuous dependent variable was Consonant-to-Consonant Duration, and the continuous predictor variable was Insertion (No Vowel Insertion vs. Vowel Insertion), with Subject and Word as random intercepts.

Table 7. Linear Mixed Effects model for Consonant-to-Consonant Duration as predicted by Insertion, with Subject and Word as random intercepts.

	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	0.037	2.94	18	0.01	0.990
Vowel Insertion	284.100	1.14	4442	248.21	<.0001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

Overall, the inserted vowel significantly adds length to the Consonant-to-Consonant Duration. This finding is consistent with a classification of the vowel as phonological, not excrescent.

3.5.2 Diagnostic criteria for the vowel being epenthetic (not lexicalized)

For each of the diagnostics used in the previous section, the results indicated that the vowel patterned as a phonological vowel, rather than as an excrescent vowel. Now that I have eliminated the possibility that the vowel is excrescent (i.e. not present in the underlying or surface representation but resulting from gestural alignment), I will use two diagnostics to determine whether the vowel is lexicalized (i.e. in the underlying representation) or epenthetic (i.e. in the surface representation), since both types of vowels are phonological (i.e. originating in the phonology module of the framework).[67]

3.5.2.1 Widespread distribution across lexical items

The first diagnostic used to determine whether the vowel behaved as epenthetic was the distribution of vowel insertion across lexical items. Since prior reports of vowel insertion

[67] As mentioned earlier, the diagnostics in this section (3.5.2) cannot establish whether a vowel is epenthetic or lexicalized when used in isolation. Only *after* independently establishing that the vowel is phonological can one use these diagnostics to determine whether the vowel is epenthetic or lexicalized, specifically.

in Scottish English demonstrate that this process minimally affects a handful of lexical items at each end of the bipolar linguistic continuum, it is possible that the vowel is lexicalized, i.e. cemented, in the underlying representations of this handful of words. To exclude this possibility, the distribution of vowel insertion across all lexical items was measured.

In words that are part of a Standard English lexis, the vowel surfaced in high-frequency words (e.g. *film, arm*), but also in lower-frequency items (e.g. *squirm, helm*). This was also true of words from a Scots lexis. Vowel insertion occurred in high-frequency words like *wirm* 'worm', *ferm* 'farm', *bairn* 'baby', but also in words like *whalm* 'engulf', *folm* 'turn upside down' (glosses from *Dictionary of the Scots Language*, www.dsl.ac.uk).[68] The vowel also surfaced in novel morphological environments. Published accounts of vowel insertion in Scottish English largely report vowels surfacing in monomorphemic words (e.g. *film, farm, world*). In the current study, vowel insertion occurred not only in monomorphemic words, but also in multimorphemic words (e.g. *film+s, farm+er, curl+ing, twirl+ed, arn+a* 'are not', *barm+ie* 'passionate', *gorm+less* 'stupid'), in compound words (e.g. *iver#mair* 'ever more'), place names (e.g. *Thirlstane, Nethermuir, Locharmoss, Dirleton, Talmine*), and surnames (e.g. *Carmicheal, Kilmer, Carloway, Carmyle*) (glosses from *Dictionary of the Scots Language*, www.dsl.ac.uk).[69,70] The widespread presence of vowel insertion in 254 items suggests that this process is epenthetic, not lexicalized.[71]

[68] Note that frequencies are impressionistic, since a frequency corpus could not be accessed.
[69] Morpheme boundaries are indicated by "+" and word boundaries by "#".
[70] Note that the vowel's morpho-phonological distribution will be discussed in detail in Chapters 5 and 6.
[71] Note that this number does not conflate different suffixes added to the same root word. For example, vowel insertion in *film, films,* and *filmed* were counted as 3 separate items.

3.5.2.2 Inserted versus underlying vowel durations

Mean durations (ms) of lexical, underlying vowels situated between liquid+liquid and liquid+nasal sequences (e.g., *Callum, forum)* were compared with mean durations of inserted vowels to determine whether they were significantly different.[72] As discussed in Section 3.2.2.2, the expectation was that a significant difference should emerge between underlying vowels and non-underlying vowels if the vowel is epenthetic, but not if it is lexicalized.

The average raw durations (ms) for each Vowel Type (Inserted vs. Underlying) are visualized in Figure 14, and suggest that inserted vowels have a shorter average duration than underlying vowels.[73]

[72] Vowels of interest are underlined in these examples.
[73] Raw durations were used (rather than normalized durations) to avoid losing generalizability by eliminating the unit of measurement (ms).

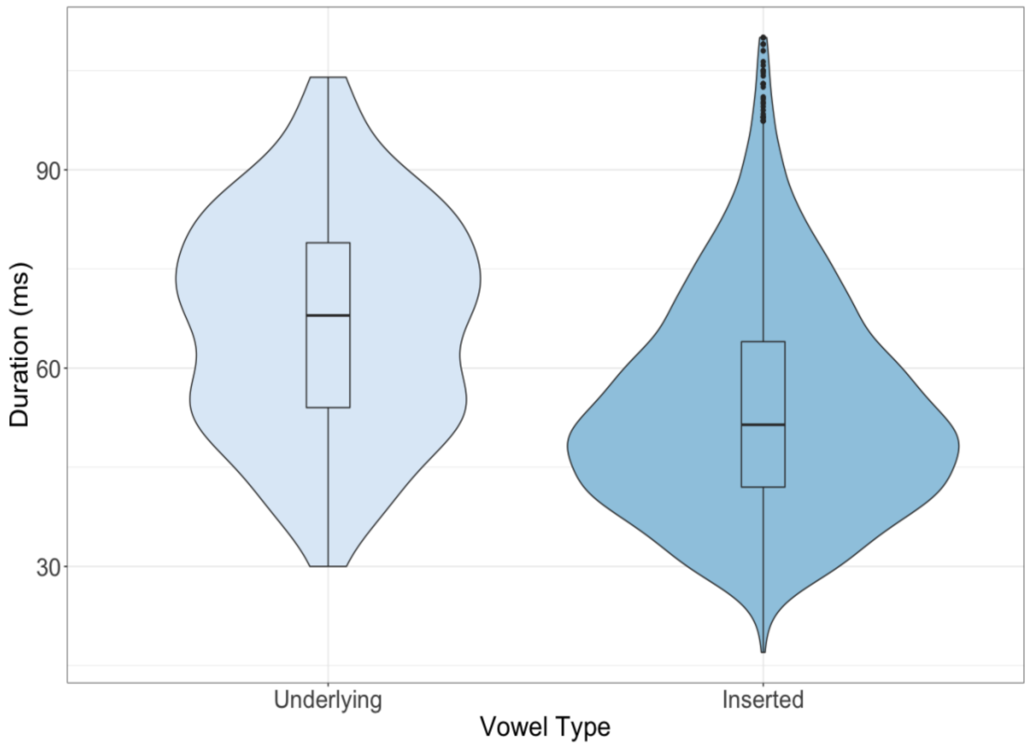

Figure 14. Average duration in milliseconds (ms) for inserted and underlying vowels.

The average value for underlying vowels was 68.16 ms (N=161, SD=19.14, Range=30-150) and 55.04 ms for inserted vowels (N=2151, SD=18.51, Range=17-167)—less than 50% shorter than the average value of underlying vowels. A two-sample Welch test revealed a significant difference between their normalized durations ($t = 8.17$, $p < .0001$***). The significant difference between underlying and inserted vowels suggests that the vowel is not lexicalized.

Taken together with the results of Section 3.5.1, in which I argued that the vowel is phonological, the results of both diagnostics used in this section (3.5.2) suggest that the inserted vowel is epenthetic.

3.5.3 Exploratory result: Inserted vowel duration and preceding vowel duration

If the vowel is phonological, as the previous diagnostics suggest, it should be visible to planning processes which originate at a higher stage of speech planning, like adjustments made for changes in speaking rate. Even at a normal pace, subtle variation in speaking rate can occur between items for the same speaker or across speakers. Phonological segments should adjust with variations in speaking rate to maintain a stable perceptual distance between segments, while excrescent vowels should not (Solé, 1992; Solé & Ohala, 2010). To determine whether the vowel participates in the global timing of gestures, comparing the duration of the inserted vowel with that of a local, underlying gesture can serve as a proxy for estimating the effects of speech rate on the inserted vowel. If the duration of the preceding, underlying vowel and the inserted vowel (e.g. *farVm*) are significantly correlated in a positive direction, I will interpret this as further evidence for the vowel being phonological.[74] If they are *not* positively correlated, I will interpret this as evidence for the vowel being excrescent.

Inserted Vowel Duration (e.g. CVC(V)C) and Preceding Vowel Duration (e.g. CVC(V)C) were measured for each word that contained an inserted vowel.[75] A linear mixed effects model was fit to the data, in which Inserted Vowel Duration was used as the dependent outcome variable and Preceding Vowel Duration as the continuous predictor variable, with Subject and Word as random intercepts. Results indicated that Preceding Vowel Duration significantly predicted Inserted Vowel Duration in tokens with inserted vowels (Table 8).

74 Target vowels are underlined, and the inserted vowel is represented as "V".
75 The inserted vowel is in parentheses to distinguish it from the preceding, underlying vowel which is marked as "V" for "vowel" (as opposed to "C" for "consonant").

Table 8. Linear Mixed Effects Model for Vowel Duration predicted by Preceding Vowel Duration with Subject and Word as random intercepts.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.921	0.02	2149	38.78	0.001***
Preceding Vowel Duration	0.078	0.02	2149	3.30	0.001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

The significant relationship between Preceding Vowel Duration and Inserted Vowel Duration summarized in Table 8 is positive-going, as illustrated in Figure 15.

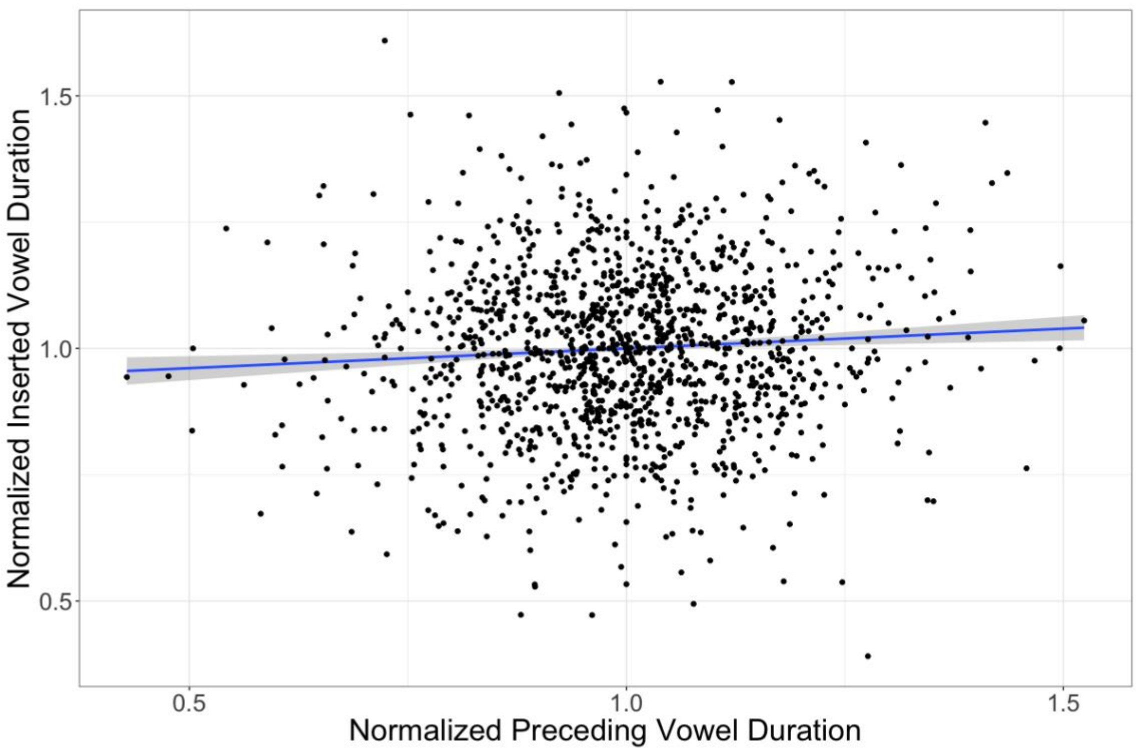

Figure 15. Linear regression for preceding vowel duration and inserted vowel duration. *Note.* The blue line represents the slope of regression, while gray shading represents a 95% confidence interval.

As a baseline comparison, I compared vowel durations in tokens containing an underlying vowel between the consonant sequences of interest (e.g. *forum*) to determine whether Underlying Vowel Duration is significantly predicted by Preceding Vowel

Duration in the same direction as was found for inserted vowels. A linear mixed effects model was fit to a subset of the data containing words with an underlying vowel (e.g. *forum*), with normalized inserted vowel duration as the continuous dependent variable and normalized preceding vowel duration as the continuous predictor variable, with Subject and Word as random intercepts. Results indicated that Preceding Vowel Duration significantly predicted Inserted Vowel Duration (Table 9).

Table 9. Linear Mixed Effects Model for Vowel Duration predicted by Preceding Vowel Duration with Subject and Word as random intercepts.

	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	0.858	0.06	150	13.64	<0.001***
Preceding Vowel Duration	0.142	0.06	150	2.30	<0.05*

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

Results indicate that, in tokens containing an underlying vowel, the vowel's duration was significantly predicted by the duration of the preceding vowel. A significantly positive relationship was obtained for tokens containing an underlying vowel and for tokens containing an inserted vowel, which suggests that the inserted vowel behaves in a manner similar to segments which are underlying. This will be explored further by directly manipulating speech rate in Chapter 4.

3.6 Summary of findings and conclusions

3.6.1 The vowel is phonological (not excrescent)

A comparison of the results against each diagnostic prediction is provided in Table 10.

Table 10. Diagnosis of the vowel as phonological, not excrescent.

Diagnostic	**Prediction for phonological vowels**	**Prediction for excrescent vowels**	**Diagnosis of current experiment**
Durational relations of adjacent gestures	Significantly negative correlation for C_1 and V. Significantly positive correlation for V and C_2. Significantly negative correlation for C_1 and C_2.	No significantly negative correlation for C_1 and V. No significantly positive correlation for V and C_2. Significantly positive correlation for C_1 and C_2.	Phonological
Bimodal distribution of durations	Bimodal.	Unimodal.	Phonological
Consonant-to-consonant duration	Duration of consonant sequence significantly lengthened in words with vowel insertion, compared to those without.	Duration of consonant sequence not significantly lengthened in words with vowel insertion, compared to those without.	Phonological

3.6.2 The vowel is epenthetic (not lexicalized)

A comparison of the results against each diagnostic prediction is provided in Table 11.

Table 11. Diagnosis of the vowel as epenthetic, not lexicalized.

Diagnostic	**Prediction for epenthetic vowels**	**Prediction for lexicalized vowels**	**Diagnosis of current experiment**
Widespread distribution across lexical items	Widespread distribution across lexical items, including novel contexts.	Narrow distribution across lexical items, occurring primarily in a handful of high frequency items.	Epenthetic
Durations of inserted versus underlying vowels	Significantly different.	Not significantly different.	Epenthetic

3.6.3 Conclusions

The current experiment sought to determine whether the inserted vowel in Scottish English liquid+liquid and liquid+nasal consonant sequences should continue to be classified as phonological epenthesis, since alternatives exist. Since prior reports of vowel insertion in Scottish English have relied exclusively on impressionistic transcriptions made by researchers, it was possible that the vowel was actually an articulatory artefact (i.e. excrescent) or absent etymologically but present underlyingly (i.e. lexicalized), as all three types of inserted vowels may superficially sound similar in the acoustics.

The goal of this experiment was to use the relatively understudied Scottish English inserted vowel as a case study to: (1) utilize converging evidence from a host of independent diagnostic criteria to classify the process and (2) demonstrate how a multi-step exclusion process, used on a high volume of acoustic data, can aid in the future experiments which seek to determine where inserted vowels originate in the theoretical representation: in the underlying representation (i.e. lexicalized vowels), surface representation (i.e. epenthetic vowels), or from gestural alignment (i.e. excrescent vowels).

The diagnostics which have not previously been used to assess whether inserted vowels are excrescent or phonological are: (a) durational relations of adjacent gestures modeled by a gestural coupling framework, and (b) the exploratory finding presented in Section 3.5.3—the relationship between the inserted vowel and preceding underlying vowel. Additionally, the widespread distribution of vowel insertion across lexical items and the significant difference between the duration of the inserted vowel and the duration of underlying vowels have not been used before to diagnose inserted vowels as lexicalized or epenthetic. These diagnostics, then, are novel contributions which I encourage researchers use to assess the nature of inserted vowels in future work.

Since epenthetic and lexicalized vowels are both phonological, while excrescent vowels are articulatory, the first step was to determine whether the vowel was phonological. The results of three diagnostics excluded the possibility that the vowel is excrescent. After determining that the vowel was phonological, I assessed the distribution of vowel insertion across lexical items and compared the of duration inserted vowels against underlying vowels to determine whether the vowel was lexicalized, and underlying, or epenthetic, and non-underlying. The results, assessed against two sets of diagnostic criteria, suggest that vowel insertion in Scottish English liquid+liquid and liquid+nasal consonant sequences is a process of phonological epenthesis.

Chapter 4:
The vowel interacts with speech rate.

One of Hall's (2006) criteria for excrescent vowels is that they should shorten or delete in fast speech and lengthen in slow speech. Hall generates this criterion using perceptions and transcriptions from researchers who report that excrescent vowels sound shorter and delete in fast speech.[76,77] Since very few studies have directly examined the effects of speech rate on excrescent vowel duration, I develop a novel speech rate criteria which can be used to distinguish between phonological and excrescent vowels (D'Apolito & Fivela, 2016; Heselwood et al., 2015; Shitaw, 2014) in the current chapter.

In Chapter 3, I used a proxy for speech rate and found that the duration of the inserted vowel in Scottish English varies with subtle adjustments in speech rate that occur even at a normal pace.[78] Comparing the inserted vowel with respect to the preceding vowel in the same lexical item (e.g. *farVm*), I found a positive correlation between their durations across subjects and items.[79] This positive correlation was also found in words that had an underlying vowel situated between liquid+liquid and liquid+nasal sequences (e.g. *forum*), which I used to propose that the vowel behaves like a phonological segment.[80] Crucially, though, if underlying vowels (i.e. lexicalized or lexical) and non-underlying vowels (i.e. epenthetic or excrescent) can each vary with speech rate adjustments (i.e. lengthening in

[76] For example, Hall (2006) cites Harms (1976), who states that, "in fast speech…these vocoids are very short; most frequently they are dropped altogether" (p. 77), and Gafos (2002), who cites a personal conversation with (Heath, 1987), who notes that an epenthetic vowel in Moroccan Arabic is not subject to syncope.

[77] One exception is Jannedy (1994), who provides experimental evidence, but for underlying vowels, not inserted vowels.

[78] In this chapter, I will continue using the term "inserted" until I diagnose the vowel as epenthetic.

[79] The preceding vowel is underlined in this example.

[80] The underlying vowel is underlined in this example.

slow speech and shortening in fast speech), Hall's (2006) criterion is not a diagnostic that can be used to exclude the possibility that a vowel is phonological. In this chapter, I create a testable diagnostic which makes different predictions for the phonetic consequences of phonological vowels and excrescent vowels at different speech rates and use this diagnostic to establish that the inserted vowel in Scottish English is not excrescent.

4.1 The current experiment

The aim of the current experiment is to demonstrate how the inserted vowel in Scottish English is affected by changes in speech rate by asking the following question: Does the duration of the vowel lengthen in slow speech and shorten in fast speech, like phonological segments should? To address this question, I use the behavior of the vowel at different speech rates, along with the diagnostics used in Chapter 3, to establish that the vowel is phonological, not excrescent.

In this chapter, I will argue that Hall's (2006) criterion that excrescent vowels should shorten and delete in fast speech due to increased gestural overlap is not a direct consequence of the underlying theory, and that reliable evidence for increased gestural overlap of any pair of gestures at fast speech is not available from empirical studies which directly investigate overlap and speech rate. I will first review the empirical research that Hall (2006) uses to develop the speech rate criterion, and then discuss how a handful of experiments have found different patterns from what Hall (2006) suggests. I will use the gestural coordination of consonants and vowels under a gestural coupling framework (Goldstein et al., 2009; Nam & Saltzman, 2003; Saltzman & Byrd, 2000) to discuss why it is important to use empirical patterns for specific phonological environments to inform expectations for how gestural overlap is modeled at different speech rates (rather than assuming the same pattern of overlap for complex codas, complex onsets, and consonants which do not form a complex cluster).[81] I will then use empirical patterns relevant to *coda* consonants to diagnose the inserted vowel which primarily surfaces in Scottish English coda clusters as phonological, rather than excrescent.

After determining that the vowel behaves as a phonological vowel across different speech rates, I will then assess the behavior of the vowel against the criteria developed in Chapter 3. I will show that the diagnosis made in Chapter 3 for the vowel in Scottish English being phonological is supported in the current experiment, and demonstrate the reliability of the diagnostics employed in Chapter 3 (i.e. durational relations of adjacent gestures, bimodal distribution of duration, and consonant-to-consonant duration).[82] I will conclude the chapter by discussing how, taken together, the results of Chapters 3 and 4 suggest that the vowel in Scottish English is epenthetic.

This chapter makes the following contributions:

(15) It provides a testable diagnostic for speech rate which can be used for inserted vowels which surface in coda clusters.

(16) It provides direct evidence for how epenthetic vowel duration in coda clusters is affected by speech rate.[83]

In Section 4.2, I will review Hall's (2006) criterion for excrescent vowels. In Section 4.3.1, I will review how consonants are coordinated under a gestural coupling framework. In Section 4.3.2, I will overview how consonants separated by a morpheme

[81] Where "#" represents a word boundary.
[82] For limitations of space, I refer the reader to Chapter 3 for a full description of these diagnostics. The predictions for each diagnostic will be discussed in each result section in the current chapter.
[83] Following Shitaw (2014).

boundary are coordinated, since Hall (2006) uses overlap between consonants situated across a morpheme boundary to generalize to how all gesture should overlap. In Section 4.3.3, I provide empirical evidence for how the degree of overlap shared by anti-phase gestures is affected by speech rate, as this is the relevant gestural coupling for coda consonants. In Section 4.4, I will use theoretical modeling and empirical patterns to develop a diagnostic for how excrescent vowels should behave compared to phonological vowels in coda position.

4.2 Hall's (2006) criterion for excrescent vowels at different speech rates

One of the criteria that Hall (2006) provides for a vowel being excrescent is that the vowel lengthens in slow speech and shortens or deletes in fast speech, which Hall claims is a "direct prediction of the gestural analysis" (Hall, 2003, p. 11). Hall (2003, 2006) uses patterns identified in cross-linguistic studies to develop these criteria, however, the evidence that Hall (2006) provides almost exclusively comes from impressionistic transcriptions. Hall (2006) argues that this pattern is due to the flanking consonant gestures overlapping more in fast speech and eliminating the excrescent vowel. However, Hall's (2006) criterion for how excrescent vowels are affected by speech rate is problematic for a few reasons.

