A Description of the Rhythm of Barunga Kriol using Rhythm Metrics and an Analysis of Vowel Reduction

Amit German

BSc/BA (University of Sydney)

A thesis submitted for the Degree of

Master of Philosophy

The MARCS Institute for Brain, Behaviour and Development
ARC Centre of Excellence for the Dynamics of Language
Western Sydney University
August 2018
Statement of Authentication

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material, either in full or in part, for a degree at this or any other institution.
Abstract

Kriol is an English-lexifier creole language spoken by over 20,000 children and adults in the Northern parts of Australia, yet much about the prosody of this language remains unknown. This thesis provides a preliminary description of the rhythm and patterns of vowel reduction of Barunga Kriol - a variety of Kriol local to Barunga Community, NT – and compares it to a relatively standard variety of Australian English.

The thesis is divided into two studies. Study 1, the Rhythm Metric Study, describes the rhythm of Barunga Kriol and Australian English using rhythm metrics. Rhythm metrics quantify durational variability across a speech sample (e.g. DeltaV, VarcoV) or within (e.g. nPVI-V), and discriminate rhythms of different languages and varieties (e.g. White & Mattys, 2007). Ten young adult female participants (5 Barunga Kriol, 5 Australian English) were recorded telling a story to a familiar peer using picture books as stimuli. Recordings were orthographically transcribed then force-aligned (using WebMAUS) and rhythm metrics (VarcoV, VarcoC, nPVI-V and %V) were calculated. It was found that Australian English had significantly higher vocalic and consonantal variability at a local level than Barunga Kriol, and lower overall ‘vocalic-ness’, suggesting that Barunga Kriol may be more ‘syllable-timed’ than Australian English.

Study 2, the Vowel Reduction Study, compared patterns of vowel reduction in Barunga Kriol and Australian English. Cross-linguistically, reduced vowels tend to have shorter duration and shifts in spectral characteristics towards a mid-central, neutral point in the F1 by F2 vowel space. In English, vowels in unstressed positions such as in function words and in medial position in an utterance are most likely to undergo reduction. Duration and spectral measurements were taken for a selection of single syllable words in Barunga Kriol and Australian English, and it was found that Barunga Kriol showed similar patterns to English, with vowels in function words and in medial position in Barunga Kriol showing shorter duration and spectral shifts towards a central point in the vowel space. The results of the first study suggest some durational rhythmic differences between Barunga Kriol and Australian English, but that these differences are not rooted in the patterns of vowel reduction explored in the second study.

This thesis contributes the first in depth studies of vowel reduction patterns and rhythm using rhythm metrics in any variety of Kriol or Australian English. The research also sets an adult
baseline for metric results and patterns of vowel reduction for Barunga Kriol and Australian English, useful for future studies of child speech in these varieties. As rhythm is a major contributor to intelligibility, the findings of this thesis have the potential to inform teaching practice in English as a Second Language.
Acknowledgments

I would like to acknowledge that this thesis was written in Naarm (Melbourne), on the lands of the Wurundjeri and Boonwurrung People of the Kulin Nation, and recognise their continuing connection to land, water and community. This land has never been ceded, and no treaty has ever been written for Indigenous people in Australia. I pay my respects to Kulin Nation elders past, present, and emerging, as well as their ongoing resistance to colonisation for over two hundred years.

I would like to thank my primary supervisor and mentor Associate Professor Caroline Jones, who time and again, has gone above and beyond to support me throughout my candidature. Caroline, thank you for your gentleness, patience, humour, guidance and encouragement. I am indebted to you.

Thank you also to my co-supervisors, Dr. Rachel Hendery, Dr. Ann Burchfield and Dr. Vincent Aubanel. My time with each of you has been short but invaluable, and your unique insights and perspectives have supported this work to be as robust as possible.

I would like to express my sincere gratitude to all who were recorded and participated in this research, from Barunga Community and Melbourne. This research would not have been possible without your contribution. Thank you to Tiarnah Ahfat, Delvean Ahfat, Jaidine Fejo, Anita Painter and Lee Rosas at Barunga Community and in Katherine, NT for help with transcriptions and making me feel so welcome on fieldwork. It has been a pleasure to work with you.