The primary reason why Hall's (2006) criterion is problematic is that the shortening and deletion of vowels at fast speech cannot serve to exclude the possibility that the vowel is phonological. This is because plenty of evidence exists for underlying vowels deleting at fast speech (in English #CəC, e.g. *potato* /pə̥teɪɾoʊ/ [pteɪɾoʊ], Davidson, 2006; in

German /ka̲nən/ [knən], Jannedy, 1994; in Mandarin, e.g. /ʂə́mə̲/ [ʂə́m] 'what', Weinberger, 1996).[84]

Another reason is that transcribers may fail to perceive vowels that are present in the acoustic signal or perceive illusory vowels that are not (Davidson & Shaw, 2012; Dupoux et al., 1999). Relying on evidence that excrescent vowels shorten and/or delete at fast speech which does not come from a controlled experiment, then, may be misleading.

Although Hall (2006) does not cite cross-linguistic *acoustic* evidence for how excrescent vowels should behave at different speech rates, Hall cites articulatory and acoustic evidence for increased gestural overlap at fast speech.[85] However, this evidence comes from consonants which straddle a morpheme or word boundary (Byrd & Tan, 1996; Davidson, 2003; Munhall & Löfqvist, 1992; Zsiga, 1994). It is important to note, though, that increased gestural overlap of the consonants in C#C sequences is not the only consequence that has been observed in research on rate-induced gestural overlap.[86] For example, Luo (2017) found that gestural overlap in C#C sequences did *not* significantly increase in fast speech. If there is no reliable pattern of overlap in studies which directly investigate the coordination of C#C sequences, it is also likely that empirical patterns for overlap in C#C sequences will not generalize to onset (i.e. #CC) and coda (i.e. CC#) clusters.

[84] Underlying vowels subject to deletion at fast speech are underlined. “#” indicates a word boundary.
[85] In consonant sequences that excrescent vowels do not intrude upon.
[86] Where “#” represents a word boundary.

4.3 Theoretical modeling and empirical patterns for gestural overlap

In this section, I will summarize how onset and coda clusters are coordinated in a gestural coupling framework (Section 4.3.1) and how consonants separated by a word or morpheme boundary are coordinated using the same framework (Section 4.3.2). In Section 4.3.3 I will argue that generalizations for consonantal overlap should not be made for consonants in varying phonological environments, i.e. onset clusters, coda clusters, and consonants separated by a word (i.e. C#C) or morpheme boundary (i.e. C+C).[87] I will then discuss why I will only use modeling that is directly pertinent to coda clusters to generate a diagnostic for excrescent vowels, since the focus of this dissertation is vowel insertion which primarily occurs between consonants which would otherwise form a coda cluster.

4.3.1 Organization of gestures in a gestural coupling framework

Under a gestural coupling framework, the timing of individual gestures and coupled gestures is modeled using the coupling of *individual* oscillators, which become *relative* oscillator phases (Goldstein et al., 2009; Nam & Saltzman, 2003; Saltzman & Byrd, 2000). Gestures can have one of two target specifications: in-phase or anti-phase. An in-phase target specification is one in which both individual gestures occur simultaneously, while an anti-phase target specification is one in which individual gestures are triggered in a stepwise fashion. Gestural coupling relationships for complex consonant clusters are illustrated in Figure 16.

[87] Where "#" indicates a word boundary and "+" indicates a morpheme boundary.

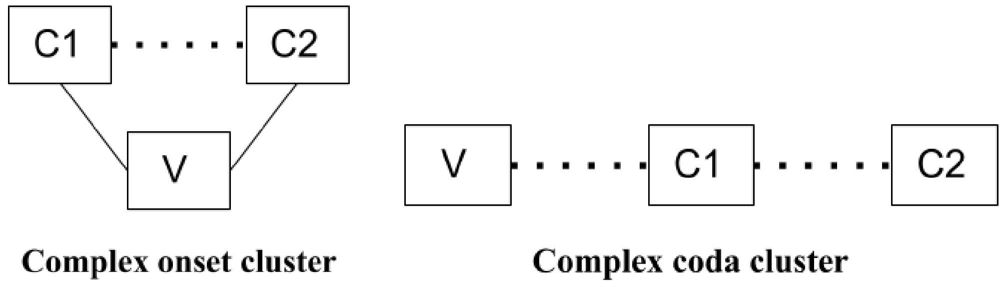

Figure 16. Gestural coupling of complex consonant clusters. Solid lines represent in-phase relationships, while dotted lines indicate anti-phase relationships. Recreated from Marin & Pouplier (2010, p. 382).
Note. C_1=first consonant in the cluster, C_2=second consonant in the cluster, V=vowel.

In Figure 16, the consonants in complex onset clusters and complex coda clusters have an anti-phase relationship with one another (i.e. are timed sequentially), as indicated by the dotted line between the first (C_1) and second consonant (C_2) in each cluster. However, the consonants in complex onset clusters *also* have an in-phase relationship with the following vowel (i.e. C_1 and V are initiated at the same time, and C_2 and V are timed together), while the consonants in complex coda clusters do not. As discussed in Chapter 3, since in-phase gestures are timed simultaneously, the degree to which they overlap may vary, while the degree to which gestures with an anti-phase relationship overlap should not vary. Since the degree to which consonant gestures in complex onset clusters exhibit overlap may be affected by their in-phase relationship with the following vowel, their anti-phase relationship is different from the consonant gestures in complex coda clusters, which *only* have an anti-phase relationship (i.e. V-C_1 and C_1-C_2). For this reason, I refer to the coordination of the gestures in coda position (i.e. V-C_1 and C_1-C_2) as having a "pure" anti-phase relationship.

4.3.2 Gestural timing of consonants across morpheme and word boundaries

As mentioned in Section 4.2, Hall (2006) argued that excrescent vowels should shorten and delete in fast speech using the gestural coordination of consonants separated by a word boundary, yet Luo (2017) found that, across a word boundary, gestural overlap did not significantly increase with increased speaking rate, rather, gestures overlapped to the same degree in fast speech as in slow speech. Under a gestural coupling framework, the consonants which straddle a word boundary do not have a timing relationship (i.e. V_1-C_1#C_2-V_2), like consonants which are not separated by a word boundary (Nam & Saltzman, 2003). For this reason, it is not clear why the consonants separated by a word boundary should exhibit consistently greater overlap at fast speech, as Hall (2006) claims.

In a study of inter-gestural timing within and across morphemes and lexical items, Cho (2001) found that Korean consonant sequences separated by a morpheme boundary in non-lexicalized compounds (e.g. /hak#pi/) exhibited more variability in their degree of overlap at fast speech than did consonants which were not separated by a morpheme boundary in lexicalized compounds (e.g. /hakpi/).[88,89] Nam & Saltzman (2003) use the patterns found by Cho (2001) as an example of how gestures intervened by a morpheme boundary (i.e. V_1-C_1#C_2-V_2) have a different coupling than tautosyllabic consonants (i.e. V_1-C_1-C_2-V_2).

Although consonants that form a cluster (i.e. onsets and codas) are organized differently than those that do not (i.e. C#C or C+C) (Nam & Saltzman, 2003), the consonants in onsets and the consonants in codas *also* have a different coordination from

[88] Consonants measured for gestural overlap are underlined.
[89] Lexicalized compounds are those in which the whole compound word is stored in the underlying representation as a single unit, rather than by its constituent morphemes /hak/ and /pi/.

one another. Although C_1 and C_2 have an anti-phase relationship in complex onsets *and* in complex codas, the consonants in complex onsets *also* have an in-phase relationship with the vowel, while the consonants in coda clusters do not. In the next section I will argue that, in order to generate a diagnostic for the behavior of excrescent vowels in coda clusters across speech rates, evidence of gestural coordination at varying speech rates needs to come from gestures which have a "pure" anti-phase relationship (i.e. V-C_1 or C_1-C_2#).

4.3.3 Gestural timing of anti-phase gestures

In order to determine how phonological and excrescent vowels should behave at different speech rates in coda clusters, as is the focus of the current experiment, evidence of gestural coordination at varying speech rates should come from gestures which have a pure anti-phase relationship. Studies of vowel-consonant (i.e. nucleus+coda) coordination at fast speech provide such an opportunity, since the vowel and the following coda consonant have a pure anti-phase relationship.[90] In the next three subsections, I will overview three studies which investigate gestural overlap at varying speech rates in sequences with an anti-phase relationship (namely, vowel-consonant sequences) and do *not* find increased gestural overlap at fast speech. I will argue that Hall's (2006) claim that excrescent vowels

[90] Note that Byrd (1996) directly investigated gestural overlap between consonants in coda clusters at increased speech rate, finding less inter-gestural stability than onset clusters in fast speech. However, the lexical items used were *mask, bagged,* and *backs*, which Byrd represents as: [sk#], [gd#], and [ks#], with "#" used to indicate a word boundary. Since Byrd's study is primarily concerned with the difference between tautosyllabic consonants (i.e. codas and onsets) versus heterosyllabic consonants (i.e. consonants across a word boundary), and not whether consonants belong to the same morpheme, it is understandable that Byrd used *bagged* and *backs* to represent tautosyllabic codas. Since these are multimorphemic words (i.e. /back+s/ and /bag+d/, with "+" used to indicate a *morpheme* boundary), empirical data from Byrd (1996) will not be used to inform the current study. This is because I am concerned with vowel insertion between what I call "pure" anti-phase gestures that are modeled by a gestural coupling framework as having a direct link between them (i.e. V_1-C_1-C_2-V_2), unlike gestures separated by a morpheme boundary (i.e. V_1-C_1#C_2-V_2). For this reason, I rely on empirical studies of V-C (i.e. anti-phase gestures) overlap to make generalizations about how coda consonants (i.e. anti-phase gestures) are coordinated at fast speech.

should shorten and delete in fast speech is not generalizable to all phonological environments in which excrescent vowels may surface, and, more importantly, that no pattern can be predicted using a gestural coupling framework. I will argue that a criterion which can be used to distinguish excrescent from phonological vowels in coda clusters is that phonological vowels should lengthen in slow speech and shorten in fast speech, while excrescent vowels should not.

4.3.3.1 de Jong (2001)

In an experimental study on rate-induced resyllabification, de Jong (2001) found that the coda consonant in vowel-consonant (i.e. anti-phase) sequences like *eep* (i.e. [ip]) became more onset-like (e.g. *pea* [pi]) as speaking rated increased—a result which others have found (Stetson, 1951; Tuller & Kelso, 1991, 1990). Measuring the duration of consonant+vowel and vowel+consonant sequences, de Jong (2001) found that CV syllables shortened more in fast speech than VC syllables. de Jong (2001) also measured the duration of the coda consonant and the vowel of the following syllable (i.e. V<u>C.V</u>), and found that C.V was still shorter than VC. Authors interpreted this as a possible resyllabification of the coda consonant to the onset of the following syllable. In this case, the distance between vowel and consonant gestures in a VC sequence (i.e. those with an anti-phase relationship) increased *more* in fast speech, subsequently shortening the distance between the coda consonant and the following vowel of the next syllable (i.e. V<u>C.V</u>). This result suggests that at fast speech, gestures with an anti-phase relationship may separate more.[91]

[91] The relevant pair of gestures is underlined.

4.3.3.2 Durvasula & Huang (2017)

Durvasula and Huang (2017) investigated the syllable affiliation of word-medial nasal consonants in English which are often regarded as ambisyllabic, i.e. affiliated with both the preceding syllable and following syllable (e.g. *Danny* [dæni]). Since prior research found that coda nasals affect the percentage of nasalization found on the preceding vowel (i.e. anti-phase vowel+nasal sequences) in English across varying speech rates (Solé, 1992, 1995, 2007), Durvasula and Huang (2017) measured the percentage of nasalization on the vowel preceding the word-medial nasal consonant (i.e. vowel+nasal) across slow, normal, and fast speech. Results indicated that, as speech rate increased, the percentage of nasalization on the preceding vowel decreased. This suggests that the degree of coarticulation (i.e. overlap) between vowel-consonant (i.e. anti-phase) sequences may *decrease* with increases in speech rate.

4.3.3.3 Solé (1992)

Speech rate has been used to distinguish whether a process of vowel nasalization in vowel+nasal (i.e. anti-phase) sequences in Spanish arises from physiological, mechanistic constraints or from phonological conditioning (Solé, 1992, 1995, 2007). Solé proposed that mechanical, co-articulatory effects should not adjust to speech rate, since low-level articulatory phenomena “do not participate in the higher-level reorganization of timing and durational factors” (Solé, 2007, p. 306). Comparing the nasalized portion of the vowel across speech rates, Solé (1992) hypothesized that the duration of vowel nasalization (contrasted with the duration of the oral portion) should vary with speech rate if nasalization was a deliberate phonological process. If the duration of nasalization was similar across different speech rates, then it could be classified as an articulatory effect. Solé (1992) found that the proportion of nasalization (measured with respect to the oral portion) of the vowel remained stable across speech rates, and was subsequently able to conclude that the process of vowel nasalization in Spanish was a low-level articulatory effect. Unlike the results found by de Jong (2001) and Durvasula and Huang (2017), that gestures overlapped less in fast speech, the results of Solé's (1992) study provide evidence for co-articulation (i.e. overlap) between anti-phase gestures—in this case, vowel+nasal (i.e. anti-phase) sequences—remaining *constant* across speech rates.

4.4 Developing a speech rate diagnostic for phonological and excrescent vowels

The empirical results of the three experiments reviewed in the previous section suggest that gestures in an anti-phase relationship (i.e. vowel-consonant or consonant-consonant) may exhibit *decreased* overlap at fast speech or exhibit a *constant* degree of overlap across fast, normal, and slow speech rates. Based on the above observations, excrescent vowels in coda clusters should either lengthen in fast speech and shorten in slow speech, or remain the same length across fast, normal, and slow speech, contra Hall (2006), since the flanking consonant gestures (i.e. those with an anti-phase relationship) should *either* overlap less in fast speech and more in slow speech *or* maintain a constant degree of overlap across all speaking rates.

Theoretically, if changes in the duration of excrescent vowels arise from variable gestural overlap of the flanking consonants, a gestural coupling model cannot predict that the vowel will shorten and/or delete in fast speech and lengthen in slow speech—at least in coda clusters—as Hall (2006) argues. Only the results from Solé (1992) can be predicted from the gestural coupling framework. That is because the degree to which gestures in anti-phase relationships shorten or lengthen should not affect the degree of overlap shared by the gestures (Shaw & Kawahara, 2018) (Figure 17).

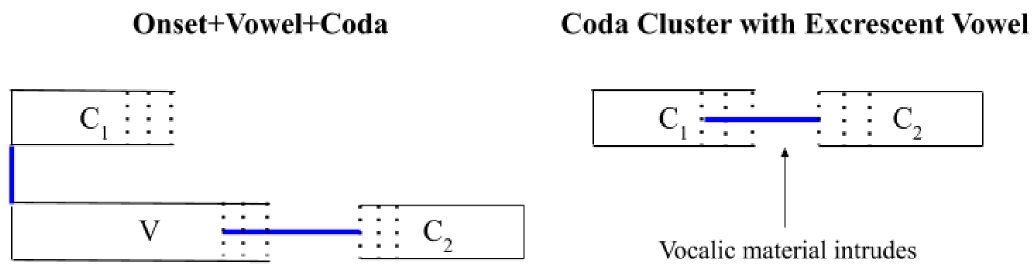

Figure 17. Acoustic consequences of gestural coordination for onset+vowel+coda sequences and for coda clusters with excrescent vowels intruding between the consonant gestures. Based on Shaw & Kawahara (2018, p. 108)
Note. C_1=first consonant, V=vowel, C_2=second vowel. Dotted lines indicate different relative gesture lengths. A vertical blue line represents the simultaneous timing of gestures. Horizontal blue lines represent the sequential timing between gestures.

In Figure 17, the degree to which the first consonant shortens affects the degree to which the following vowel is overlapped by the flanking consonants in gestures which have an in-phase coordination (C_1 and V in left panel), while the degree to which the first consonant of an anti-phase gestural coordination shortens should not affect the degree to which the adjacent gestures overlap with one another (V and C_2 in left panel; C_1 and C_2 in right panel).

The gestural coupling framework predicts that gestures with an anti-phase relationship should have a constant degree of overlap across different speech rates. Since excrescent vowels arise only in clusters which overlap to a low degree, if they are present in normal speech, they should be present in fast and slow speech also, and their duration should remain constant. On the other hand, *phonological* segments should adjust with changes in speech rate—lengthening in slow speech and shortening in fast speech—as they have their own gesture associated with them and should undergo changes in their relative duration to ensure that the absolute duration of the entire lexical item is produced in a shorter or faster amount of time. Additionally, in onset+vowel+coda sequences (i.e. consonant sequences produced with a phonological vowel), the degree to which V and C_2 lengthen or shorten should not affect the degree of overlap shared between them. However, the degree to which C_1 and V overlap can vary, since shortening C_1 can expose more of the intervening vowel and lengthening C_1 can overlap the vowel more, while the vowel can shorten and lengthen, also (as indicated by the dotted lines in Figure 17). This can result in phonological vowel duration being affected by changes in speaking rate.

Taken together, the way that gestures are coordinated in a gestural coupling framework, along with the empirical findings which do not provide support for Hall's (2006) claim, make a testable diagnostic for phonological and excrescent vowels which surface between consonants which would otherwise form a coda cluster. Phonological vowels—underlying or epenthetic—should lengthen in slow speech and shorten in fast speech, while excrescent vowels that surface between coda consonants should not. In the current experiment, then, I will use this as a criterion to diagnose the vowel in Scottish English as phonological, not excrescent.

I will overview the methodology used in the current experiment in Section 4.5, the experimental results in Section 4.6, and my discussion and conclusions in Section 4.7.

4.5 Methods

4.5.1 Participants

29 Scottish English native speakers (Range=46-82, Mean age=63, SD=10.46) born and residing in the Central Belt region of Scotland at the time of testing were recruited.[92] Only

[92] See APPENDIX E for the recruitment form used.

participants over the age of 45 who self-identified as male were recruited. The goal of selecting an older male demographic was to increase the chance of having speakers who retain a Scots lexis (Scobbie et al., 2006), since this process has been reported to occur at both ends of the linguistic continuum.[93]

Testing sites included The University of Edinburgh in Edinburgh, UK, and the Mitchell Library in Glasgow, UK. Participants were recruited via Scots language Facebook groups, physical flyers posted at the Mitchell Library, email advertisements to faculty and staff mailing lists at The University of Edinburgh, and the Philosophy, Psychology, and Language Sciences Volunteer Panel at The University of Edinburgh.[94] Facebook groups included the *Scots Language Forum* (~5,500 members; public group) and the *Scots Language Society* (387 members; public group). Eligibility requirements included: being a native speaker of either Broad Scots or Scottish Standard English; having normal or normal-to-corrected vision and hearing; having no speech disorders; being over 45 years old; self-identifying as a male; being born in the Central Belt region of Scotland, which contains the major cities Edinburgh and Glasgow; and at the time of testing (June 2019), residing in the Central Belt region of Scotland. Speakers were paid £15 for their participation.[95]

4.5.2 Stimuli

Stimuli included 20 unique sentence frames, each 8-10 words in length, repeated 30 times. The presentation of stimuli was sectioned into 3 speech rate blocks: *fast*, *normal*, and *slow*.

[93] See Chapter 2 for a discussion of the usage of Scots and the presence of vowel insertion in both SSE and Scots.

[94] I'd like to express my gratitude for access to this resource.

[95] I'd like to express my gratitude to the Michigan State University Phonetics and Phonology Group for providing this funding.

Each of the 20 sentences were repeated 10 times within each of the three blocks of the experiment, for a total of 600 target utterances (20 sentences x 10 repetitions x 3 speech rates). Target words consisted of monosyllabic, monomorphemic items containing liquid+liquid and liquid+nasal tautosyllabic consonant clusters (i.e. /ɹl/, /ɹm/, /ɹn/, /lm/) occupying the coda position. Consonant clusters occupying a tautosyllabic coda were selected since this is the environment which conditioned vowel insertion the most in the prior study (Chapter 3).[96] In all but one word (i.e. *world*), the cluster was situated word-finally (i.e. *film*, *farm*). In each sentence frame, the target word was located in the second half of the sentence, in varying positions (e.g. *I hope I'll get better marks next* <u>*term*</u>*; On the horizon, a big* <u>*storm*</u> *approached the beach*) (APPENDIX F).[97] This was done to ensure that participants had time to adjust to the appropriate speaking rate (i.e. *fast*, *normal*, *slow*). In all sentence positions, target words were nouns.

4.5.3 Procedure

Participants were seated in a quiet booth with the experimenter in the adjacent room.[98] At the start of the session, participants completed an informed consent. Participants were instructed to read aloud from a series of sentences presented on a 2010 Macbook Pro laptop as naturally as possible while being recorded with a Samson Go Mic Portable USB Condenser Microphone. Recordings were made using a Cardoid 10db setting. Sound files

[96] Since the current experiment is concerned with how the duration of inserted vowels produced at a normal speaking rate differ from the duration of inserted vowels produced in fast and slow speech, these items were selected to increase the chances that vowel insertion would be present at a normal speech rate.
[97] Target words are underlined.
[98] I'd like to express my gratitude to the School of Informatics at The University of Edinburgh and the Mitchell Library in Glasgow for providing me with access to two soundproof booths and an individual study carrel, respectively. Access was made possible via a Postgraduate Researcher opportunity in the School of Philosophy, Psychology, and Language Sciences at The University of Edinburgh and a library membership provided by the Mitchell Library.

were generated using Audacity 2.3.2 (Audacity Team, 2019) with a sampling rate of 44100 Hz (16-bit resolution; mono channel).

Sentences were presented visually via PsychoPy version 3.1.0 (Peirce et al., 2019). Each sentence frame was displayed at the center of the screen, for a fixed rate of time, and replaced by a fixation cross before the presentation of the next stimulus, to ensure that participants paused between each stimulus. The duration that each sentence was displayed on the screen varied by each block of the experiment—*slow, normal,* and *fast*. Sentences were displayed for 5.5 seconds in the *slow* speech rate block, 4 seconds in the *normal* speech rate block, and 2.5 seconds in the *fast* speech rate block.

Instructions were displayed visually for ten seconds at the start of the experiment, detailing that each sentence would appear for a fixed amount of time and be replaced by a fixation cross for 2.5 seconds. Participants were instructed to "speak as you would with your family or close friends", but that there would be three speech rate blocks, in which speakers would need to alter their speaking rate. The instructions detailed that each block of the experiment would include a set of instructions and handful of practice items, and that, once participants had completed each speech rate block, a break would be provided. Participants were instructed to try to match their speaking rate with the duration that each sentence was displayed on the screen, but that a visual reminder would appear under each stimulus indicating which speech rate to use. For example, in the *slow* speech rate block, sentences would be displayed for 5.5 seconds, and the word "SLOW" would appear below the sentence the whole time.

At the start of each block of the experiment, participants were told which speech rate to speak at and instructed to read aloud from three practice sentences to familiarize them with the rate at which each sentence would be displayed on the screen. Sentences were displayed in a white font against a gray background, while the word "SLOW", "FAST", or "NORMAL" would appear in capital letters in yellow font below each sentence, depending on the speech rate block. A Latin-Squares design was used to counterbalance the order in which the three blocks of the experiments were displayed. The order of presentation for each sentence was randomized within each block of the experiment, for each participant. In between blocks, participants were provided an automated break of 12 seconds and encouraged to drink water. At the end of the experiment, participants completed a brief demographic questionnaire (APPENDIX G).[99] Sessions lasted about 45 minutes in total.

4.5.4 Acoustic measurements

For each sound file, maximum formant values for males were set to 5000 Hz. Segmentation was done manually by myself and one paid assistant using the speech analysis software Praat (Boersma, 2001). Each segment was identified by visual inspection of the spectrogram and waveform.[100] The phonetic variable measured in the current experiment was duration (in milliseconds). After all of the *.wav* files were segmented, a Praat script was written which extracted the start time and end time of the segment boundaries. Segment duration was calculated by subtracting the start time from the end time. In the following subsections, I will detail how the segmentation was performed.

[99] I refer the reader to APPENDIX H for a summary of participant demographics.

[100] Forced alignment (i.e. the automatic detection of segment boundaries) was not used because the current experiment used sentence-level speech and only one target word per sentence needed to be segmented. Additionally, I determined that forced alignment did not save time in Chapter 3 since every word of each file (over 1,000 words in a file ~45 minutes in length) was hand-corrected after forced alignment.

4.5.4.1 Vowel measurements

The presence of a vowel between consonants which would otherwise form coda clusters in monosyllabic, monomorphemic words with a CVC(V)C (e.g. *far(V)m*) or CVC(V)CC (e.g. *wor(V)ld*) structure was determined by manual inspection of the spectrogram and waveform.[101] The following visual cues were used to indicate the presence of a vowel: U-shaped curvature in the waveform, clear dark formants, and vertical striations in the spectrogram. An example of a token with vowel insertion and a token without vowel insertion is provided in Figure 18 and in Figure 19, respectively.

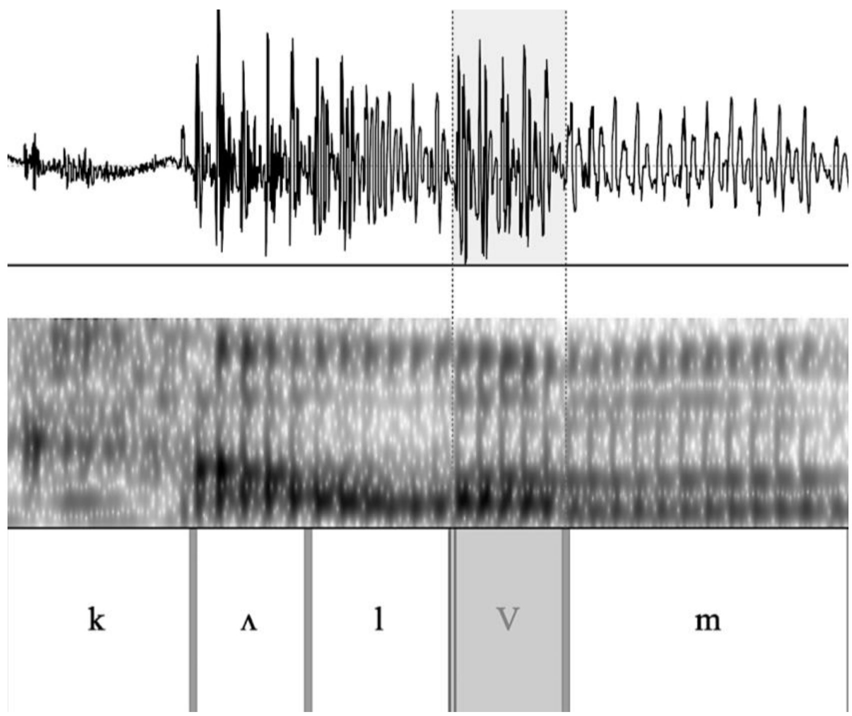

Figure 18. Example of [l] + inserted vowel (Subject 3, *culm*, normal speaking rate). *Note.* Dashed lines mark the offset of C_1 and onset of C_2, with vocalic material intervening (“V”).

[101] In each example provided here, “V” represents an inserted vowel and parentheses indicate optionality.

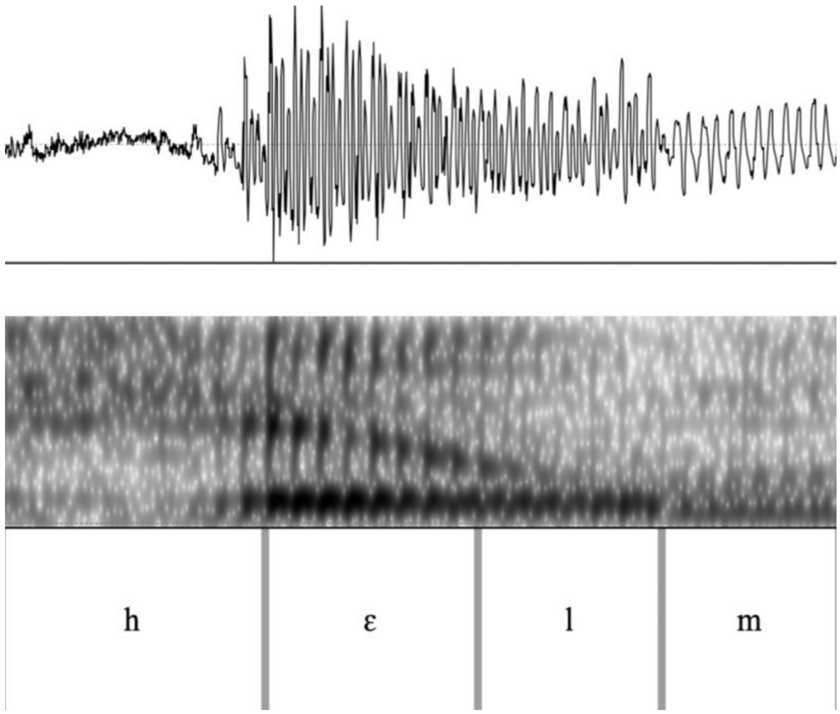

Figure 19. Example of [l] + no vowel insertion (Subject 3, *helm*, normal speaking rate).

Boundaries were marked at the start and end of the vowel, at the point at which there was a clear change in formant structure in the spectrogram and a change in amplitude in the waveform from the preceding and following segments.

4.5.4.2 Consonant measurements

We coded for five possible rhotic realizations: alveolar approximant [ɹ], trilled [r], tapped [ɾ], derhoticized (i.e. acoustically-weakened) alveolar approximant /ɹ/, and fricativized approximant /ɹ/, along with /ɹ/-deletion.[102,103] Acoustic quality of segments was determined using the first (F1) and second formants (F2), which were extracted from the midpoint of each segment, along with formant trajectories in the spectrogram.

102 "We" refers to myself and one trained annotator.
103 See Jauriberry et al. (2015) for a description of fricativized approximant /ɹ/.

A key visual cue used to identify rhotics, especially approximant [ɹ], was the lowering of the third formant (F3). As for tapped [ɾ], we looked for brief reduction in energy for F1 and F2 (Stuart-Smith & Lawson, 2017), as well as a reduction in amplitude and pitch across all formants, slight frication after the tapped portion, and a duration of approximately 25 milliseconds. For trilled [r], we visually inspected the spectrogram for multiple reductions in amplitude, which slightly resemble the dark vertical striations seen in vocalic segments. For derhoticized alveolar approximant /ɹ/, we looked for a flat or rising F3 (Stuart-Smith et al., 2014). For fricativized approximant /ɹ/, we searched for noisy waveform, dark, unstructured formants. Lastly, for /ɹ/-deletion, we looked for an absence of a rhotic consonant.

Syllabic C_2s were identified as consonants with high amplitude, intensity, and longer duration than non-syllabic C_2s. /l/-initial consonant sequences displaying /l/-vocalization or /l/-deletion were identified by acoustically weak formants, and an acoustic absence of a lateral consonant, respectively.

Note that syllabic C_2s and cases where C_1 was deleted or weakened (i.e. rhotic deletion, derhoticization, /l/-deletion, /l/-vocalization) were excluded from all analyses since I was only concerned with vowel insertion. I will briefly discuss their distribution in Chapters 5 and 6.[104]

4.5.5 Data analysis

Data from 14 participants were analyzed (M=14, Range=46-82, Mean age=65, SD=11.33). Fifteen participants were excluded from the analysis for not meeting the eligibility

104 This will be to discuss how they arise in tokens without vowel insertion, and how phonological vowel epenthesis typically occurs in places which are avoided by way of other phonological processes (Hall, 2006).

requirements (i.e. not being native Scottish English speakers but speakers born and raised in England), or for not producing an inserted vowel in any speech rate. Since the primary aim of the current study was to determine whether differences in duration existed across each level of speech rate, it was imperative that participants produced inserted vowels in order to do within-subjects comparisons. Durations were subject to normalization, in which the average segment duration by Subject and by Word were calculated in the *normal* speech rate and this value was divided from the duration of each vowel uttered across all tokens and all three levels of speech rate.[105]

After the Praat script was run for each speaker, the output was saved as a *.txt* file. Each *.txt* file was imported into R 3.6.2 (R Core Team, 2019) and merged into the same data frame to be visually inspected via histograms, bar plots, and violin plots. The *dplyr* 0.8.4 (Wickham et al., 2020), *tidyr* 1.0.2 (Wickham & Henry, 2020), and *ggplot2* 3.2.1 (Wickham, 2016) packages were used to clean and reshape the data, and to create data visualizations. The *lme4* 1.1-21 (Bates et al., 2015) and *lmerTest* 3.1-1 (Kuznetsova et al., 2017), and *nnet* 7.3-12 (Ripley et al., 2016) packages were used for linear mixed effects modeling.

4.6 Results

In Section 4.6.1 I present the results of the speech rate diagnostic. In Section 4.6.2 I present the replicated results of the three diagnostics used in Chapter 3 used to establish that the vowel is phonological.

[105] The only places where normalization was not used are in plots of vowel duration, so as to not lose the unit of measurement. Where normalized durations were not used, a footnote will be included to note this.

4.6.1 Vowel duration and speech rate

The aim of this experiment was to determine whether the duration of the vowel shortens in fast speech and lengthens in slow speech, as phonological vowels should. Across the 14 subjects, the raw average duration of inserted vowels at each speech rate is presented in Figure 20.[106]

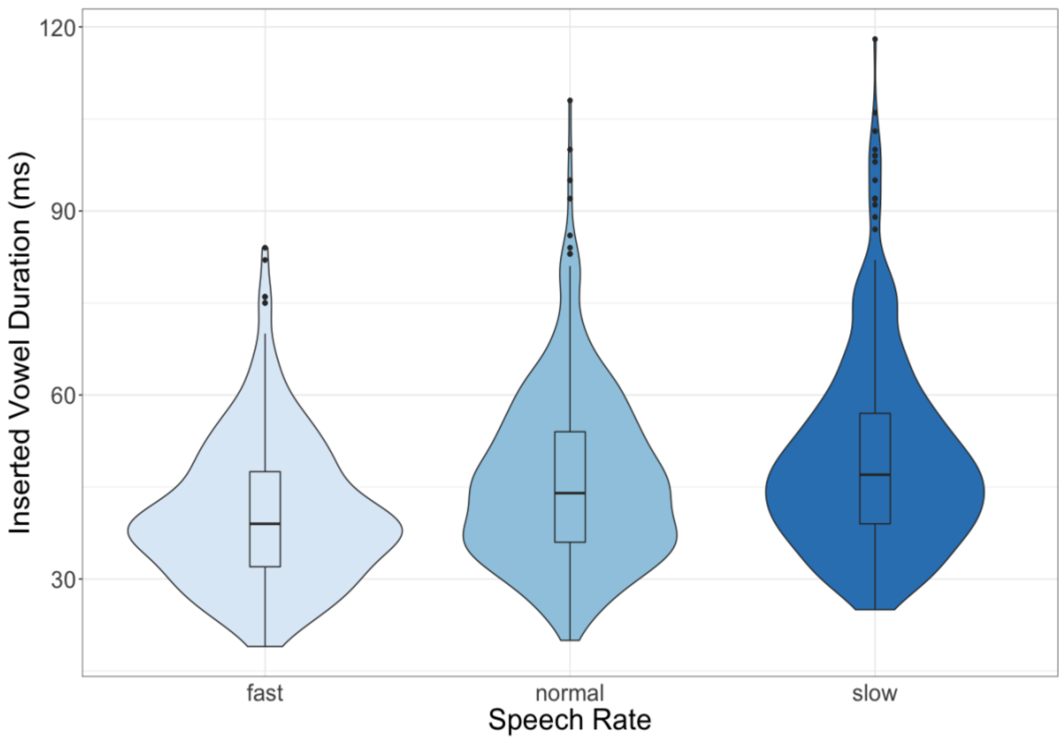

Figure 20. Average inserted vowel duration (ms) across subjects, by speech rate.
Note. Box and whisker plots display median values (horizontal line), along with distribution in quartiles.

A linear mixed effects model was fit to the data, using normalized Inserted Vowel Duration as the continuous dependent variable, Speech Rate as the categorical predictor variable,

[106] Raw durations were used (rather than normalized durations) to avoid losing generalizability by eliminating the unit of measurement (ms).

and Subject and Word as random intercepts. The results of the linear mixed effects model are summarized in Table 12.

Table 12. Linear Mixed Effects model for Inserted Vowel Duration as predicted by Speech Rate, with Subject and Word as random intercepts.

	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	1.009	0.02	26	44.29	<.0001***
Fast	-0.089	0.02	1195	-4.51	<.0001***
Slow	0.133	0.02	1190	7.45	<.0001***

Significance codes: <0.1 ‘.’ <0.05 ‘*’ 0.01 ‘**’ 0.001 ‘***’ (Alpha=.05)

The results summarized in Table 12 indicate that, as speaking rate increased, vowel duration increased, and as speaking rate decreased, vowel duration decreased. Since the theory makes no prediction for increased gestural overlap at fast speech between the consonants in coda positions (i.e. anti-phase gestures), but makes the prediction that phonological gestures lengthen in slow speech and shorten in fast speech, the effect found here provides support for the classification of the vowel as phonological, rather than excrescent.

4.6.2 Additional diagnostics for the vowel being phonological, not excrescent

The same diagnostic criteria used to establish the vowel as phonological in Chapter 3 were used in the current experiment. The diagnostics and relevant predictions are repeated in (17), below.

(17) Diagnostic criteria for the vowel being phonological (not excrescent)

a. Durational relations of adjacent gestures

(i) The inserted vowel will have the same acoustic consequences predicted for underlying gestures by a gestural coupling model. Specifically, the durations of the vowel and the flanking consonants will overlap more or less (i.e. have a negative or positive correlation), based on their syllable position (i.e. onset+nucleus, nucleus+coda).

b. Bimodal distribution of durations

(i) When plotting the distribution of the duration of the vowel in tokens with vowel insertion against tokens without vowel insertion (coded as having a duration of zero milliseconds), their distributions will be bimodal, not unimodal.

c. Consonant-to-consonant duration

(i) The total consonant-to-consonant duration (i.e. the duration from the acoustic onset of the first consonant to the acoustic offset of the second consonant) will be significantly larger in tokens with vowel insertion than in those without vowel insertion.

I will discuss the results of each diagnostic in the next three subsections.

4.6.2.1 Durational relations of adjacent gestures

The acoustic consequences of a gestural coupling model summarized in Chapter 3 were used in the current chapter to determine whether the duration of coupled gestures reflects their relationship as anti-phase gestures or in-phase gestures.

If the inserted vowel is excrescent, it is predicted that the flanking consonant gestures (i.e. C_1 and C_2) should have an anti-phase relationship, since excrescent vowels are not segments and do not offer a gesture to break up this relationship. This should result in the form of a positive correlation between C_1 and C_2, since anti-phase gestures are produced serially. If the inserted vowel is phonological, it is predicted that C_1 and C_2 would not have an anti-phase relationship, since the vowel would provide a gesture to break up the relationship between C_1 and C_2. This should result in a negative correlation or an absence of a correlation between the durations of C_1 and C_2. Since phonological vowels add a gesture to the sequence, C_1 and V should have an in-phase relationship, resulting in a negative correlation between their durations, while V and C_2 should have an anti-phase relationship, resulting in a positive correlation between their durations.

Pearson's correlations were conducted for normalized segment durations in each of the following gestural couples: C_1-V, V-C_2, and C_1-C_2. This was assessed in words with an inserted vowel between C_1 and C_2 (e.g. *farVm*) and in words without vowel insertion between C_1 and C_2 (e.g. *farm*), both produced in the *normal* speech rate.[107] For each subset of the data (Inserted Vowel, No Vowel) and for each gesture, outliers greater than 3 standard deviations above the mean were removed. A summary table is provided below (Table 13).

Table 13. Acoustic consequences of gestural coupling predicted for underlying vowels, phonologically inserted vowels, excrescent vowels, and vowel-less tokens in the normal speech rate.

	Correlation between durations		
Vowel Type in CC	**C_1-V**	**V-C_2**	**C_1-C_2**
Inserted Vowel (e.g. *farVm*) N=2151	0.194(.0001*)	-0.003(.9567)	-0.172 (.0001*)
No Vowel (e.g. *farm*) N=8248	--	--	-0.014(.6958)

Note. Table displays Pearson's r, followed by the p-value in parentheses. C_1=first consonant, V=intervening vowel, C_2=second consonant.

In words with vowel insertion, V and C_2 had a non-significantly negative correlation, while C_1 and V had a significantly positive correlation—an effect not predicted by a gestural coupling model. Assuming the vowel is phonological, two possible explanations for this might be that the intervocalic consonant either does not resyllabify to the onset (e.g.

[107] Where "V" represents an inserted vowel.

far.Vm) of the new syllable that the vowel creates or it acts as an ambisyllabic consonant (e.g. *fil.lVm*).[108] If this were true, it would mean that the intervocalic consonant is not timed with the inserted vowel because its timing has begun prior to the start of the new syllable.

Although the (non-significant) negative correlation between C_1 and C_2 in words *without* vowel insertion was not predicted for adjacent consonants in coda position, the negative correlation for C_1 and C_2 in words *with* vowel insertion is larger than in words without vowel insertion and highly significant. Since the relationship between C_1 and C_2 durations in words without vowel insertion is only used as a baseline for comparison, this set of results indicate that the inserted vowel creates an even more negative correlation between the durations of C_1 and C_2 than was found for vowel-less tokens. This suggests that the vowel acts to break up the coordination between these segments by offering a gesture to break up the sequence.

Since an excrescent vowel should not add a gesture to break up the anti-phase relationship that should exist for C_1 and C_2 in a coda, the significantly negative correlation between C_1 and C_2 in words with vowel insertion suggests that the vowel is phonological, not excrescent.

4.6.2.2 Bimodal distribution of durations

The second diagnostic I used was the distribution of the inserted vowel's duration plotted against zero-duration tokens (i.e. No Vowel Insertion). Bellik (2018) used this as a diagnostic for determining whether an inserted vowel was excrescent or epenthetic, arguing that a unimodal distribution would support a classification of the vowel as excrescent, in

108 Where "V" represents an inserted vowel and "." indicates a syllable boundary.

which a vowel is not categorically inserted and does not have a durational target. Excrescent vowels should have a length that is highly constrained by the degree to which the flanking consonants overlap, as they are the result of a brief period in which the vocal tract is left open between consonant production. Their distribution of duration should be much closer to zero-duration tokens as they are not deliberately inserted by the speaker and do not reflect a categorical process of vowel insertion. Conversely, a bimodal distribution would exclude the possibility that the vowel is excrescent, as this would suggest that the vowel is categorically inserted in some tokens and not others (i.e. phonological) and has a durational target noticeably dissimilar from vowel-less tokens (Bellik, 2018; Bürki et al., 2007). A histogram of raw vowel duration (in milliseconds), separated by Insertion (No Vowel Insertion vs. Vowel Insertion) is provided in Figure 21.

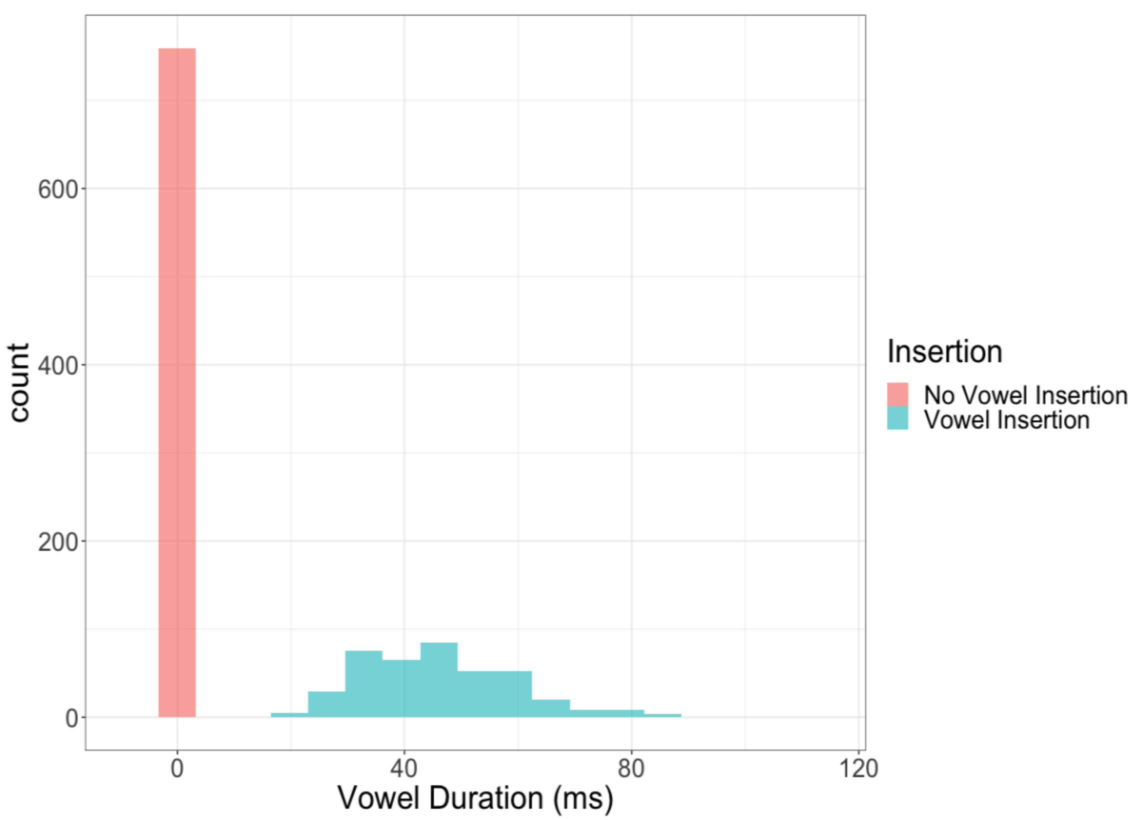

Figure 21. Bimodal distribution of counts for Vowel Duration (ms) by Insertion.

As was found in Chapter 3, there is a clear separation between the distribution of zero-duration tokens (i.e. No Vowel Insertion) and the distribution of tokens with an inserted vowel. Since the duration of excrescent vowels should have a distribution which merges with zero-duration tokens, forming a unimodal distribution, this result suggests that the vowel is phonological, not excrescent.