Thank you to my colleagues and friends at Western Sydney University, Dr. Anne Dwyer and Mark Richards who provided me with feedback on some of my earliest drafts, as well as support and encouragement. I am also grateful Denise Angelo who gave me insightful feedback and suggestions on my chapter about Kriol, as well as life in general. Also, thank you to Dr. Weicong Li who assisted me with a Matlab script for my second study, and Dr. James Whang who saved me at the last hour with statistics guidance.

A warm thank you to the linguistics community at the University of Melbourne who were a supportive student body and helped me combat the isolation of being an interstate student - particularly to Dr. Debbie Loakes, who always has time for me.
I would like to acknowledge Western Sydney University who supported me financially and awarded me a generous Higher Degree Research stipend with Caroline Jones’ Future Fellowship, and also the ARC Centre of Excellence for the Dynamics of Language (CoEDL), for supporting me with travel and fieldwork funding. Thank you to CoEDL for providing fantastic events and opportunities to learn, network and collaborate with like-minded researchers and academics.

Thank you to my friends and family for the endless words of encouragement, for believing in me, and for being my biggest fans. Thank you to my queer and gender-diverse companions for seeing and holding me through these years. Thanks in particular to my parents, Dr. Ester Senderey and Gustavo German, who instilled in me from a young age the importance of knowledge, learning and challenging myself. This thesis is dedicated to you both.

And lastly, Fielding Frost: always by my side through the best of times and the not-so-best, who colours the minutiae of my days, pushes me to be the best version of myself. Thank you for existing.
Table of Contents

Chapter 1: Introduction .................................................................................................................. 1

Chapter 2: Kriol ............................................................................................................................ 5

  2.1. Social and Linguistic Context of Barunga Kriol ................................................................. 5
  2.2. Emergence and Historical Context of Kriol ................................................................. 6
  2.3. Descriptions of Kriol ............................................................................................................ 7
  2.4. Phonetics and Phonology of Kriol ...................................................................................... 8
      2.4.1. Superstrate and Substrate Influence ........................................................................ 11
          2.4.1.1. Aboriginal Languages ..................................................................................... 12
          2.4.1.2. Superstrate language – English in Australia .................................................... 13
  2.5. Conclusion .......................................................................................................................... 15

Chapter 3: Rhythm Background .................................................................................................. 16

  3.1. The Concept of Rhythm ..................................................................................................... 16
  3.2. The Rhythm Metrics ......................................................................................................... 22
      3.2.1. Global metrics .......................................................................................................... 23
      3.2.2. Local metrics ........................................................................................................... 26
  3.3. Issues Surrounding the Use of Rhythm Metrics .............................................................. 31
      3.3.1. Rhythm Metrics and Speech Rate ........................................................................ 31
      3.3.2. Focus on Duration, and Disregard of Rhythm as a Perceptual Entity ..................... 31
      3.3.3. Methodological Issues ............................................................................................. 34
      3.3.4. Issues of Interpretation ............................................................................................ 37
  3.4. Conclusion .......................................................................................................................... 38

Chapter 4: Rhythm Metric Study ............................................................................................... 39

  4.1. Research Hypotheses ........................................................................................................ 39
      4.1.1 Research Question 1 ............................................................................................... 40
      4.1.2. Research Question 2 .............................................................................................. 40
  4.2. Method ............................................................................................................................... 41
      4.2.1. Participants .............................................................................................................. 41
      4.2.2. Data Collection ...................................................................................................... 41
      4.2.3. Recordings .............................................................................................................. 42
      4.2.4. Data Preparation .................................................................................................... 42
      4.2.5. Analysis .................................................................................................................. 43
  4.4. Results ............................................................................................................................... 47
      4.4.1. Variability of Vocalic and Consonantal Intervals in Kriol and English ................. 48
          4.4.1.1. VarcoV .............................................................................................................. 48
          4.4.1.2. VarcoC ............................................................................................................. 49
Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix 1:</td>
<td>Coding Protocols</td>
</tr>
<tr>
<td>Appendix 2:</td>
<td>Rhythm Metric Results Including Final Segment</td>
</tr>
<tr>
<td>Appendix 3:</td>
<td>Speaker Variation for Rhythm Metric Results</td>
</tr>
<tr>
<td>Appendix 4:</td>
<td>Function and Content Words Analysed in the Vowel Reduction Study</td>
</tr>
<tr>
<td>Appendix 5:</td>
<td>Vowel Reduction Study Figures using Raw Durational Data</td>
</tr>
</tbody>
</table>
Appendix 6: Vowel Reduction Study Results of Duration and Vowel Interaction Model....................136
Appendix 7: Vowel Reduction Study Results for Dispersion and Vowel Interaction Model............139
List of Figures