4.6.2.3 Consonant-to-consonant duration

Previous research has found that other putative cases of excrescent vowels do *not* significantly add length to the consonant clusters that they intrude upon (Ridouane & Fougeron, 2011). Obtaining a significant difference between the Consonant-to-Consonant Duration of tokens with an inserted vowel and of those without would provide further evidence that the vowel in Scottish English is not excrescent.

Consonant-to-Consonant duration was calculated by adding the normalized duration (in milliseconds) of C_1 and C_2 together in tokens that do not contain an inserted vowel, and adding the normalized duration of C_1, V, and C_2 together in the tokens containing an inserted vowel. A linear mixed effects model was fit to the data to determine whether Consonant-to-Consonant Duration was significantly different in tokens containing an inserted vowel from tokens without a vowel between C_1 and C_2 (Table 14). The continuous dependent variable was Consonant-to-Consonant Duration, the continuous predictor variable was Insertion (No Vowel Insertion vs. Vowel Insertion), with Subject and Word as random intercepts.

Table 14. Linear Mixed Effects model for Consonant-to-Consonant Duration as predicted by Insertion, with Subject and Word as random intercepts.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	2.012	0.011	1067	179.71	<.0001***
Vowel Insertion	0.929	0.024	1067	39.23	<.0001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

The results of the linear mixed effects model indicate that the vowel significantly adds length to the Consonant-to-Consonant Duration. This finding is consistent with a classification of the vowel as phonological, not excrescent.

4.7 Summary of findings and conclusions

4.7.1 The vowel is phonological (not excrescent)

4.7.1.1 Speech rate diagnostic

The result of the speech rate diagnostic is summarized in Table 15.

Table 15. Diagnosis of the vowel as phonological, not excrescent.

Diagnostic	Prediction for phonological vowels	Prediction for excrescent vowels	Diagnosis of current experiment
Speech Rate	The vowel will lengthen in slow speech and shorten in fast speech.	The vowel will not lengthen in slow speech or shorten in fast speech.	Phonological

4.7.1.2 Replication of Chapter 3 findings

A comparison of the results against each diagnostic prediction is provided in Table 16.

Table 16. Diagnosis of the vowel as phonological, not excrescent.

Diagnostic	**Prediction for phonological vowels**	**Prediction for excrescent vowels**	**Diagnosis of current experiment**
Durational relations of adjacent gestures	Significantly negative correlation for C_1 and V. Significantly positive correlation for V and C_2. Significantly negative correlation for C_1 and C_2.	No significantly negative correlation for C_1 and V. No significantly positive correlation for V and C_2. Significantly positive correlation for C_1 and C_2.	Phonological
Bimodal distribution of durations	Bimodal.	Unimodal.	Phonological
Consonant-to-consonant duration	Duration of consonant sequence significantly lengthened in words with vowel insertion, compared to those without.	Duration of consonant sequence not significantly lengthened in words with vowel insertion, compared to those without.	Phonological

4.7.2 The vowel is epenthetic (not lexicalized)

The diagnostics used in Chapter 3 to diagnose the vowel as epenthetic and not lexicalized were: (1) the vowel's duration compared with underlying vowels in similar environments and (2) the productivity of vowel insertion across a wide range of lexical items. Although this experiment did not include words with underlying vowels situated in similar environments (e.g. *forum*), and included a smaller range of test items (i.e. 20 experimental items in current experiment versus 317 in previous experiment), vowel insertion was observed in six lexical items which have not previously been reported to undergo vowel insertion. Taken together, the results of Chapters 3 and 4 indicate that the vowel in Scottish

English is epenthetic, since it is shorter than underlying vowels in similar environments (Chapter 3) and productive across a wide range of lexical items (Chapters 3 and 4).

4.7.3 Conclusions

The primary goal of this chapter was to use speech rate as a diagnostic for whether the inserted vowel was phonological or excrescent. I discussed Hall's (2006) criterion, which states that excrescent vowels should lengthen in slow speech and shorten in fast speech as a result of the flanking consonant gestures overlapping more in fast speech and less in slow speech. I noted that Hall cited empirical evidence for increased gestural overlap at fast speech for consonants separated by a morpheme or word boundary, and explained why empirical evidence for rate-induced gestural overlap should be specific to the phonological position where the consonants are situated (i.e. onset, coda, or C#C). I argued that Hall's (2006) criterion for how excrescent vowel duration in coda clusters should be affected by speech rate is not an obvious consequence of the way that gestures are coordinated under a gestural coupling framework. It is for this reason that I relied on evidence from controlled experiments which directly investigate vowel+consonant (i.e. anti-phase) gestures to generate a diagnostic for how the duration of phonological and excrescent vowels which surface in consonants that would otherwise form a coda should be affected by changes in speaking rate.[109] In the current experiment, I directly manipulated speech rate and measured the duration of the inserted vowel, and found that the vowel lengthened in slow

109 Note that, because the result of this diagnostic suggests that the vowel is phonological, this dissertation does not find direct evidence for whether *excrescent* vowels behave in a manner consistent with the prediction generated by the novel speech rate diagnostic that I provide. I encourage other researchers to utilize this diagnostic in tandem with other independent diagnostic criteria on vowels that display the phonetic and phonological properties associated with excrescent vowels in order to determine whether the theoretical prediction I make for excrescent vowels is supported by empirical evidence.

speech and shortened in fast speech, consistent with the predictions made for phonological vowels.

A secondary goal of this chapter was to re-assess the vowel against three diagnostics from Chapter 3, to establish that the vowel is phonological, not excrescent. The results of the three diagnostics used in Chapter 3 were replicated in the current experiment.

In Section 4.7.2, I argued that the vowel must be epenthetic, as its distribution across the (albeit limited) range of stimuli used in the current experiment was widespread. Vowel insertion in the current experiment occurred in six words that have not been previously reported to undergo vowel insertion in Scottish English.

In this chapter, I used a multi-step exclusion process to diagnose the vowel as phonological, and subsequently as epenthetic. I generated a novel speech rate diagnostic that can be used to distinguish between phonological and excrescent vowels. The results of the speech rate diagnostic indicated that the inserted vowel lengthened in slow speech and shortened in fast speech, consistent with the prediction for phonological vowels. In this chapter, the reliability of the diagnostic criteria used in Chapter 3 as well as the robust nature of the effects found in Chapter 3 were demonstrated in a separate sample of the population. Taken together, the results of the two sets of diagnostics used in this chapter, along with the classification in Chapter 3, suggest that vowel insertion in Scottish English is a process of phonological epenthesis.

Chapter 5:
The vowel is morpho-phonologically conditioned.

Previous reports of vowel insertion in Scottish English describe the vowel being limited to liquid+liquid and liquid+nasal sequences, specifically in complex coda clusters situated in monosyllabic words (see Chapter 2). If the vowel is epenthetic, one might infer based on this set of patterns that the vowel is inserted to repair the sonority cline of the coda cluster to satisfy the Sonority Sequencing Principle (see Chapter 2). However, Maguire (2017) notes that vowel insertion has been recorded in one disyllabic word where the consonant sequence (i.e. /ɹn/) is not in the coda position of the same syllable—*morning.* It is possible, then, that vowel insertion in Scottish English is epenthetic but not solely motivated to repair coda clusters, or alternatively, excrescent or lexicalized to certain consonant sequences—regardless of their syllable position. In this chapter, I will again use a multi-step exclusion process (as in Chapters 3 and 4) to diagnose the vowel as phonological, and subsequently as epenthetic using *phonological* diagnostics (i.e. those that assess the vowel's visibility to phonology).

5.1 The current chapter

In the previous two chapters, I used *phonetic* diagnostics (i.e. those that assess the vowel's acoustic, phonetic properties) to conclude that the inserted vowel in Scottish English liquid+liquid and liquid+nasal sequences was phonological.[110,111] The goal of the current chapter is to use *phonological* diagnostics (i.e. those that assess the vowel's visibility to

110 All but the "Widespread distribution across lexical items" diagnostic.
111 In this chapter, I will continue using the term "inserted" until I diagnose the vowel as epenthetic.

phonology) to diagnose the vowel as phonological, and subsequently, as epenthetic.[112] In this chapter, I will use the following diagnostics to classify the vowel as phonological (18).

(18) Diagnostic criteria for the vowel being phonological (not excrescent)

a. Morpho-phonological distribution

(i) The inserted vowel will have a systematic distribution within predictable morpho-phonological environments.

b. Interaction with other phonological processes

(i) The inserted vowel will be phonologically visible (i.e. interact with separate phonological processes).

After using the diagnostics in (18) to establish that the vowel is phonological, I will use the diagnostic in (19) to conclude that the vowel is epenthetic:

(19) Diagnostic criteria for the vowel being epenthetic (not lexicalized)

a. Environment targeted by other phonological repair strategies.

(i) The inserted vowel will appear in environments which are otherwise repaired by way of separate phonological processes.

I will use data from the two production experiments summarized in Chapters 3 and 4 to provide further evidence that the vowel is phonological, and subsequently, epenthetic. Data from the first production experiment (summarized in Chapter 3) will be used to demonstrate that the vowel has a systematic morpho-phonological distribution (18a). Data from the second language production experiment (summarized in Chapter 4) will be used to demonstrate that the vowel interacts with other phonological processes (18b). After establishing that the vowel is phonological using the diagnostics in (18), I will use data

[112] Hall (2006) describes epenthetic vowels as being "visible" to the phonology, as they originate in the phonology module of the framework.

from the second language production experiment to demonstrate that the vowel does not co-occur with other phonological processes which target the same consonant sequences (19a). Doing so will allow me to exclude the possibility that the vowel is lexicalized, and conclude that the vowel is epenthetic.

5.2 Organization of the chapter

I will summarize the diagnostics in (18) and (19) in Sections 5.3 and 5.4, respectively. In each section (5.3 and 5.4), I will repeat the stimuli used in each experiment, and results of each diagnostic. I will conclude in section 5.5.

5.3 Diagnostic criteria for the vowel being phonological (not excrescent)

5.3.1 Systematic morpho-phonological distribution

The first diagnostic I used was the vowel's distribution across varying morpho-phonological environments. Since phonological vowels should have a predictable morpho-phonological environment, but excrescent vowels should not, I used the vowel's distribution in words with different morphological suffixes (e.g. *farm*, *farm+ed*), in consonant sequences in coda and non-coda positions, and in a consonant sequence which violates the universal Sonority Sequencing Principle (SSP) (Selkirk, 1984).[113,114] The data that I assessed against this diagnostic come from the first language production experiment

113 Where "+" indicates a morpheme boundary.

114 Note that lexicalized vowels can exhibit phonological sensitivity, even if they are not as productive as epenthetic vowels, since they may have been epenthetic vowels before becoming fossilized in the underlying representation.

(summarized in Chapter 3).[115] I present the stimuli in Section 5.3.1.1, and the results in Section 5.3.1.2.[116]

5.3.1.1 Stimuli

The stimuli were those used in the first language production experiment (summarized in Chapter 3). Stimuli were sectioned into two blocks: one containing words from a Scottish Standard English lexis and one containing words specific to a Scots lexis, along with Scottish place names and surnames. Both blocks of the experiment contained words where the consonant sequence was situated within a syllable coda and across a syllable boundary.

5.3.1.1.1 Block 1: Scottish Standard English words

38 five-word sets (e.g. *farm*, *farms*, *farmed*, *farmer*, *farming*) (APPENDIX B) were used. Each item was repeated twice in the experiment, and all items were randomized. The experimental stimuli consisted of words containing one of the following Consonant Sequences: /ɹl/ (e.g. *twirl*), /ɹm/ (e.g. *farm*), /ɹn/ (e.g. *warn*), or /lm/ (e.g. *film*). Control items included 29 five-word sets containing other attested onset and coda clusters. Control items were included to ensure that participants did not guess the target region of each word, i.e. consonants in coda clusters (e.g. *fa<u>rm</u>*) or consonants situated across a syllable boundary (e.g. *fa<u>r.m</u>ing*).[117]

Items contained a bare root or a root plus a suffix. The suffixes used were: the past tense /d/ (i.e. -ed), the plural /z/ (i.e. -s), agentive /əɹ/ (i.e. -er), comparative /əɹ/ (i.e. -er),

[115] Data from the first experiment were used because they include consonant sequences which are situated within and across syllables, and this diagnostic is concerned with the distribution of vowel insertion in coda and non-coda positions.
[116] Please refer to Chapter 3 for remainder of the methodology used in this experiment.
[117] Where "." indicates a syllable boundary.

and /ɪŋ/ (i.e. -ing).[118] Consonants in words that consisted of a bare root and consonantal suffixes (i.e. /z/ and /d/) are a part of a syllable coda (e.g. *farms*, *farmed*), while consonants in words containing vowel-initial suffixes (i.e. /əɹ/ and /ɪŋ/) are separated by a syllable boundary (e.g. *far.ming*, *far.mer*), since these suffixes provide a new syllable nucleus to the word they attach to.[119] Example stimuli are provided in (20).

(20)

Root (+grammatical morpheme)	*Word*	*Phonetic Transcription*
root	*farm*	[fɑɹm]
root + /z/	*farm*s	[fɑɹm+z][120]
root + /d/	*farmed*	[fɑɹm+d]
root + /ɪŋ/	*farming*	[fɑɹ.m+ɪŋg]
root + /əɹ/ (agentive)	*farmer*	[fɑɹ.m+əɹ]
root + /əɹ/ (comparative)	*warmer*	[wɔɹ.m+əɹ]

5.3.1.1.2 Block 2: Scottish place names, surnames, and Scots words

The second block of the study was included to increase the number of relevant words and encourage speakers to utilize both ends of the bipolar linguistic continuum, since vowel insertion has been reported in Scottish Standard English and Scots.

Stimuli included 127 experimental items and 70 control items (APPENDIX B). Words included Scottish place names, surnames, and Scots words collected from *Scotland's Places* (scotlandsplaces.gov.uk), *Scotland's People* (scotlandspeople.gov.uk), and the *Dictionary of the Scots Language* (dsl.ac.uk), respectively. Each item contained

118 The comparative morpheme /əɹ/ was used in only a handful of stimuli (e.g. *warmer, firmer, calmer*) since most of the items were verbs and could only be combined with agentive /əɹ/ morphemes (e.g. *filmer, farmer*).
119 Although some argue that these consonants are word-final appendices (Borowsky, 1986; Vaux & Wolfe, 2009), I will regard them as coda clusters, since prior reports of vowel insertion in Scottish English have regarded them as such.
120 Where "+" indicates a morpheme boundary.

one of the Consonant Sequences (/ɹl/, /ɹm/, /ɹn/, /lm/, and /ln/) in one of two syllable positions: coda or non-coda position. Example stimuli are provided in (21).

(21)

Word Type	*Position*	*Word*	*Translation*	*Phonetic Transcription*
Scots word	Coda	skirl	‘to shriek’	[skʌɾɫ]
Scots word	Non-Coda	contermashious	‘obstinate’	[kɔn.tər.mɑ(:).ʃəs]
Place name	Coda	Leckmelm		[ɫɛk.mɛɫm][121]
Surname	Non-Coda	Kilmer		[kIɫ.mər]

A summary table of the number of items for Word Type and Position of Consonant Sequence are presented in Table 17.

Table 17. Summary of experimental stimuli by Word Type and Position of Consonant Sequence.

Word Type	**Number of Items**	**Position of Consonant Sequence**	
		Coda	**Non-Coda**
Place Name	32	9	23
Scots Word	79	43	36
Surname	10	4	6
Total	121	56	65

5.3.1.2 Results

In this section, I will summarize the distribution of vowel insertion across varying phonological environments from the first (5.3.1.2.1) and second (5.3.1.2.2) block of the experiment.

121 Note that the word-initial lateral is transcribed with a dark [ɫ] because Scottish English has dark [ɫ] in all syllable positions, unlike other dialects of English (Johnston, 1997; Wells, 1982).

5.3.1.2.1 Block 1: Scottish Standard English words

If vowel insertion is a phonological repair strategy used to break up complex coda clusters, it should surface most often in places where the consonants are a part of a coda cluster, i.e. monosyllabic, monomorphemic words where the consonant sequence is situated in the coda position (e.g., [fIłIm], [wɑɾʌm]) or in words which have undergone consonantal suffixation (i.e., +/z/, +/d/), as the consonantal suffix cannot offer a syllable nucleus to break up the coda cluster. Conversely, vowel insertion should surface least often in words which have undergone vowel-initial suffixation (i.e. +/əɹ/, +/ɪŋ/) because the vowel-initial suffixes offer a viable syllable nucleus that may trigger a resyllabification. A resyllabification would incorporate the consonant [m] into the second syllable as an onset, thus breaking up the dis-preferred coda cluster and blocking the deliberate insertion of a vowel (i.e. [fIł.mɪŋ], [wɑɾ.mər]). If the inserted vowel is not phonological, but excrescent, it should not have a systematic distribution across specific phonological environments, as excrescent vowels are not deliberately inserted as a phonotactic repair, but are an unintended consequence of low-level articulatory constraints.

To determine whether vowel insertion occurred most often between coda consonants, I measured the distribution of vowel insertion across Suffix type (root, +/z/, +/d/, +/əɹ/, +/ɪŋ/). Vowels were detected more often in words where the consonants were a part of the syllable coda (i.e. root, +/z/, +/d/) than in words where the consonant sequence was separated by a syllable boundary (i.e. +/əɹ/, +/ɪŋ/) (Figure 22).

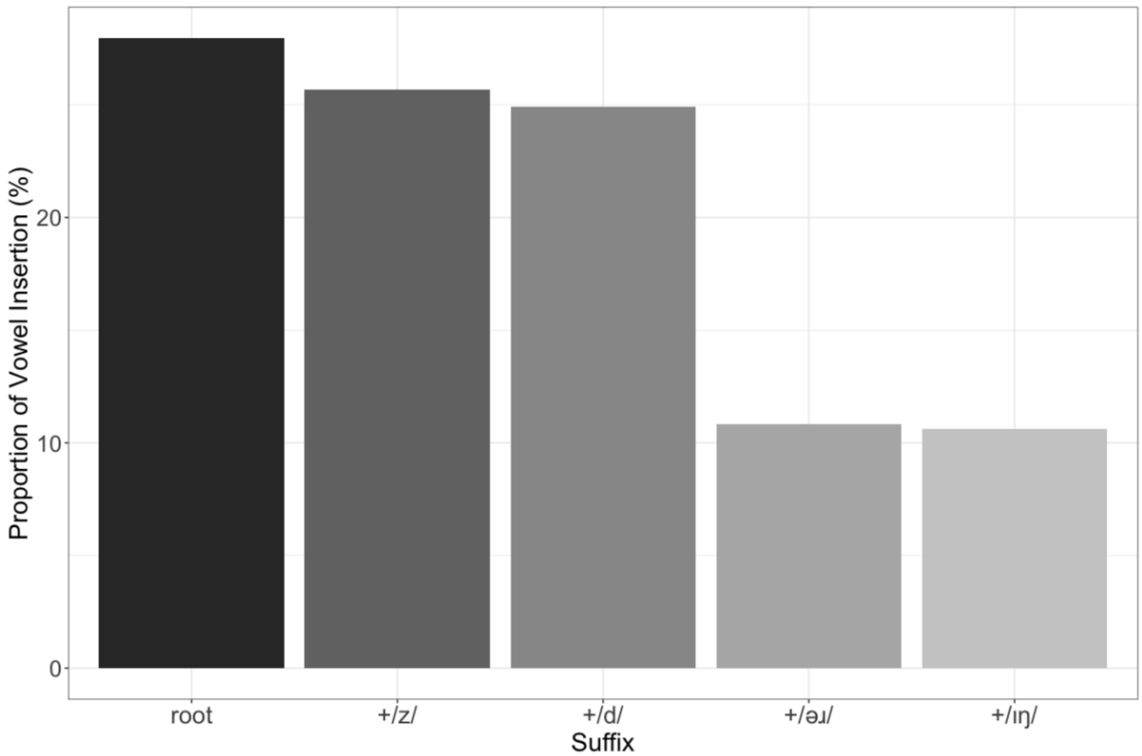

Figure 22. Proportion (%) of Vowel Insertion by Suffix.

This pattern was expected if the vowel is phonological, since excrescent vowels should not have a systematic distribution across varying morpho-phonological environments. Although Figure 22 demonstrates that vowel insertion occurred mainly in coda clusters, it also occurred in non-coda positions a little over 20% of the time.[122]

A mixed effects logistic regression confirmed that the difference in the distribution of vowel insertion between coda and non-coda positions observed by visual inspection of Figure 22 was statistically significant, with vowel insertion occurring more in coda clusters. The mixed effects logistic regression was conducted using Insertion (No Vowel Insertion vs. Vowel Insertion) as the binary, categorical dependent variable, and Position

[122] I will discuss a possible motivation for vowel insertion in non-coda positions in Chapter 6.

of Consonant (Coda vs. Non-Coda) as the categorical predictor variable, with Subject and Word as random intercepts (Table 18).

Table 18. Mixed Effects Logistic Regression for Insertion as predicted by Position of Consonant Sequence using Subject and Word as random intercepts.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-3.323	0.36	-9.11	<.0001***
Coda	1.709	0.28	5.99	<.0001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

To determine whether this pattern was consistent across speakers, I visualized the distribution of vowel insertion by Subject and Suffix (root, +/z/, +/d/, +/əɹ/, +/ɪŋ/). Doing so revealed an implicational hierarchy, in which, if speakers produced a vowel between consonants which straddle a syllable boundary (e.g. *far.ming*, *far.mer*), speakers also produced a vowel between consonants which formed a coda cluster (e.g. *farm*, *farms*, *farmed*) (Figure 23).[123]

123 Note that Subject labels are "S" for "subject" plus the number corresponding to the list in APPENDIX D. Missing numbers (e.g. "S2") indicate subjects which were excluded from the analysis for not satisfying the recruitment criteria.

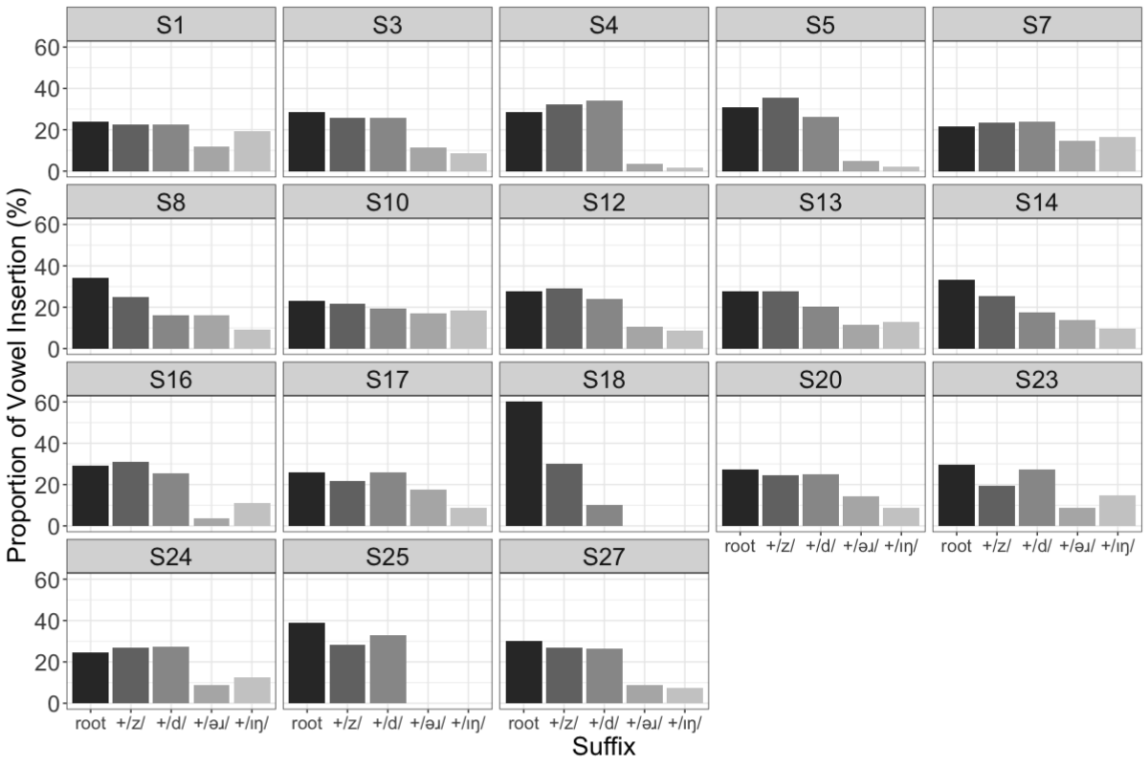

Figure 23. Proportion (%) of Vowel Insertion by Suffix and Subject.

In Figure 23, vowels were not identified between consonants separated by a syllable boundary (i.e. +/əɹ/, +/ɪŋ/) for only two subjects—S18 and S25; however, they were identified in all of the possible environments where the consonants constitute a coda cluster (i.e. root, +/z/, +/d/).[124]

If the vowel is inserted to repair the sonority relationship of liquid+liquid and liquid+nasal coda clusters, it should surface more often in clusters that violate the universal Sonority Sequencing Principle (SSP), which dictates that segments in a syllable coda should decrease in sonority with respect to the syllable nucleus (Selkirk, 1984).[125] Of the Consonant Sequences (/ɹl/, /ɹm/, /ɹn/, /lm/) which are subject to vowel insertion, /ɹl/ should

124 I use by-subject data to visualize the implicational hierarchy which emerged for vowel insertion by Suffix type, however, a thorough investigation of the individual differences obtained in either experiment is outside of the scope of the current dissertation.

125 I refer the reader to Chapter 1 for a complete discussion of epenthesis motivated by sonority violations.

be the most likely candidate for vowel insertion, since the consonants in /ɹl/ form a sonority plateau, while the others decrease in sonority from the syllable nucleus, in line with the SSP. Although the *least* preferential sonority sequencing in a coda is a reverse sonority cline (i.e. consonants rise in sonority away from the syllable nucleus), sonority plateaus are also a violation of the SSP (Kreitman, 2006). Visualizing the distribution of vowel insertion by Consonant Sequence (/ɹl/, /ɹm/, /ɹn/, /lm/) reveals that /ɹl/ received more vowel insertion in places where the cluster occupied a coda (i.e. root, +/z/, +/d/) than did the other consonant sequences, as expected (Figure 24).

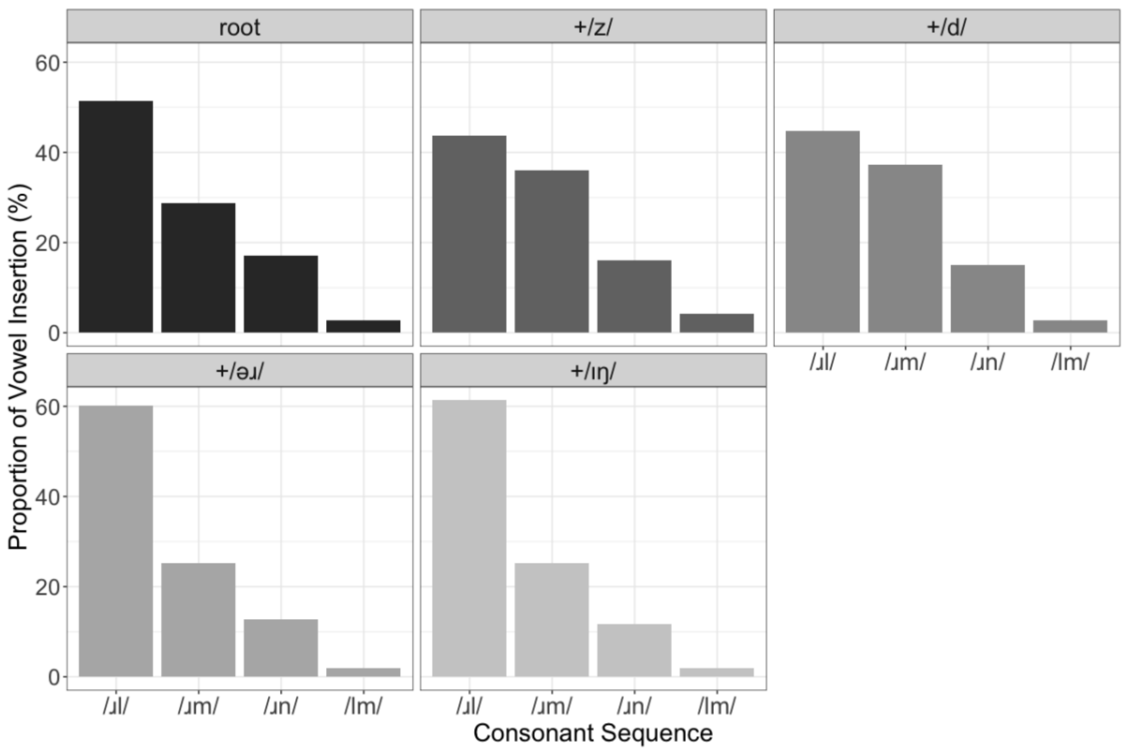

Figure 24. Proportion (%) of vowel insertion by Suffix, grouped by Consonant Sequence.

Since Figure 24 reveals that vowel insertion also occurred in non-coda positions (i.e. +/əɹ/, +/ɪŋ/), I visualized the distribution of vowel insertion in /ɹl/ consonant sequences by Suffix to see whether vowel insertion in /ɹl/ sequences was maintained across a syllable boundary (i.e. +/əɹ/, +/ɪŋ/). If the high degree of vowel insertion in /ɹl/ is due to a sonority sequencing violation, this pattern should be maintained in /ɹl/ consonant sequences situated across a syllable boundary, since the Syllable Contact Law states that consonants should decline in their sonority, maximizing their sonority dispersion across syllables (Vennemann, 1988). Since the consonants in an /ɹl/ sequence have the same sonority, it may be the case that inserting a vowel in this consonant sequence when a syllable boundary intervenes is a repair to a violation of the Syllable Contact Law. Figure 24 reveals that, not only did the /ɹl/ consonant sequences receive the most vowel insertion, but vowel insertion which occurred in words with /əɹ/ and /ɪŋ/ suffixes was present most in /ɹl/ consonant sequences. Not only was the pattern established in Figures 23 and 24—that vowel insertion occurs more in coda positions (i.e. root, +/z/, +/d/) than in consonants separated by a syllable boundary (i.e. +/əɹ/, +/ɪŋ/)—maintained in /ɹl/ consonant sequences; this patterned was maintained for all consonant sequences.[126]

The high distribution of vowel insertion in the /ɹl/ consonant sequence across a syllable boundary (i.e. +/əɹ/, +/ɪŋ/), implicational hierarchy in Figure 23, and high distribution of vowel insertion in /ɹl/ clusters within and across a syllable suggests that the vowel has a predictable morpho-phonological distribution, as expected for phonological vowels.

5.3.1.2.2 Block 2: Scottish place names, surnames, and Scots words

In this block of the experiment, mono- and di- syllabic words where the consonant sequence was situated word-finally (e.g. *skoolm, Dachalm*) were expected to undergo more

[126] Note that this pattern was identified using visual inspection of the data.

vowel insertion than words where the consonant sequence was situated word-internally (e.g. *Kilmer, contermashious*) if the vowel is phonological.

Visualizing the percentage of vowel insertion by Position of Consonant Sequence (Coda vs. Non-Coda), vowel insertion occurred more frequently in complex coda clusters, as expected if the vowel is phonological (Figure 25).

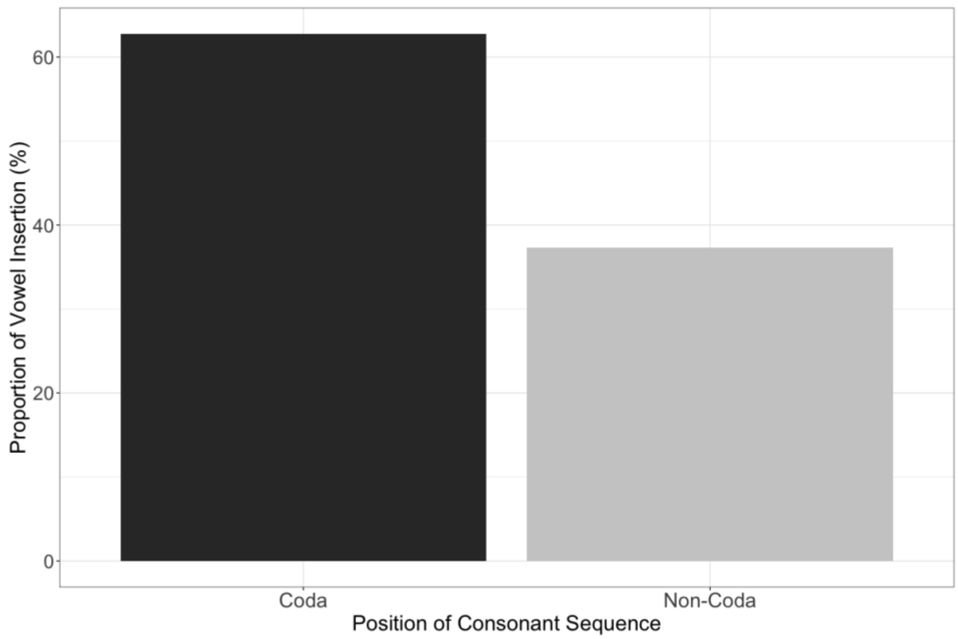

Figure 25. Proportion (%) of Vowel Insertion by Position of Consonant Sequence.

To determine whether this pattern was maintained across Word Type (Place Names, Scots Words, Surnames), I calculated the proportion of words with and without vowel insertion for each level of Word Type by Position of Consonant Sequence (Coda vs. Non-Coda). Doing so revealed that the pattern of vowel insertion occurring more often in complex coda clusters was maintained in Scots words, but not place names and surnames (Table 19).

Table 19. Proportion of Insertion by Word Type and Position of Consonant Sequence.

Insertion	Word Type	Position of Consonant Sequence	
		Coda	**Non-Coda**
No Vowel Insertion N=3,486	Place Name (N=953)	31%	69%
	Scots Word (N=2,199)	51%	49%
	Surname (N=334)	39%	61%
	Total	44%	56%
Vowel Insertion N=689	Place Name (N=135)	14%	86%
	Scots Word (N=539)	76%	24%
	Surname (N=15)	47%	53%
	Total	63%	37%

Table 19 illustrates that vowel insertion occurred in consonants that were a part of a coda cluster more often than in consonants separated by a syllable boundary, but that this difference is primarily driven by vowel insertion which occurs in Scots words. This is likely due to low number of experimental items for place names and surnames (see Table 17).

To determine whether there was a difference across the two levels of Position of Consonant Sequence (Coda vs. Non-Coda), a mixed effects logistic regression was run using Insertion as the binary dependent variable, Position of Consonant Sequence as the categorical predictor variable, and Subject and Word as random intercepts. Result indicate that Coda position significantly predicted vowel insertion (Table 20).

Table 20. Mixed Effects Logistic Regression for Insertion as predicted by Position of Consonant Sequence, with Subject and Word as random intercepts.

	Estimate	**Std. Error**	**z value**	**Pr(>|t|)**
(Intercept)	-3.156	0.33	-9.67	<.0001***
Coda	1.301	0.31	4.14	<.0001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

In the previous subsection, I investigated the distribution of vowel insertion in each Consonant Sequence to determine whether the distribution of vowel insertion is sensitive to sonority sequencing. In the first block of the experiment, /ɹl/ received the most amount of vowel insertion out of all Consonant Sequences, and this was also the case in the second block of the experiment: the /ɹl/ consonant sequence received vowel insertion in 52% of all /ɹl/ tokens, while the other clusters received vowel insertion less than 20% of the time (i.e. 15% for /ɹm/, 16% for /ɹn/, 4% for /lm/, 0% for /ln/).

Visualizing the proportion of vowel insertion by Consonant Sequence and Position of Consonant Sequence revealed that the /ɹl/ consonant sequence received more vowel insertion in non-coda positions, while the other consonant sequences received more vowel insertion in coda positions (Figure 26).

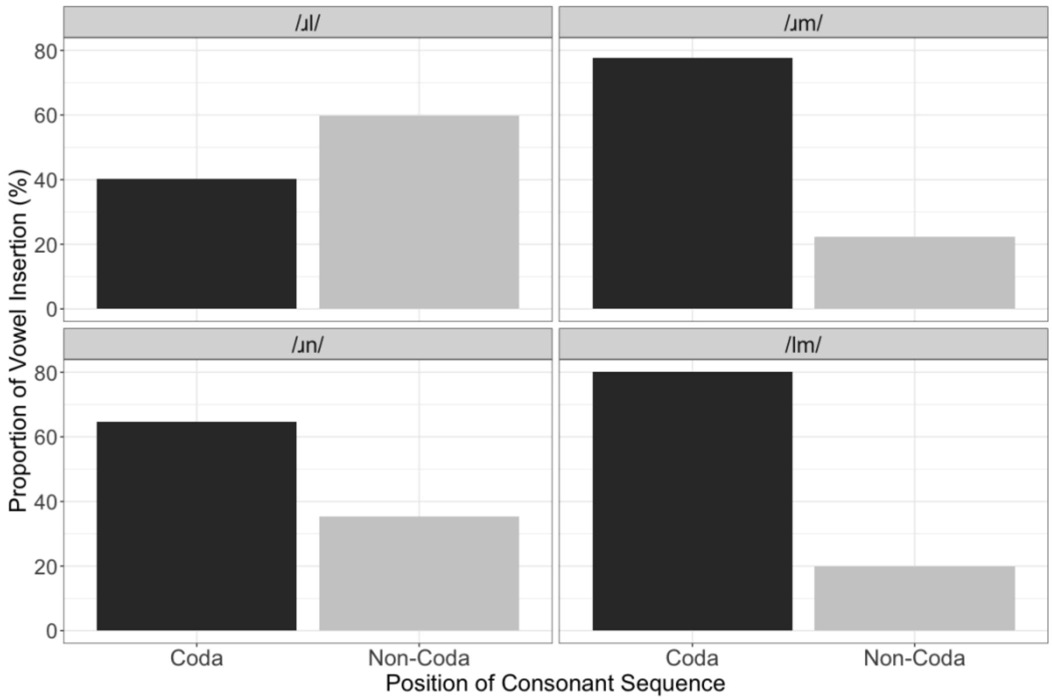

Figure 26. Proportion (%) of vowel insertion by Suffix, grouped by Position of Consonant Sequence.

The tendency for vowel insertion to occur more in non-coda positions for /ɹl/ suggests that vowel insertion in /ɹl/ may be a repair to the consonant sequence having a sonority plateau—both within the same syllable and across a syllable boundary.

The results of both blocks of the experiment suggest that vowel insertion in Scottish English is sensitive to syllable structure and sonority sequencing. The systematic distribution of vowel insertion in predictable phonological environments suggests that this process is phonological, not excrescent.

5.3.2 Interaction with other phonological processes

The second diagnostic I used was whether the vowel interacted with other phonological processes. The data that I assess against this diagnostic come from the second language production experiment (summarized in Chapter 4).[127] I present the stimuli in Section 5.3.2.1, and the results in Section 5.3.2.2[128] To determine whether the vowel interacts with other phonological processes, I investigated the co-occurrence of vowel insertion with rhotic allophony.

Phonological allophony involves the categorical switch between distinct allophones which is conditioned by phonological environment—for example, /l/-allophony in English onsets versus codas (Chomsky & Halle, 1968) (see Chapter 2). Although some have argued that /l/-allophony in English is not categorical, but a gradient effect in which speakers produce darker or lighter /l/ allophones (Sproat, Richard & Fujimura, 1993), rhotic allophony between approximant [ɹ] and tapped [ɾ], for example, involve very different articulations.

Since, in Scottish English, tapped [ɾ] prefers intervocalic positions (Jauriberry et al., 2015), I expect that, if the vowel is phonological, it should co-occur most with tapped [ɾ], while tokens without vowel insertion should be produced with a tapped [ɾ] allophone less often. If the insertion of a vowel triggers rhotic allophony, in which speakers produce a tapped [ɾ] allophone in tokens with vowel insertion and a separate rhotic allophone in tokens without vowel insertion, I will interpret this as evidence that the vowel is phonological. Since excrescent vowels should not be phonologically visible in the sense that they interact with other phonological processes, the degree to which speakers use one rhotic allophone versus another in tokens with and without vowel insertion should not be significantly different if the vowel is excrescent.

[127] Data from the second experiment were used because they only include consonant sequences situated in coda position. Since this diagnostic is concerned with the vowel's interaction with phonological processes affecting C_1 and C_2, I controlled for syllable position to ensure there were no confounds of syllable structure.
[128] Please refer to Chapter 4 for remainder of the methodology used in this experiment.

5.3.2.1 Stimuli

The stimuli were those used in the second language production experiment (summarized in Chapter 4). Stimuli included 20 unique sentence frames, each 8-10 words in length. Each of the 20 sentences was repeated 10 times within each of the three speech rate blocks of the experiment, for a total of 600 target utterances (20 sentences x 10 repetitions x 3 speech rates). Target words consisted of monosyllabic, monomorphemic items containing liquid+liquid and liquid+nasal consonant sequences (/ɹl/, /ɹm/, /ɹn/, /lm/) occupying the coda position. The /ln/ sequence was not included since no vowel insertion was observed in this consonant sequence in the second block of the first experiment. In all but one word (i.e. *world*), the relevant consonant sequence was situated word-finally (i.e. *film, farm*). In each sentence frame, the target word was located in the second half of the sentence (e.g. *I hope I'll get better marks next* <u>*term*</u>*; On the horizon, a big* <u>*storm*</u> *approached the beach*) (APPENDIX F).[129] This was done to ensure that participants had time to adjust to the appropriate speaking rate (i.e. fast, normal, slow). In all sentence positions, target words were nouns.

5.3.2.2 Results

I tested whether the C_1 allophone used in rhotic-initial consonant sequences (/ɹl/, /ɹm/, /ɹn/) differed between tokens with and without vowel insertion to determine whether the inserted vowel interacted with rhotic allophony.

Plotting the percentage of C_1 Allophone (Approximant [ɹ], Tapped [ɾ], Trilled [r], Fricativized /ɹ/) used by each level of Insertion (No Vowel Insertion vs. Vowel Insertion)

[129] Target words are underlined.

revealed that approximant [ɹ] and tapped [ɾ] were used most often in all utterances, but the percentage of tapped [ɾ] used was doubled in words with vowel insertion (Figure 27).[130]

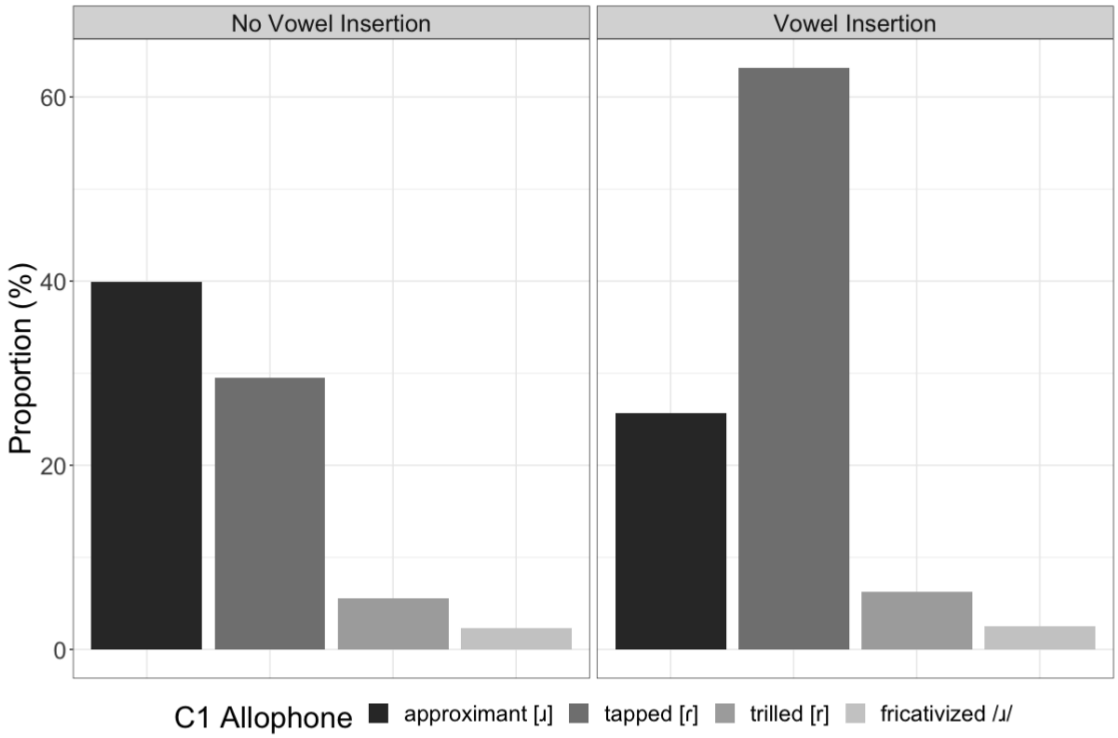

Figure 27. Proportion of C_1 Allophone used by Insertion.

Visualizing the results by Subject, the strong use of tapped [ɾ] in words with vowel insertion is maintained across subjects (Figure 28).[131, 132]

130 See Jauriberry et al. (2015) for a discussion of fricativized approximant /ɹ/ in Scottish English.
131 Note that Subject labels are "S" for "subject" plus the number corresponding to the list in APPENDIX H. Missing numbers (e.g. "S2") indicate subjects which were excluded from the analysis for not satisfying the recruitment criteria.
132 Data here include tokens in all three levels of the Speech Rate variable, since visual inspection of C1 Allophone across the three levels of Speech Rate in tokens with vowel insertion compared to tokens without vowel insertion confirmed this pattern. Specifically, speakers used approximant [ɹ] in tokens without vowel insertion >50% of the time in each speech rate, and tapped [ɾ] in tokens with vowel insertion >60% of the time.

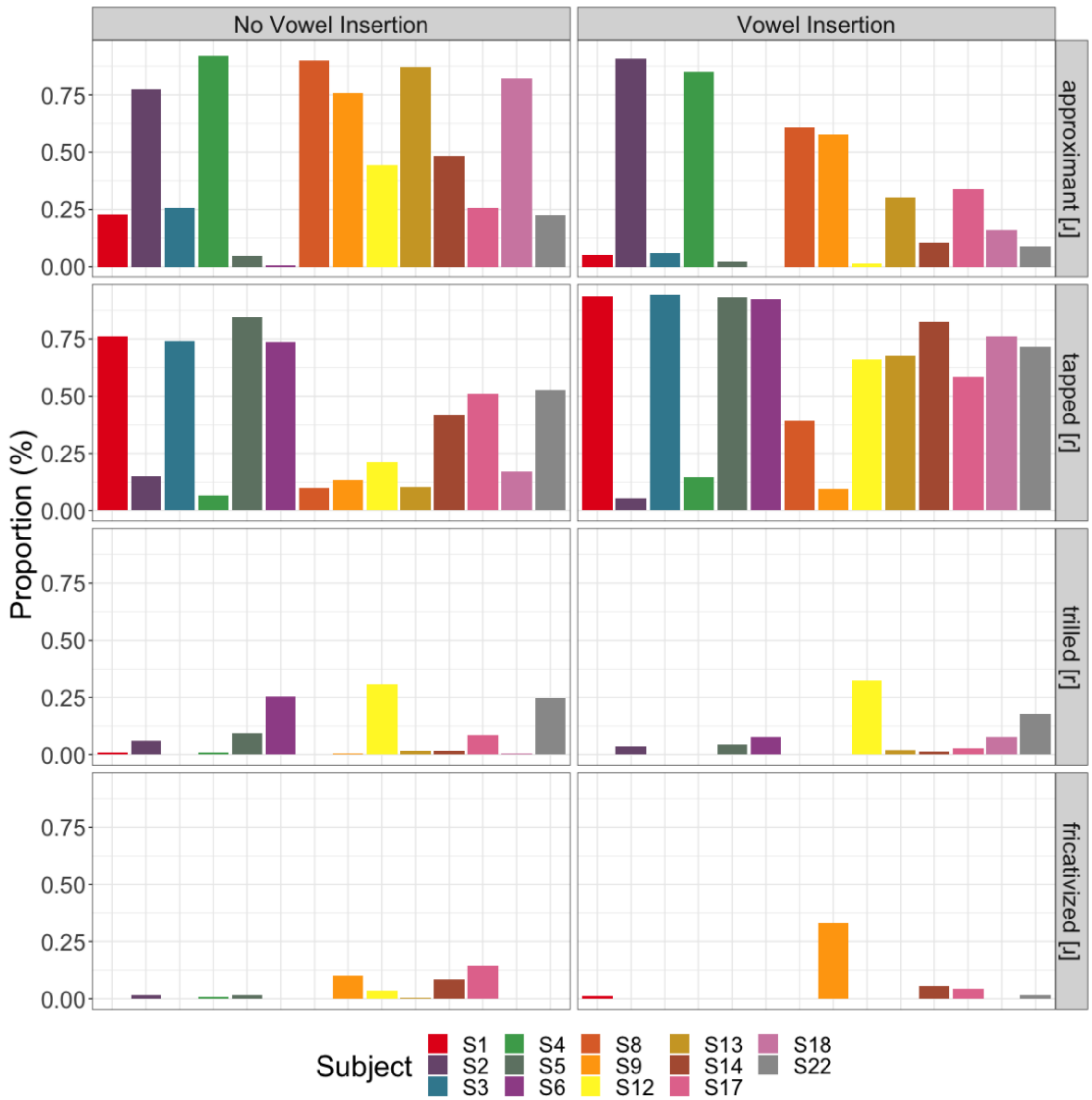

Figure 28. C_1 Allophone usage by Subject and Insertion.

Visual inspection of Figure 28 reveals that, of all rhotic-initial utterances containing an inserted vowel (Right panel), tapped [ɾ] was used over 50% of the time for 10 of the 14 subjects. Visually, there appears to be a tradeoff in which speakers use approximant [ɹ] more often than tapped [ɾ] in tokens without vowel insertion (e.g. *farm*) and use tapped [ɾ] more than approximant [ɹ] in words which contain inserted vowels (e.g. *forVm*).[133]

133 Note that Figure 28 also illustrates that some speakers do not exhibit rhotic allophony. This pattern may suggest that the vowel is not merely an excrescent vowel which surfaces in tapped [ɾ]-initial sequences because of the way that tapped [ɾ] is coordinated with the following consonant. The vowel is also produced

To determine whether vowel insertion interacts with rhotic allophony (e.g. the categorical switch between approximant [ɹ] and tapped [ɾ]), I ran a binary logistic regression for Insertion (No Vowel Insertion vs. Vowel Insertion) as predicted by C_1 Allophone (Tapped [ɾ] vs. Approximant [ɹ]) and found that Tapped [ɾ] was significantly used more in words with vowel insertion (Table 21).

Table 21. Logistic Regression for Insertion as predicted by C_1 Allophone with Subject and Word as random intercepts.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**
(Intercept)	-1.532	0.76	-2.00	<.05*
Tapped [ɾ]	1.123	0.14	8.15	<.001***

Significance codes: <0.1 '.' <0.05 '*' 0.01 '**' 0.001 '***' (Alpha=.05)

This interaction with rhotic allophony is further support for the classification of the vowel as phonological, since excrescent vowels are not visible to the phonology and should not interact with rhotic allophony.

5.4 Diagnostic criteria for the vowel being epenthetic (not lexicalized)

5.4.1 Absence of vowel insertion in words with C_1 deletion, derhoticization, /l/-vocalization, or syllabic C_2s

If the phonological vowel is a deliberate repair for dis-preferred consonant sequences, and in this experiment reflects an avoidance of coda clusters, specifically, it may not be the only phonological repair that speakers use.[134] One of Hall's (2006) criteria for epenthetic

with approximant [ɹ], which should theoretically have a greater length and potentially a greater degree of gestural overlap with the following consonant than tapped [ɾ], since approximant [ɹ] is produced with continuant voicing, while tapped [ɾ] is a short, non-continuant sound.

134 Note that lexicalized vowels can exhibit phonological sensitivity, even if they are not as productive as epenthetic vowels, since they may have been epenthetic vowels before becoming fossilized in the underlying representation.

vowels is that they arise in phonological environments which other phonological processes target in that language. I expected that, if the vowel was phonological, it would not co-occur with C_1 deletion, C_1 weakening, or syllabic C_2s, since they may independently function as a repair for dis-preferred coda clusters. Since the deletion of C_1 would eliminate the sequence of consonantal gestures that an excrescent vowel could intrude upon, I use this diagnostic to distinguish between epenthetic and lexicalized vowels, rather than to distinguish phonological and excrescent vowels, since lexicalized vowels should target sequences that are not repaired by way of other phonological processes, since lexicalized vowel should surface most, if not all, of the time.

Coda clusters can be repaired by way of: (1) deleting the C_1 so that there is no consonant sequence, (2) weakening the articulation of C_1 so that it is less perceptible and/or the consonant sequence is easier to articulate (i.e. derhoticized /ɹ/, /l/-vocalization), or (3) using a syllabic C_2 (i.e. a consonant that can constitute a syllable nucleus) to trigger a resyllabification which incorporates C_2 into a new syllable (e.g. *farm* [fɑɹ.m̩]).[135,136]

Calculating the counts of vowel insertion in words that had C_1 deletion (i.e. rhotic or lateral deletion) or a syllabic C_2 revealed that vowel insertion did not co-occur with these processes at all. Vowel insertion co-occurred with a derhoticized /ɹ/ five times, but this was out of 67 possible utterances containing a derhoticized /ɹ/ (Table 22).

135 Where “ ̩ ” indicates a syllabic consonant (i.e. a consonant which can be a syllable nucleus).
136 For more information on derhoticization and /l/-vocalization in Scottish English, see Chapter 2.

Table 22. Absence of vowel insertion in words with C_1 deletion, derhoticization, /l/-vocalization, or syllabic C_2s.

Process	No Vowel Insertion	Vowel Insertion
Derhoticized /ɹ/	62	5
/l/-vocalization	29	0
Rhotic deletion	23	0
Lateral deletion	174	0
Syllabic C_2	250	0
Total (N=543)	538	5

The results summarized in Table 22 reveal a categorical avoidance of vowel insertion in places where the coda cluster was broken up by way of C_1 deletion, C_1 weakening (i.e. derhoticized /ɹ/, /l/-vocalization), or syllabic C_2s, suggesting that the vowel is epenthetic, not lexicalized.

5.5 Summary of findings and conclusions

5.5.1 The vowel is phonological (not excrescent)

A comparison of the results against each diagnostic prediction is provided in Table 23.

Table 23. Diagnosis of the vowel as phonological, not excrescent.

Diagnostic	Prediction for phonological vowels	Prediction for excrescent vowels	Diagnosis of current experiment
Morpho-phonological distribution	Systematic (occurring in predictable phonological environments)	Not systematic	Phonological
Interaction with other phonological processes	Interacts with other phonological processes	Does not interact	Phonological

5.5.2 The vowel is epenthetic (not lexicalized)

A comparison of the results against each diagnostic prediction is provided in Table 24.

Table 24. Diagnosis of the vowel as epenthetic, not lexicalized.

Diagnostic	**Prediction for epenthetic vowels**	**Prediction for lexicalized vowels**	**Diagnosis of current experiment**
Environment targeted by other phonological repair strategies	Systematically present in environments avoided by way of other phonological processes.	Not systematically present in environments avoided by way of other phonological processes.	Epenthetic

5.5.3 Conclusions

In Chapters 3 and 4, I used *phonetic* diagnostics (i.e. those assessing the acoustic/phonetic properties of the vowel) to diagnose the inserted vowel as phonological. In the current chapter, I used *phonological* diagnostics (i.e. those assessing the phonological properties of vowel insertion) to diagnose the inserted vowel as phonological, and subsequently as epenthetic.

Since epenthetic and lexicalized vowels are both phonological, while excrescent vowels are articulatory, the first step was to determine whether the vowel was phonological. The results of two phonological diagnostics used in the current chapter excluded the possibility that the vowel is excrescent. After determining that the vowel was phonological, I investigated whether vowel insertion occurred in environments avoided by way of other phonological processes, and found that speakers weakened and deleted C_1 and used syllabic C_2s in tokens where a vowel was not present, indicating that vowel insertion in Scottish English is a process of phonological epenthesis.

Chapter 6:
Discussion and Conclusions

6.1 Summary of findings

In this dissertation, I presented two language production experiments which probed the phonetic and phonological properties of inserted vowels (i.e. vowels that are not present etymologically but are present in the acoustics). Since underlying (i.e. lexicalized) and non-underlying (i.e. epenthetic or excrescent) vowels can both be present in the acoustic signal but originate at different points in the derivation, I set out to improve the diagnostic criteria and process by which these vowel types are distinguished from one another. Regardless of whether experimental (e.g. acoustic or articulatory studies) or non-experimental (e.g. perception or fieldworker transcriptions) methods are employed, phonologists who have attempted to characterize inserted vowels have faced a challenge—that these vowels can appear to be quite similar to one another on the surface. I argued that the difficulty in diagnosing inserted vowels comes from diagnostics which are not informed by theory, misclassifications, and from not excluding the alternatives when making a diagnosis. In this dissertation, I sought to demonstrate how our diagnostics could be improved and how a multi-step exclusion process could be used to circumvent the challenges posed to phonologists investigating the origins of inserted vowels which arise in the acoustic signal.

In this dissertation, I asked the following questions about the inserted vowel which arises in Scottish English liquid+liquid and liquid+nasal consonant sequences:

(22) Is the vowel phonological or excrescent?

(23) If the vowel is phonological, is it lexicalized or epenthetic?

I used converging evidence from a host of independent diagnostic criteria to classify the vowel as phonological and then as epenthetic in Chapters 3, 4, and 5. In Tables 25 and 26, I summarize the diagnostics used in this dissertation, predictions for each diagnostic, and the diagnosis made for each criterion.

Table 25. Phonetic and phonological diagnostics used to classify the vowel as phonological.

Properties assessed	**Diagnostic**	**Prediction for phonological vowels**	**Prediction for excrescent vowels**	**Diagnosis**
Phonetic properties	Durational relations of adjacent gestures	Significantly negative correlation for C_1 and V. Significantly positive correlation for V and C_2. Significantly negative correlation for C_1 and C_2.	No significantly negative correlation for C_1 and V. No significantly positive correlation for V and C_2. Significantly positive correlation for C_1 and C_2.	Phonological (Chapters 3 and 4)
Phonetic properties	Bimodal distribution of durations	Bimodal.	Unimodal.	Phonological (Chapters 3 and 4)
Phonetic properties	Consonant-to-consonant duration	Duration of consonant sequence significantly lengthened in words with vowel insertion, compared to those without.	Duration of consonant sequence not significantly lengthened in words with vowel insertion, compared to those without.	Phonological (Chapters 3 and 4)
Phonetic properties	Speech Rate	The vowel will lengthen in slow speech and shorten in fast speech.	The vowel will not lengthen in slow speech and shorten in fast speech.	Phonological (Chapter 4)

Table 25 (cont'd)

Phonological properties	Morpho-phonological distribution	Systematic (occurring in predictable phonological environments).	Not systematic.	Phonological (Chapter 5)
Phonological properties	Interaction with other phonological processes	Interacts (co-occurring with processes which do not independently repair dis-preferred coda clusters and not co-occurring with processes that do).	Does not interact.	Phonological (Chapter 5)

Table 26. Phonetic and phonological diagnostics used to classify the vowel as epenthetic.

Diagnostic Type	**Diagnostic**	**Prediction for epenthetic vowels**	**Prediction for lexicalized vowels**	**Diagnosis**
Phonological properties	Widespread distribution across lexical items	Widespread distribution across lexical items, including novel contexts.	Narrow distribution across lexical items, occurring primarily in a handful of high frequency items.	Epenthetic (Chapters 3 and 4)
Phonetic properties	Durations of inserted versus underlying vowels	Significantly different.	Not significantly different.	Epenthetic (Chapter 3)
Phonological properties	Environment targeted by other phonological repair strategies	Distribution in environments which are otherwise repaired by way of separate phonological processes.	Distribution in environments which are *not* otherwise repaired by way of separate phonological processes.	Epenthetic (Chapter 5)

In the following sections, I will discuss implications and conclusions of this work.

6.2 Implications

6.2.1 Implications for Scottish English phonology

The language production experiments that constitute the basis of this dissertation are the first to experimentally investigate the function of vowel insertion in Scottish English liquid+liquid and liquid+nasal consonant sequences. Although many have regarded the vowel as epenthetic or lexicalized (see Chapter 1), this characterization has never been argued for using a detailed analysis until now. In this dissertation, I assessed the phonetic and phonological properties of the vowel against nine diagnostics and excluded the possibility that the vowel was lexicalized or excrescent. This dissertation makes the following contributions to Scottish English phonology: (1) the inserted vowel in liquid+liquid and liquid+nasal clusters is epenthetic, (2) vowel insertion occurs in coda and non-coda positions, (3) vowel insertion occurs in a very large number of lexical items, and (4) vowel insertion interacts with other processes which affect the flanking consonants.[137]

In this dissertation, I considered the possibility that the prior reports of vowel insertion regarded the process as being restricted to coda position because the distribution of vowel insertion in non-coda positions was not surveyed. In Chapter 5, I investigated the distribution of vowel insertion in coda and non-coda positions, and found that, although the vowel surfaced between consonants which would otherwise form a coda to a high degree, it also surfaced between consonants which were separated by a syllable boundary. I also considered the possibility that researchers regarding vowel insertion as lexicalized

[137] I will discuss a possible motivation for vowel insertion in coda and non-coda positions in Section 6.2.3.

might have done so because the distribution of vowel insertion had never been investigated across a wide range of lexical items. In Chapter 3, I included high- and low- frequency items, SSE and Scots words, and multimorphemic, polysyllabic tokens. The results of Chapter 3 revealed that this process applies to a considerably larger number of items than was previously thought.

Since this phenomenon involves liquid consonants (i.e. rhotics and laterals) and vowel length, and widely-discussed phenomena in Scottish English phonology include the Scottish Vowel Length Rule, derhoticization, and /l/-vocalization, this research has the potential to supplement work on these processes. In the first production experiment, vowel insertion co-occurred with rhotic allophony, and in the second production experiment, vowel insertion did not co-occur with C_1 weakening or deletion (i.e. derhoticization, /l/-vocalization, rhotic deletion or lateral deletion).

6.2.2 Implications for methodology

To classify the inserted vowel in Scottish English as epenthetic, I used a multi-step exclusion process which made use of nine diagnostics, three of which have not been used before to distinguish different inserted vowel types. In the following subsections, I will discuss each of the three novel diagnostics in turn.

6.2.2.1 Phonological vowels and transgestural coordination

In this dissertation, I used a gestural coupling framework to model how phonological and excrescent vowels should be coordinated with the flanking consonants. Shaw and Kawahara (2018) used the acoustic consequences of gestural coupling to diagnose a process of vowel devoicing as phonological deletion, rather than the result of the flanking gestures overlapping the intervening vocalic material. In a similar fashion, I used the acoustic consequences of gestural coupling to diagnose the vowel in Scottish English as phonological, rather than excrescent. This diagnostic has not yet been used to classify inserted vowels, so this is a novel application of the diagnostic. The results of Chapters 3 and 4 revealed that the consonants flanking inserted vowels (i.e. C_1-C_2) were negatively correlated, as was predicted for consonants which are not a part of a coda cluster. This diagnostic, then, served to exclude the possibility that the vowel is excrescent. However, considering the fact that not all of the results were in the predicted direction or statistically significant, it is worth noting that this diagnostic may need to be adapted in future research. For example, a duration *ratio* may better capture the greater overlap shared between the gestures in an in-phase relationship when the duration of the first gesture increases.

6.2.2.2 Excrescent vowels and speech rate[138]

In Chapter 4, I directly investigated the effects of speech rate on the phonetic properties of the inserted vowel in a second language production experiment. Hall (2003, 2006) identifies cross-linguistic patterns for how epenthetic and excrescent vowels are affected by speech rate, however, these patterns come from studies where speech rate is not directly manipulated. Hall (2003) maintains that the pattern for excrescent vowels shortening (and sometimes deleting) in fast speech and lengthening in slow speech is a "direct prediction of the gestural analysis" (p. 11) in which gestures should overlap more in fast speech. Since underlying (i.e. lexicalized or lexical) *and* non-underlying vowels (i.e. epenthetic or

138 Note that the exploratory result presented in Section 3.5.3 used a proxy for speech rate which allowed me to determine whether the vowel behaved as a phonological segment, lengthening and shortening with minor adjustments in speech rate which occur even at a normal pace.

excrescent) have the potential to shorten in fast speech, Hall's (2006) criteria cannot be used as a diagnostic to distinguish an excrescent vowel from any other type of vowel.

Using a gestural coupling framework (Goldstein et al., 2009; Nam & Saltzman, 2003; Saltzman & Byrd, 2000), I described why making generalizations about the degree to which gestures overlap at varying speech rates should be specific to the phonological environment in which the gestures are situated is necessary. Hall (2006) cites experimental evidence of greater overlap at fast speech to support the claim that excrescent vowels should shorten/delete at fast speech and lengthen in slow speech, however, the empirical evidence that Hall (2006) provides is for gestural overlap between consonants separated by a word boundary. Since vowel insertion in Scottish English occurs between consonants which would otherwise form a coda cluster, I used experimental evidence for gestural coordination of anti-phase gestures, since coda consonants have an anti-phase relationship. The three studies that I reviewed (de Jong, 2001; Durvasula & Huang, 2017; Solé, 1992) found that gestures overlapped *less* in fast speech or maintained a *stable* degree of overlap across varying speech rates—contra the predictions made by Hall (2006). Relying on theoretical considerations and experimental evidence for the coordination of gestures which have an anti-phase coupling, I developed a criterion for how excrescent vowel length should be affected by speaking rate. This criterion stated that, in coda clusters, excrescent vowels should *not* lengthen in slow speech and shorten in fast speech, since anti-phase gestures should not be more or less overlapped as a result of varying speaking rate, since they are produced serially. This created a testable diagnostic—that excrescent vowels should *not* behave like phonological vowels, lengthening in slow speech and shortening in fast speech, as Hall (2006) argues. This diagnostic, then, is the first to make predictions for how the duration of excrescent and phonological vowels which surface between consonants which would otherwise form a coda cluster should be affected by changes in speaking rate.[139]

6.2.2.3 The duration of underlying and epenthetic vowels

As discussed in Chapter 2, many acoustic studies of vowel epenthesis have found that the vowel is often shorter than underlying vowels. Based on these cross-linguistic findings, I used this as a diagnostic to classify the vowel as epenthetic, rather than lexicalized (i.e. underlying), after having established that the vowel was phonological, not excrescent. Since *both* epenthetic and excrescent vowels can be shorter than underlying vowels, I argued that this diagnostic would be better used to distinguish between lexicalized (i.e. vowels not present etymologically but fossilized in the underlying representation over time) and epenthetic vowels, since lexicalized vowels should be the same duration as underlying vowels (i.e. those present etymologically) and epenthetic vowels should not. I anticipated that the vowel in Scottish English would be shorter than underlying vowels in similar environments if it were epenthetic, which was found in Chapter 3.

One possible explanation for this difference might be that, because the epenthetic vowel is optional, speakers may have multiple surface representations competing while planning their production. Under a cascaded activation framework, multiple, competing representations pass activation to later processing stages, regardless of whether the selection has been made at any given stage (Goldrick & Blumstein, 2006). For example, morphs may pass activation to phonological representations and articulatory planning

[139] I encourage other researchers to use this diagnostic in future studies in order to determine its reliability.

before a morph has been chosen (Bermúdez-Otero, 2010). Although most of the production literature providing support for this framework has focused on morpheme-level activation, McMillan and Corley (2010) found articulatory variability in the production of tongue twisters that was attenuated by the number of phonological features that the competing representations differed by. This attenuation suggests that partial activation of multiple phonemes may influence production. Assuming that speakers have multiple representations of words which undergo optional vowel epenthesis, it may be the case that the competition between the possible production types results in epenthetic vowels being shorter than underlying vowels which are not optionally inserted, however, more research on phoneme-level activation is needed to support this kind of proposal.

6.2.3 Implications for the theory

In Chapter 1, I highlighted how cross-linguistic patterns for the different types of inserted vowels have the power to impact the theoretical framework. A vowel which is classified as excrescent but that displays partial visibility to phonology, for example, has the power to challenge a modular feedforward framework (i.e. where the phonology module feeds the phonetics module and not vice versa) since processes which *originate* in the phonetics module cannot access the phonology module. Although it is entirely possible that vowels which display some exceptional properties may require a different phonological framework to generate these patterns, it is also possible that typological 'exceptions' are the result of methodological limitations. Improving upon our existing diagnostics and utilizing them to exclude alternatives can aid in updating the typology of inserted vowels—their phonetic and phonological properties, and the environments in which they surface—and circumvent inadvertently creating typological outliers that a modular feedforward framework is not equipped to generate.[140]

This dissertation revealed phonetic and phonological properties of epenthetic vowels which can be used to supplement the existing typology of vowel epenthesis. One such property was that the epenthetic vowel lengthened in slow speech and shortened in fast speech.[141] Another was that the epenthetic vowel displayed the expected acoustic consequences of a gestural coupling model predicted for underlying segments. Lastly, the epenthetic vowel had a significantly shorter duration than underlying vowels in similar environments. Additionally, although the patterns observed from vowel insertion in Scottish English can only directly contribute to the typology of epenthetic vowels, this dissertation generated theoretical predictions for the phonetic properties of excrescent vowels that can be used by future researchers.

Furthermore, this dissertation presented a case of vowel insertion which surfaces in coda clusters and in non-coda consonant sequences. Although liquid+liquid heterosyllabic coda+onset sequences (e.g. *ear.ly*) are sonority plateaus and liquid+nasal heterosyllabic coda+onset sequences (e.g. *mur.mur*) have minimally dispersed sonority, they are permittable sequences in English (Yip, 1991). Vowel epenthesis in English is often found in tautosyllabic clusters, while consonant epenthesis is found more often than vowel epenthesis across syllable boundaries (Warner & Weber, 2002). In Chapter 5, I discussed the possibility that vowel epenthesis in this environment may serve to repair violations of the Syllable Contact Law (Murray & Vennemann, 1983; Vennemann, 1988), which

140 Note that the novel exclusion process developed in this dissertation was developed specifically to allow one to discern between different inserted vowels which exist in a modular feedforward framework.
141 To my knowledge, only one other study (Shitaw, 2014) has investigated the effects of speech rate on epenthetic vowel length in coda clusters. Authors found the same pattern identified in the current study.

dictates that the coda preceding the syllable boundary should be more sonorous than the onset following the syllable boundary, but only for the *-rl* consonant sequence, since it forms a sonority plateau. The other sequences (i.e. liquid+nasal), however, do not violate the Syllable Contact Law, as liquids are more sonorous than nasals. The tendency for inserted vowels to appear in liquid+liquid and liquid+nasal sequences more in coda positions, but also across syllable boundaries, has also been found in Irish English (Sell, 2012), so vowel insertion in this environment may have implications for the phonology of Celtic dialects of English.

The distribution of vowel insertion in both environments, but to a higher degree in coda clusters, might reflect two separate processes of epenthesis. Durvasula and Liter (2020) found that, for phonotactic sequence patterns, learners can have multiple generalizations for sequences of segments. In the current study, it is possible that speakers have two generalizations for liquid+liquid and liquid+nasal sequences—one specifically for consonants occupying the coda position, and another for the same consonant sequences, regardless of their syllable position. If a segment sequence generalization was employed by speakers, it would target consonants which occupy codas, but also those that do not. If a coda-specific generalization was also available to speakers, it would result speakers having two generalizations for coda sequences (i.e. one for consonants in a coda and one for consonant sequences in any syllable position), but only one generalization for non-coda sequences. This would result in speakers producing epenthetic vowels more often in coda clusters than in consonants separated by a syllable boundary.

6.3 Conclusion

One of the many important questions central to the field of phonology is at what level of representation do speech sounds observable in the acoustics originate from? In this dissertation I operated under a modular feedforward framework, in which phonology and phonetics are distinct modules, and phonology "feeds" phonetics in an asymmetrical fashion (Pierrehumbert, 2002). To be able to adequately model the different inserted vowels that have been found cross-linguistically, I specifically used a featural-gestural framework (Zsiga, 1994, 1997). Since inserted vowels are present in the acoustics, but originate at different stages of representation, utilizing *both* the phonetic (e.g. duration) and phonological properties (i.e. distribution, interaction with other phonological processes) of inserted vowels can help ensure that the theoretical predictions are borne out in more ways than one. Using two language production experiments, novel diagnostics, a novel exclusion process, and an understudied case of vowel insertion in Scottish English, this dissertation compiled two sets of phonetic and phonological diagnostic criteria which were used to determine where the inserted vowel (i.e. lexicalized, epenthetic, excrescent) originates in the derivation (i.e. underlying representation, surface representation, or gestural alignment, respectively) via a process of elimination. The diagnostics, exclusion process, language-specific results, and patterns for epenthetic and excrescent vowels presented in this dissertation serve to supplement Scottish English phonology, to improve upon the methodology available to phonologists investigating inserted vowels, and to contribute to the typology of inserted vowels.

APPENDICES

APPENDIX A:
Recruitment Flyer for Experiment 1

Hello all,

My name is Kaylin Smith and I am a Visiting Postgraduate Researcher in the PPLS department. I am conducting a study which involves recording native Scottish English speakers reading aloud from a word list. Participants will be compensated £# and recording sessions should be about 1 hour long. Testing will occur on the main campus M-F (although Saturday and Sunday times may be arranged) June 8th- June 23rd.

All participants must:

1) Be native Scots and/or English speakers
2) Have lived in Scotland for the past 10 years
3) Have been born in and currently live in the Central Belt region
4) Have no speech or hearing disorders and have normal or corrected-to-normal vision
5) Be older than 18 years of age

If you would like to participate in this research, please contact me at #### and provide the following information:

1) Your full name, email address, and mobile phone number
2) Age
3) How many years you have lived in Scotland
4) Where you were born and where you live now
5) Dates and times of availability (please provide at least two options)

Additionally, please feel free to forward this email to people you feel may be qualified and interested.

Thank you!

Kaylin Smith
Visiting Postgraduate Researcher
Philosophy, Psychology, and Language Sciences
The University of Edinburgh

APPENDIX B:
Experiment 1 Stimuli

Stimuli for Block 1.

a) Experimental items

Consonant Sequence	Bare word (x)	X + [PRES PROG/ GER] *ing*	x + [AGENT] *er*	x + [PAST] *ed*	x +[PLU] *s*
lm	calm	calming	calmer	calmed	calms
lm	culm	culming	culmer	culmed	culms
lm	dishelm	dishelming	dishelmer	dishelmed	dishelms
lm	film	filming	filmer	filmed	films
lm	helm	helming	helmer	helmed	helms
lm	microfilm	microfilming	microfilmer	microfilmed	microfilms
lm	overwhelm	overwhelming	overwhelmer	overwhelmed	overwhelms
lm	psalm	psalming	psalmer	psalmed	psalms
lm	underwhelm	underwhelming	underwhelmer	underwhelmed	underwhelms
rl	curl	curling	curler	curled	curls
rl	hurl	hurling	hurler	hurled	hurls
rl	snarl	snarling	snarler	snarled	snarls
rl	twirl	twirling	twirler	twirled	twirls
rl	unfurl	unfurling	unfurler	unfurled	unfurls
rl	whirl	whirling	whirler	whirled	whirls
rm	affirm	affirming	affirmer	affirmed	affirms
rm	alarm	alarming	alarmer	alarmed	alarms
rm	arm	arming	armer	armed	arms
rm	charm	charming	charmer	charmed	charms
rm	confirm	confirming	confirmer	confirmed	confirms
rm	dorm	dorming	dormer	dormed	dorms
rm	farm	farming	farmer	farmed	farms

APPENDIX B (cont'd):
Experiment 1 Stimuli

rm	firm	firming	firmer	firmed	firms
rm	form	forming	former	formed	forms
rm	outperform	outperforming	outperformer	outperformed	outperforms
rm	perform	performing	performer	performed	performs
rm	squirm	squirming	squirmer	squirmed	squirms
rm	swarm	swarming	swarmer	swarmed	swarms
rm	term	terming	termer	termed	terms
rm	warm	warming	warmer	warmed	warms
rm	worm	worming	wormer	wormed	worms
rn	burn	burning	burner	burned	burns
rn	earn	earning	earner	earned	earns
rn	learn	learning	learner	learned	learns
rn	mourn	mourning	mourner	mourned	mourns
rn	turn	turning	turner	turned	turns
rn	warn	warming	warner	warned	warns
rn	yarn	yarning	yarner	yarned	yarns

APPENDIX B (cont'd):
Experiment 1 Stimuli

b) Control items.

Consonant Sequence	Bare word (x)	x + [PRES PROG/ GER] *ing*	x + [AGENT] *er*	x + [PAST] *ed*	x +[PLU] *s*
ng	sing	singing	singer	singed	sings
ng	ring	ringing	ringer	ringed	rings
ng	long	longing	longer	longed	longs
ng	wrong	wronging	wronger	wronged	wrongs
inger	finger	fingering	fingerer	fingered	fingers
X	skate	skating	skater	skated	skates
nt	plant	planting	planter	planted	plants
lk	sulk	sulking	sulker	sulked	sulks
lp	help	helping	helper	helped	helps
lp	gulp	gulping	gulper	gulped	gulps
rt	blurt	blurting	blurter	blurted	blurts
X	stroll	strolling	stroller	strolled	strolls
X	dream	dreaming	dreamer	dreamed	dreams
X	drool	drooling	drooler	drooled	drools
X	glide	gliding	glider	glided	glides
X	smell	smelling	smeller	smelled	smells
X	skip	skipping	skipper	skipped	skips
st	cast	casting	caster	casted	casts
st	list	listing	lister	listed	lists
st	test	testing	tester	tested	tests
st	post	posting	poster	posted	posts
st	dust	dusting	duster	dusted	dusts
nt	point	pointing	pointer	pointed	points

APPENDIX B (cont'd):
Experiment 1 Stimuli

mpt	tempt	tempting	tempter	tempted	tempts
pt	script	scripting	scripter	scripted	scripts
sk	whisk	whisking	whisker	whisked	whisks
sp	lisp	lisping	lisper	lisped	lisps
sp	grasp	grasping	grasper	grasped	grasps
mp	stamp	stamping	stamper	stamped	stamps

c) Words with underlying vowels in similar environments.

1. forum
2. eardrum
3. fulcrum
4. vellum
5. column
6. skellum
7. Callum

APPENDIX B (cont'd):
Experiment 1 Stimuli

Stimuli used in Block 2.

a) Experimental items (with consonant sequences in coda position).

1. folm
2. skilm
3. skoolm
4. stolm
5. whalm
6. helm
7. colm
8. McColm
9. Inchcolm
10. Christholm
11. Chrisholm
12. Denholm
13. Broomholm
14. langholm
15. Dachalm
16. Leckmelm
17. ferm
18. aarm
19. barm
20. chirm
21. corm
22. gurm
23. karm
24. nyarm
25. sherm
26. thairm
27. yirm
28. yarm
29. wirm
30. wairm
31. berm
32. alairm
33. erm
34. Glenarm
35. Kailwirm

APPENDIX B (cont'd):
Experiment 1 Stimuli

36. lew-wairm
37. hurl
38. nairn
39. bairn
40. morn
41. corn
42. burn
43. warld
44. birlt
45. dairn
46. murn
47. kirn
48. pirn
49. curn
50. airn
51. Coburn
52. girn
53. hern
54. lairn
55. Broxburn
56. Kinghorn
57. forfairn
58. skirl

b) Experimental items (with consonant sequences in non-coda position).

1. Helmsdale
2. Talmine
3. Balmoral
4. Chalmers
5. Malcolmson
6. Pilmuir
7. Gilmour
8. Kilmer
9. contermashious
10. Locharmoss
11. Carmyllie
12. fermer
13. Carmyle
14. Permeision

APPENDIX B (cont'd):
Experiment 1 Stimuli

15. Barmulloch
16. Yarmouth
17. airmie
18. barmie
19. betterment
20. gormless
21. evermair
22. flichtermoose
23. gormaw
24. infirmary
25. ivermair
26. murmle
27. Westermill
28. Fostermeadow
29. Nethermuir
30. Lammermuir
31. Carmichael
32. MacCormick
33. Kilninver
34. Skelmorlie
35. Caerlaverok
36. Thirlstane
37. Carloway
38. Dirleton
39. Aberlemno
40. Burleigh
41. Milngavie
42. daurna
43. warna
44. arna
45. girned
46. bairnie
47. burnewin
48. burnie
49. capernoitie
50. carnaptious
51. Cummernaud
52. herness
53. hornie
54. lairnit

APPENDIX B (cont'd):
Experiment 1 Stimuli

55. efternuin
56. fernent
57. moarnin
58. Barlinnie
59. amurny
60. barnie
61. birlin
62. willnae
63. pernicketie
64. Abernethy
65. Carnoustie
66. Kilninver

c) Control items (other attested consonant sequences that form onset and coda clusters).

1. skelp
2. braw
3. radge
4. shan
5. scran
6. skite
7. dreich
8. rank
9. spraff
10. craic
11. cowk
12. gran
13. kirk
14. scheme
15. skint
16. wheesht
17. wynd
18. smirr
19. skelf
20. schame
21. richt
22. plunk

23. pairk
24. maist
25. clowt
26. clype
27. clart
28. gang
29. reeking
30. blootered
31. scunner
32. sleekit
33. steamin
34. strang
35. stramash
36. stank
37. stane
38. splore
39. midge
40. gloamin
41. glaikit
42. ahint
43. ayont
44. bing

d) Control items (other attested consonant sequences in non-coda or word-medial positions).

1. blether
2. bucket
3. munter
4. bampot
5. pieing
6. bowfin
7. awfy
8. chankin
9. minted
10. wisnae
11. unco
12. tumshie
13. tourie

APPENDIX B (cont'd):
Experiment 1 Stimuli

14. stairtit
15. sodger
16. sempie
17. muckin
18. isnae
19. disnae
20. cludgie
21. corbie
22. inby
23. forby
24. wisnae
25. lichtnin
26. banger

APPENDIX C:
Demographic Questionnaire for Experiment 1

Please read aloud each question number and state your answer aloud. You will be recorded while doing so. When you have finished, please notify the researcher.

1. Please state aloud your age and gender.
2. Where were you born? (City, County, Country)
3. Have you ever lived in a country other than Scotland? If yes-- where? For how long?
4. What towns/villages/cities have you been a resident of in Scotland?
5. Do you speak Scots? Do you speak Scottish Standard English?
6. Who do you speak Scots with? Who do you speak Scottish Standard English with?
7. Do you speak Scots often?
8. Where were your parents born?
9. Are your parents native English speakers? What languages do your parents speak?
10. Have you ever lived with someone whose native language was not English? What about someone who spoke with a non- Scottish dialect of English?
11. What is your occupation? How long have you been in that occupation? What are some daily responsibilities in this profession?
12. What about your hobbies?
13. What education have you had? Tell me about your experiences in school.

APPENDIX D:
Speaker Demographics for Experiment 1

ID	Gender	Age	City of birth	Occupation
1	F	44	Broxburn	Computing support UoE
2	*Removed from analysis*			
3	M	52	Edinburgh	States operation support UoE
4	F	60	St Fillans	None (admin; box office manager)
5	M	71	Edinburgh	Retired (civil servant)
6	*Removed from analysis*			
7	M	68	Crossgates	Retired (chartered accountant)
8	F	64	Edinburgh	Retired (head primary teacher)
9	*Removed from analysis*			
10	F	72	Strathaven	Retired (primary school teacher)
11	*Removed from analysis*			
12	M	21	Edinburgh	Student
13	F	58	Edinburgh	Retired
14	F	36	Edinburgh	Communications officer
15	*Removed from analysis*			
16	F	46	Alexandria	Full time student
17	F	51	Falkirk	Part time examinations individuator
18	F	22	Edinburgh	Recently graduated; working in Operations
19	*Removed from analysis*			
20	M	93	Motherwell	Retired; general medical practice
21	*Removed from analysis*			
22	*Removed from analysis*			
23	F	63	Edinburgh	Retired; teacher, jeweler
24	M	80	Edinburgh	Retired; civil servant
25	M	77	Rothglen	Retired; solicitor- law apprentice until retirement
26	*Removed from analysis*			
27	M	65	Glasgow	Retired; civil servant; senior manager

APPENDIX E:
Recruitment Flyer for Experiment 2

Hello all,

My name is Kaylin Smith and I am a Visiting Postgraduate Researcher in the PPLS department. I am conducting a study which involves recording native Scottish English speakers reading aloud from a word list. Participants will be compensated £# and recording sessions should be 45 minutes-1 hour long. Testing will occur at The University of Edinburgh main campus and the Mitchell Library in Glasgow Monday-Friday May 27th-June 12th.

All participants must:

1) Be a Scots and/or English speaker, who doesn't fluently speak another language besides Scots or English
2) Have been born in and living in Scotland for the past 10 years
3) Have no speech disorders
4) Have normal or corrected-to-normal vision and hearing
5) Be 45+ years old
6) Be male

If you would like to participate in this research, please contact me at #### or #### and provide the following information:

1) Your full name, email address, and mobile phone number
2) Age
3) How many years you have lived in Scotland
4) Where you were born and where you live now
5) Dates and times of availability (please provide at least two options)

Additionally, please feel free to forward this email to people you feel may be qualified and interested.

Thank you!
Kaylin Smith
Visiting Postgraduate Researcher
Philosophy, Psychology, and Language Sciences
The University of Edinburgh

APPENDIX F:
Experiment 2 Stimuli

Note. Target words are bolded.

1. I told Pearl we'd see the **film** another time.
2. The captain stood at the **helm** of the boat.
3. He could imagine it in some other **realm**.
4. Lightning struck inches from where she planted an **elm**.
5. The ecologist inspected the **culm** of the prairie grass.
6. You need to fill out that **form** by tomorrow.
7. I hope I'll get better marks next **term**.
8. Driving through the countryside, we spotted Mr. Waterston's **farm**.
9. On the horizon, a big **storm** approached the beach.
10. I'm wondering if I should get a **perm**.
11. The caber toss is inferior to the haggis **hurl**.
12. My aunt even wears **pearl** earrings to the cinema.
13. Our dog chews on everything, so he's called **Gnarl**.
14. The castle is owned by the **Earl** of Airlie.
15. The busiest holiday in the **world** is Boxing Day.
16. Our actual house isn't as sturdy as dad's **barn**.
17. I'm worried I'll get another **burn** if I'm distracted.
18. Laverne didn't realize that it was her **turn** again.
19. My gran spends a fortune on **yarn** each week.
20. While trimming roses, she was pricked by a **thorn**.

APPENDIX G:
Demographic Questionnaire for Experiment 2

Please write your answers for each question below and let the experimenter know if you have any questions.

1. Age: _____________
2. Gender:____________________
3. Place of birth: CITY __________________ COUNTRY ________________
4. What is the city/town/village you currently stay in? _______________________
5. Have you ever lived in a country other than Scotland? Please circle one: YES/NO
   a) If YES, where did you live?
      1. CITY: __________________ COUNTRY: _____________________
      2. CITY: __________________ COUNTRY: _____________________
      3. CITY: __________________ COUNTRY: _____________________
   b) If YES, what were the dates you lived there?
      1. ARRIVAL: (month/year) ______ DEPARTURE: (month/year)_______
      2. ARRIVAL: (month/year) ______ DEPARTURE: (month/year)_______
      3. ARRIVAL: (month/year) ______ DEPARTURE: (month/year)_______

6. Please list all of the places you have lived in Scotland and for how long (including your current residence):
   PLACE: _________________LENGTH (months/years):____________________
   PLACE: _________________LENGTH (months/years):____________________
   PLACE: _________________LENGTH (months/years):____________________
   PLACE: _________________LENGTH (months/years):____________________
   PLACE: _________________LENGTH (months/years):____________________
   PLACE: _________________LENGTH (months/years):____________________

7. Are you fluent in any languages other than Scots or English? YES / NO
   a) If YES, which one(s)?
      ____

8. Was English the primary language spoken in your house when you were a child? YES / NO
   a) If NO, what other languages were spoken?
      ____

9. Have you ever lived with someone whose native language was not English? YES / NO
   a) If YES, what language did/do they speak?
      ____

b) If YES, what dates did you live with that person? (e.g., June 2012 until May 2019)

____

10. Have you ever lived with someone whose native language was English but spoke with a different dialect of English (not with a Scottish dialect, but an Australian English or American English dialect, for example)? YES / NO

a) If YES, please specify what dialect of English they used:

____

b) If YES, how long did you live with them? (months/years)____

11. Are you currently employed? YES / NO

a) If YES, please state your current occupation:

____

b) If YES, how long have you been in your current post:

____

12. Are you retired? YES / NO

a) If YES, what was/were your previous occupation(s) and for how long were you employed in that/those position(s)?
(example: "teacher, 4 years; principal of school, 5 years; nurse, 2 years")

____

____

13. Are you unemployed but not retired? YES / NO

a) If YES, what was/were your previous occupation(s) and for how long were you employed for each?
(example: "teacher, 4 years; principal of school, 5 years; nurse, 2 years")

____

____

14. What is the highest level of education you have had? Please indicate this by placing a **circle** around an option from the following table. *Note.* If none of the choices in the table describe your education level, please write in your answer in (a):

a) Other:

____

APPENDIX G (cont'd):
Experiment 2 Demographic Questionnaire

Scottish Credit and Qualification Framework (SCQF) (The Scottish Credit and Qualifications Framework, n.d.)

SCQF levels	SQA National Units, courses and group awards	Higher Education (HE) qualifications	Scottish Vocational Qualifications	SCQF levels
12		Doctorate		12
11		Masters	SVQ 5	11
10		Honours degree		10
9		Ordinary degree		9
8		HND	SVQ 4	8
		Diploma of HE		
7	Advanced Higher	HNC		7
		Certificate of HE		
6	Higher		SVQ 3	6
5	Intermediate 2/		SVQ 2	5
	Credit S Grade			
4	Intermediate 1/ General S Grade		SVQ 1	4
3	Access 3/ Foundation S Grade			3
2	Access 2			2
1	Access 1			1

15. Did you go to a public school for your education? YES / NO / N/A
16. What were your parents'/guardians' occupations? (e.g., "Mom, nurse; Dad, farmer", etc.)
____
____

17. What are your hobbies?______________________________
____

18. Please rate how often you write OR say each of these words:
0= never, 1= rarely, 2= sometimes, 3=fairly often, 4=very often
    a) Heid ______
    b) Hen ______
    c) Hunner ______
    d) Intae ______
    e) Isnae ______
    f) Muntered ______
    g) Nae ______
    h) Reekin ______
    i) Wisnae ______
    j) Yersel ______
    k) Aff ______
    l) Doon ______

19. If you had a choice of which word to use, which would you prefer (regardless of your audience)? Please circle your answer.
    a) Shouldn't **OR** shuidna(e)
    b) Right **OR** richt
    c) Why **OR** hou/whit wey
    d) Don't **OR** dinna(e)
    e) Know **OR** ken
    f) Another **OR** anither
    g) Remember **OR** mynd

20. Do you self-identify as a speaker of Scots, Scottish Standard English, or both?
Please note. By "Scots", we mean "Broad Scots", not "Scots Gaelic".
____
    a) If you identify as a Scots speaker, what regional dialect do you use? If unknown, please write "unknown":
    ____

APPENDIX G (cont'd):
Experiment 2 Demographic Questionnaire

21. Please place an "X" on the line below to indicate what best describes your speech. Please refer to the following key before making a selection:

<u>Key for question 21:</u>
1 = I only speak Scots
2 = I speak Scots 85% of the time, English 15% of the time
3 = I speak Scots 65% of the time, English 35% of the time
4 = I speak Scots 50% of the time, English 50% of the time
5 = I speak English 65% of the time, Scots 35% of the time
6 = I speak English 85% of the time, Scots 15% of the time
7 = I only speak English

only Scots			Both			only English
<-->						
1	2	3	4	5	6	7

APPENDIX H:
Speaker Demographics for Experiment 2

ID	Gender	Age	City of birth	Occupation
1	M	73	Falkirk	Retired (Police Officer)
2	M	67	Glasgow	Retired (Computers)
3	M	66	Renfrow	Retired (Social Services)
4	M	72	Glasgow	Retired (Author, Artist)
5	M	46	Coatbridge	Writer/Communications Consultant
6	M	73	Glasgow	Retired (Sales/Marketing Manager)
7	*Removed from analysis*			
8	M	50	Glasgow	Client Services at Royal Conservatory
9	M	73	Edinburgh	Retired (Civil Servant)
10	*Removed from analysis*			
11	*Removed from analysis*			
12	M	82	Edinburgh	Retired (Civil Servant)
13	M	55	Edinburgh	HR Professional
14	M	72	Edinburgh	Retired (Civil Servant)
15	*Removed from analysis*			
16	*Removed from analysis*			
17	M	69	Edinburgh	Bus Garage Labourer, Postman
18	M	49	Glasgow	Student at Open University
19	*Removed from analysis*			
20	*Removed from analysis*			
21	*Removed from analysis*			
22	M	70	Crossgates	Retired (Chartered Accountant)
23	*Removed from analysis*			
24	*Removed from analysis*			
25	*Removed from analysis*			
26	*Removed from analysis*			
27	*Removed from analysis*			
28	*Removed from analysis*			
29	*Removed from analysis*			

REFERENCES

REFERENCES

Aitken, A. (1984). Scots and English in Scotland (1984). In P. Trudgill (Ed.), *Language in the British Isles* (pp. 517–532). Cambridge University Press.

Aitken, A. J. (1981). The Scottish Vowel Length Rule. In *So Many People, Longages, and Tongues, Edinburgh: Middle English Dialect Project* (pp. 131–157).

Aitken, A. J. (1979). Scottish speech: A historical view with special reference to the Standard English of Scotland. In Adam J. Aitken & T. McArthur (Eds.), *The Languages of Scotland* (pp. 85–119). Chambers.

Bates, D., Maechler, M., Bolker, B., Walker, S., Haubo, R., Christensen, B., Singmann, H., Dai, B., Grothendieck, G., Green, P., & Bolker, B. (2015). *Package "lme4."*

Bellik, J. (2018). An acoustic study of vowel intrusion in Turkish onset clusters. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *9*(1), 16. https://doi.org/10.5334/labphon.112

Bermúdez-Otero, R. (2010). Currently available data on English t / d-deletion fail to refute the classical modular feedforward architecture of phonology. *The 18th Manchester Phonology Meeting*, *1995*, 1–16. www.bermudez-otero.com/18mfm.pdf

Blevins, J., & Pawley, A. (2010). Typological implications of Kalam predictable vowels. *Phonology*, *27*(01), 1–44. https://doi.org/10.1017/S0952675710000023

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9).

Boersma, Paul, & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, *32*(1), 45–86. https://doi.org/10.1162/002438901554586

Borowsky, T. (1986). *Topics in English phonology.* University of Massachusetts.

Bosch, A., & Jong, K. De. (1997). *The prosody of Barra Gaelic epenthetic vowels*. *27*(1).

Brinton, L. J., & Traugott, E. C. (2005). *Lexicalization and language change*. Cambridge University Press.

Browman, C., & Goldstein, L. (1990). Tiers in Articulatory Phonology, with Some Implications for Casual Speech. *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, 341–376.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview.

Phonetica, *49*(3–4), 155–180. https://doi.org/10.1159/000261913

Browman, C. P., Goldstein, L., Saltzman, E., & Smith, C. (1986). GEST: A computational model for speech production using dynamically defined articulatory gestures. *The Journal of the Acoustical Society of America*, *80*(S1), S97–S97. https://doi.org/10.1121/1.2024067

Bürki, A., Fougeron, C., & Gendrot, C. (2007). On the categorical nature of the process involved in schwa elision in French. *International Speech Communication Association - 8th Annual Conference of the International Speech Communication Association, Interspeech 2007*, *3*, 2209–2212.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics*, *24*(2), 209–244. https://doi.org/10.1006/JPHO.1996.0012

Byrd, D., & Tan, C. C. (1996). Saying consonant clusters quickly. *Journal of Phonetics*, *24*(2), 263–282. https://doi.org/10.1006/jpho.1996.0014

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, *58*(3), 129–162. https://doi.org/10.1159/000056196

Chomsky, N., & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of Linguistics*, *1*(2), 97–138. https://doi.org/10.1017/S0022226700001134

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English.* MIT Press. https://pdfs.semanticscholar.org/2ff6/0c8f2df9a57b5dfba38e9be6d6401d88a1e9.pdf

Clements, G. (1990). The role of the sonority cycle in core syllabification. *Papers in Laboratory Phonology: Volume 1, Between the Grammar and Physics of Speech*, *1*, 283.

Clements, G. N. (1992). Phonological primes: Gestures or features? *Phonetica*, *49*(3–4), 181–193. http://conf.ling.cornell.edu/plab/sites/all/themes/plabzen/papers/wpcpl7-Clements.pdf

Coleman, J. (1999). The Nature of Vocoids Associated With Syllabic Consonants in Tashlhiyt Berber. *Proceedings of the XIVth International Congress of Phonetic Sciences*, 735–738.

D'Apolito, S., & Fivela, B. G. (2016). *Investigating Speech Gesture Coordination in the Production of L2 Consonant Clusters: the Added Value of Electromagnetic Articulography*. Conference Proceedings. ICT for Language Learning.

Davidson, L. (2003). *The Atoms of Phonological Representation: Gestures, Coordination and Perceptual Features in Consonant Cluster Phonotactics*.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics*, *34*, 104–137. https://doi.org/10.1016/j.wocn.2005.03.004

de Jong, K. J. (2001). Rate-Induced Resyllabification Revisited. *Language and Speech*, *44*(2), 197–216.

Dieth, E. (1932). *A Grammar of the Buchan dialect*. Heffer.

Dupoux, E., Hirose, Y., Kakehi, K., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1568–1578. https://doi.org/10.1037/0096-1523.25.6.1568

Durvasula, K. (2009). *Understanding nasality*. University of Delaware.

Durvasula, K., & Huang, H. H. (2017). Word-internal “ambisyllabic” consonants are not multiply-linked in American English. *Language Sciences*, *62*, 17–36. https://doi.org/10.1016/j.langsci.2017.03.002

Durvasula, K., & Liter, A. (2020). There is a simplicity bias when generalizing from ambiguous data. *Phonology*, *850*, 1–64.

Gafos, A. (2002). A Grammar of Gestural Coordination. *Natural Language and Linguistic Theory*, *20*(2), 269–337. https://doi.org/10.1023/A:1014942312445

Gallois, C., Ogay, T., & Giles, H. (2005). Communication Accommodation Theory:a look back and a look ahead. In W. B. Gudykunst (Ed.), *Theorizing about intercultural communication* (pp. 121–148). Sage.

Gartenberg, R. D. (1984). *An electropalatographic investigation of allophonic variation in English /l/ articulations* (No. 4; Phonetics Laboratory University of Reading Work in Progress).

Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a Theory of Interpersonal Accommodation through Language : Some Canadian Data Author (s): Howard Giles , Donald M . Taylor and Richard Bourhis Published by : Cambridge University Press Stable URL : http://www.jstor.org/stable/4166723. *Language in Society*, *2*(2), 177–192.

Goldrick, M., & Blumstein, S. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes*, *21*(6), 649–683. https://doi.org/10.1080/01690960500181332

Goldstein, L., Nam, H., Saltzman, E., & Chitoran, I. (2009). Coupled oscillator model of speech timing and syllable structure Coupled Oscillator Planning Model of Speech

Timing and Syllable Structure. *Frontiers in Phonetics and Speech Science*, 239–250.

Gouskova, M., & Hall, N. (2009). Acoustics of epenthetic vowels in Lebanese Arabic. *Phonological Argumentation: Essays on Evidence and Motivation*, 203–225.

Hall, N. (2003). *Gestures and segments: Vowel intrusion as overlap*. University of Massachusetts, Amherst, MA.

Hall, N. (2006). Cross-linguistic patterns of vowel intrusion. In *Phonology* (Vol. 23, Issue 3, pp. 387–429). https://doi.org/10.1017/S0952675706000996

Hall, N. (2011). Vowel Epenthesis. *The Blackwell Companion to Phonology*, 1–21. https://doi.org/10.1002/9781444335262.wbctp0067

Hammond, M., Warner, N., Davis, A., Carnie, A., Archangeli, D., & Fisher, M. (2014). Vowel insertion in Scottish Gaelic. In *Phonology* (Vol. 31, Issue 01). https://doi.org/10.1017/S0952675714000050

Harms, R. T. (1976). The segmentalization of Finnish "nonrules." *Texas Linguistic Forum, Vol. 5*, 73–88.

Heath, J. (1987). *Ablaut and ambiguity: Phonology of a Moroccan Arabic dialect*. SUNY Press.

Heselwood, B., Shitaw, A., Ghummed, A., & Plug, L. (2015). *Epenthetic and Excrescent Vowels in Stop Sequences in Tripolitanian Libyan Arabic*. 10–14.

Jannedy, S. (1994). Rate effects on German unstressed syllables. *OSU Working Papers in Linguistics*, *44*, 105–124.

Jauriberry, T., Sock, R., & Hamm, A. (2015). Phonetic Variation in Standard Scottish English : Rhotics in Dundee. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.

Johnston, P. (1985). The rise and fall of the Morningside/Kelvinside accent. In M. Gorlach (Ed.), *Focus on Scotland* (pp. 37–56). John Benjamins Publishing Company.

Johnston, P. (1997). Older Scots phonology and its regional variation. In *The Edinburgh History of the Scots Language* (pp. 47–111).

Johnston, P., & Speitel, H. (1983). *A sociolinguistic investigation of Edinburgh speech* (C/00/23/0023/1; Social Science Research Council End of Grant Report).

Jones, C. (2002). *The English Language in Scotland: An Introduction to Scots*. Tuckwell Press.

Kahn, D. (1980). *Syllable-Based Generalizations*. Garland Publishers.

Kawahara, S. (2011). Experimental Approaches in Theoretical Phonology. *The Blackwell Companion to Phonology*, 1–21. https://doi.org/10.1002/9781444335262.wbctp0096

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Blackwell Publishers.

Kreitman, R. (2006). Cluster buster: A typology of onset clusters. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *42*(1), 163–179.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lawson, E., Scobbie, J. M., Stuart-Smith, J. (2014). A Socio-Articulatory Study of Scottish Rhoticity. In *Sociolinguistics in Scotland* (pp. 53–78). Palgrave MacMillan.

Lawson, E., Scobbie, J. M., & Stuart-Smith, J. (2013). Bunched /r/ promotes vowel merger to schwar: An ultrasound tongue imaging study of Scottish sociophonetic variation. *Journal of Phonetics*, *41*(3–4), 198–210. https://doi.org/10.1016/j.wocn.2013.01.004

Lawson, E., Stuart-Smith, J., & Scobbie, J. (2008). Articulatory insights into language variation and change : preliminary findings from an ultrasound study of derhoticization in Scottish English Pennsylvania Working Papers in Linguistics, 14 (2). pp. 102-110. ISSN Deposited on: 14 February 2011. *N University of Pennsylvania Working Papers in Linguistics*, *14*(February), 102–110.

Levin, J. (1987). Between epenthetic and excrescent vowels. *Proceedings of the 6th West Coast Conference on Formal Linguistics*, 187–202.

Levin, Juliette. (1985). *A Metrical Theory of Syllabicity* [Massachusetts Institute of Technology].

Löfqvist, A., & Gracco, V. L. (1999). Interarticulator programming in VCV sequences: Lip and tongue movements. *The Journal of the Acoustical Society of America*, *105*(3), 1864–1876. https://doi.org/10.1121/1.426723

Luo, S. (2017). Gestural overlap across word boundaries: Evidence from English and Mandarin speakers. *Canadian Journal of Linguistics*, *62*(1), 56–83. https://doi.org/10.1017/cnj.2016.37

Macafee, C. (1997). Ongoing change in modern Scots: The social dimension. In *The Edinburgh History of the Scots Language* (pp. 514–548).

Macafee, Caroline. (1994). *Traditional Dialect in the Modern World: A Glasgow Case Study*. Lang.

Maguire, W. (2017). Epenthesis in liquid+consonant clusters in Scots. In J. Cruickshank & R. McColl Millar (Eds.), *Before the Storm: Papers from the Forum for Research on the Languages of Scotland and Ulster triennial meeting, Ayr 2015* (pp. 156–183).

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control, 14*(3), 380–407. https://doi.org/10.1123/mcj.14.3.380

Marshall, J. (2004). *Language Change and Sociolinguistics: Rethinking Social Networks*. Palgrave MacMillan.

Maté, I. (1996). *Scots Language: A Report on the Scots Language Research Carried Carried out by the General Register Office for Scotland in 1996*.

McMillan, C. T., & Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition, 117*(3), 243–260. https://doi.org/10.1016/j.cognition.2010.08.019

Munhall, K., & Löfqvist, A. (1992). Gestural aggregation in speech: laryngeal gestures. *Journal of Phonetics*, *20*(1), 111–126. https://doi.org/10.1016/s0095-4470(19)30242-6

Murdoch, S. (1995). *Language Politics in Scotland*. AUSLQ.

Murray, J. (1873). *The dialect of the southern counties of Scotland*. The Philological Society.

Murray, R. W., & Vennemann, T. (1983). Sound change and syllable structure in Germanic phonology. *Language*, 514–528.

Myers, S. (2000). Boundary disputes: The distinction between phonetic and phonological sound patterns. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge: Conceptual and empirical issues* (p. 272). Oxford University Press.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. *15th ICPhS Barcelona*, *3*, 2253–2256.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of*

Neuroscience Methods, *162*(1–2), 8–13. https://doi.org/10.1016/J.JNEUMETH.2006.11.017

Pierce, M. (2007). Vowel epenthesis vs. schwa lexicalization in classical Armenian. *Journal of Indoeuropean Studies*, *35*(1/2), 111.

Pierrehumbert, J., Beckman, M. E., & Ladd, D. R. (2000). Conceptual foundations of phonology as a laboratory science. In *Phonological knowledge: Conceptual and empirical issues* (pp. 273–303).

Pierrehumbert, J. B. (2002). Word-specific phonetics. In C. Gusshoven & N. Warner (Eds.), *Laboratory Phonology* (7th ed., pp. 1–24). Mouton de Gruyter.

Plug, L., Shitaw, A., & Heselwood, B. (2019). Inter-consonantal intervals in Tripolitanian Libyan Arabic : Accounting for variable epenthesis. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, *10*(1), 1–33.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. https://doi.org/10.7282/T34M92MV

R Development Core Team. (2016). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, https://www.R-project.org. http://www.r-project.org/

Ridouane, R., & Fougeron, C. (2011). Schwa elements in Tashlhiyt word-initial clusters. *Laboratory Phonology*, *2*(2), 275–300. https://doi.org/10.1515/labphon.2011.010

Ripley, B., Venables, W., & Ripley, M. B. (2016). *Package "nnet". R package version 7*.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499–526. https://doi.org/10.1016/S0167-9457(00)00030-0

Scobbie, J. M., Gordeeva, O. B., & Matthews, B. (2006). Acquisition of Scottish English Phonology: an overview. In *QMUC Speech Science Research Centre Working Paper* (Vol. 7).

Scobbie, J. M., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. In *Contrast in Phonology: Theory, Perception, Acquisition* (QMUC Speech Science Research Centre Working Paper WP-8, Issue 8). https://doi.org/10.1515/9783110208603.1.87

Scobbie, J. M., & Wrench, A. A. (2003). An articulatory investigation of word final /l/ and /l/-sandhi in three dialects of English. *Proceedings of the 15th International Congress of Phonetic Sciences.* .

Scobbie, J., & Stuart-Smith, J. (2008). Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English. *Contrast: Perception and Acquisition: Selected Papers from the Second International Conference on Contrast in Phonology*, 87–113.

Scotland, N. R. of. (2011). *2011 Census*. Digitised Boundary Data (Scotland) [Computer File]. https://borders.ukdataservice.ac.uk

Sebregts, K. (2015). Boundary disputes and sociophonetic variation: Schwa-epenthesis in Dutch r-C clusters. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.

Selkirk, E. (1984). On the major class features and syllable theory. In *Language sound structure* (Vol. 5079, Issue 6, pp. 107–136). MIT Press. https://doi.org/10.1037//0003-066X.46.5.506

Sell, K. (2012). Sociolinguistic findings on schwa epenthesis in Galway English. In M. N. C. Bettina Migge (Ed.), *New Perspectives on Irish English* (Vol. 44, p. 47). John Benjamins Publishing Company.

Shaw, J. A., & Kawahara, S. (2018). The lingual articulation of devoiced /u/ in Tokyo Japanese. *Journal of Phonetics*, *66*, 100–119. https://doi.org/10.1016/j.wocn.2017.09.007

Shaw, J., Gafos, A., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure. *Phonology*, *28*(3), 455–490. https://doi.org/10.25932/publishup-41247

Shitaw, A. (2014). An instrumental phonetic investigation of timing relations in two-stop consonant clusters in Tripolitan Libyan Arabic. *Linguistics and Phonetics*, *Ph.D*(June).

Sievers, E. (1881). *Grundzüge der Phonetik.* Breitkopf und Hartel.

Sole, M.-J. (1995). Spatio-Temporal Patterns of Velopharyngeal Action in Phonetic and Phonological Nasalization. *Language and Speech*, 38–39. https://journals.sagepub.com/doi/pdf/10.1177/002383099503800101

Solé, M. J. (2007). Controlled and mechanical properties of speech: A review of the literature. In M-J. Solé, P. Beddor, & M. Ohala (Eds.), *Experimental Approaches to Phonology* (pp. 302–321).

Solé, M., & Ohala, J. J. (2010). *What is and what is not under the control of the speaker: Intrinsic vowel duration*. https://doi.org/10.1515/9783110224917.5.607

Solé, Maria-Josep. (1992). Phonetic and Phonological Processes: The Case of

Nasalization. *Language and Speech*, *35*(1–2), 29–43.

Sproat, Richard & Fujimura, O. (1993). *Allophonic_variation_in_English_l_and_it.pdf* (pp. 291–311).

Stanley, R. (1967). Redundancy Rules in Phonology. *Language*, *43*(2), 393–436. http://www.jstor.com/stable/411542

Stanton, J., & Zukoff, S. (2018). Prosodic identity in copy epenthesis: Evidence for a correspondence-based approach. *Natural Language and Linguistic Theory*, *36*(2), 637–684. https://doi.org/10.1007/s11049-017-9385-9

Stetson, R. H. (1951). Motor phonetics: A study of speech movements in action. Amsterdam: North-Holland. In J. A. S. Kelso & K. G. Munhall (Eds.), *R. H. Stetson's motor phonetics: A retrospective edition.* North Holland.

Stuart-Smith, J. (2008). Scottish English: Phonology. In Bernd Kortmann & C. Upton (Eds.), *Varieties of English: The British Isles* (pp. 48–70). Walter de Gruyter.

Stuart-Smith, Jane. (2003). The phonology of modern Urban Scots. In J. Corbett, J. D. McClure, & J. Stuart-Smith (Eds.), *The Edinburgh Companion to Scots*. Edinburgh University Press.

Stuart-Smith, J. (2004). Scottish English: Phonology. In B. Kortmann, K. Burridge, E. W. Schneider, R. Mesthrie, & C. Upton (Eds.), *Varieties of English: The British Isles* (pp. 47–67). Mouton de Gruyter.

Stuart-Smith, J. (2007). A Sociophonetic Investigation of post-vocalic /r/ in Glaswegian Adolescents. *International Congress of Phonetic Sciences (ICPhS)*, *August*, 1449–1452.

Stuart-Smith, J., & Lawson, E. (2017). Scotland: Glasgow and The Central Belt. In R. Hickey (Ed.), *Listening to the Past: Audio Records of Accents of English* (pp. 171–198). Cambridge UP. https://doi.org/10.1017/9781107279865.009

Stuart-Smith, J., Lawson, E., & Scobbie, J. M. (2014). Derhoticisation in Scottish English: a sociophonetic journey. In *Advances in Sociophonetics* (pp. 59–96). John Benjamins. https://doi.org/10.1002/14651858.CD001920.pub3/abstract

Stuart-Smith, Jane, Timmins, C., & Tweedie, F. (2006). Conservation and innovation in a traditional dialect L-vocalization in Glaswegian*. In *English World-Wide* (Vol. 27, Issue 1).

Team, A. (2016). *Audacity (R): Free Audio Editor and Recorder* (2.1.2). http://www.audacityteam.org

Team, A. (2019). *Audacity (R): Free Audio Editor and Recorder (Version 2.3.2)*.

Team, R. C. (2019). *R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing*.

The Scottish Credit and Qualifications Framework. (n.d.). Framework, Scottish Credit and Qualifications. https://scqf.org.uk/interactive-framework/

Tuller, B., & Kelso, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research*, *34*(3), 501–508. https://doi.org/10.1044/jshr.3403.501

Tuller, B., & Kelso, J. S. (1990). Phase transitions in speech production and their perceptual consequences. *The Journal of the Acoustical Society of America*, *86*.

Vaux, B., & Wolfe, A. (2009). The Appendix. *Contemporary Views on Architecture and Representations in Phonology*, *48*, 1–33.

Vennemann, T. (1988). The rule dependence of syllable structure. *On Language: Rhetorica, Phonologica, Syntactica: A Festschrift for Robert P. Stockwell from His Friends and Colleagues*, 257–283.

Warner, N. & Weber, A. (2002). Stop epenthesis at syllable boundaries. *Seventh International Conference on Spoken Language Processing*.

Warner, N., Jongman, A., Cutler, A., & Mücke, D. (2001). The phonological status of Dutch epenthetic schwa. *Phonology*, *18*(3), 387–420. https://doi.org/10.1017/S0952675701004213

Wells, J. C. (1982). Accents of English. In *Accents of English*. Cambridge UP. https://doi.org/10.1017/cbo9780511611766

Wettstein, P. (1942). *The phonology of a Berwickshire dialect*. Schuler.

Wickham, H. (2016). ggplot2 - Elegant Graphics for Data Analysis (2nd Edition). *Journal of Statistical Software*, *77*(Book Review 2), 3–5. https://doi.org/10.18637/jss.v077.b02

Wickham, H. (2017). Tidyverse: Easily install and load the “tidyverse”. *R Package Version*, *1*(1).

Wickham, H., François, R., Henry, L., & Müller, K. (2020). *dplyr: A Grammar of Data Manipulation. R package version 0.8. 0.1.* https://cran.r-project.org/package=dplyr

Wickham, H., & Henry, L. (2020). *tidyr: Easily tidy data with “spread()” and “gather()” functions*.

Wilson, J. (1915). *Lowland Scotch as spoken in the Sthrathearn district of Perthshire*. Oxford University Press.

Yip, M. (1991). *Coronals, consonant clusters, and the coda condition.* (Paradis & Prunet (Eds).

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Proceedings - European Conference on Noise Control*, 5687–5690. https://doi.org/10.1121/1.2935783

Zai, R. (1942). *The phonology of the Morebattle dialect*. Raeber.

Zsiga, E. C. (1994). Acoustic evidence for gestural overlap in consonant sequences. *Journal of Phonetics*, *22*(2), 121–140. https://doi.org/10.1016/s0095-4470(19)30189-5

Zsiga, E. C. (1997). Features, Gestures, and Igbo Vowels : An Approach to the Phonology-Phonetics Interface. *Language*, *73*(2), 227–274.