Figure 1 Vowel inventory of Kriol, orthographic (Schultze-Berndt et al., 2013, p. 242).................. 9
Figure 2 Consonant inventory of Kriol (Schultze-Berndt et al., 2013, p. 243).................................10
Figure 3 Distribution of languages over the (%V, ∆C) plane (Ramus et al., 1999, p. 273).............25
Figure 4 Successive vowel durations of two hypothetical languages A and B (Low et al., 2000, p. 382)......................................................................................................................28
Figure 5 PVI profiles for eighteen languages (Grabe & Low, 2002, p. 24)......................................29
Figure 6. VarcoV Scores for Kriol and English..................................................................................48
Figure 7 VarcoC scores for Kriol and English..................................................................................49
Figure 8 nPVI-V Scores for Kriol and English..................................................................................50
Figure 9 %V for Kriol and English....................................................................................................51
Figure 10 Spectral patterns for full and potentially reduced vowels in British English (reproduced from Low, 1998, p. 57)..................................................................................56
Figure 11 Diagram for calculating Euclidean distance for an arbitrary token...............................65
Figure 12 Duration by Word Type (Kriol and English).................................................................69
Figure 13 Vowel Duration (Log Transformed) by Word Type and Vowel (Kriol and English)....70
Figure 14 F1 by F2 Vowel Plot: Vowels /ɪ/, /ɛ/ and /ʌ/ by Word Type (Kriol and English)...........72
Figure 15 Dispersion by Word Type (Kriol and English)..............................................................73
Figure 16 Dispersion by Word Type for vowels /ɪ/, /ɛ/ and /ʌ/ (Kriol and English)......................74
Figure 17 Duration by Utterance Position (Kriol and English).......................................................78
Figure 18 Duration by Utterance Position and Vowel (Kriol and English)......................................79
Figure 19 Dispersion by Utterance Position (Kriol and English)...................................................81
Figure 20 F1 by F2 Vowel Plot: vowels /ɪ/, /ɛ/ and /ʌ/ by Utterance Position (Kriol and English)....82
Figure 21 Dispersion by Vowel and Utterance Position, Kriol and English.................................83
Figure 22 Duration (Log Transformed) by Word Type for Data without Finals (top) and with Finals (bottom)..............................................................................................................90
Figure 23 Duration by Dispersion, with Regression line for Dispersion as Fixed Effect, Duration as Predictor..................................................................................................................................92
Figure 24 VarcoV Scores for Kriol and English (Final Segment Included).................................122
Figure 25 VarcoC Scores for Kriol and English (Final Segment Included).......................................123
Figure 26 nPVI-V Scores for Kriol and English (Final Segment Included)....................................124
Figure 27 %V for Kriol and English (Final Segment Included)......................................................125
Figure 28 VarcoV Scores for Speakers of Kriol and English..........................................................126
Figure 29 VarcoV Scores for Speakers of Kriol and English..........................................................127
Figure 30 nPVI-V Scores for Speakers of Kriol and English.........................................................128
Figure 31 %V for Speakers of Kriol and English.............................................................................129
Figure 32 Duration (Raw) by Word Type ................................................................. 133
Figure 33 Duration (Raw) by Word Type and Vowel ............................................. 134
Figure 34 Duration (Raw) by Utterance Position ................................................... 134
Figure 35 Duration (Raw) by Utterance Position and Vowel ................................. 135

List of Tables
Table 1 Descriptive Statistics for Rhythm Metric Scores by Language (estimates from statistical models) .................................................................................................................. 48
Table 2 Summary of all Rhythm Metric Study Results. ............................................. 52
Table 3 Descriptive Statistics for Duration/Dispersion by Word Type (estimates are from statistical models) .................................................................................................................. 68
Table 4 Descriptive Statistics for Duration by Word Type and Vowel with Interaction (estimates are from statistical models) ........................................................................................................ 71
Table 5 Descriptive Statistics for Dispersion by Word Type and Vowel with Interaction (estimates are from statistical models) ........................................................................................................ 75
Table 6 Descriptive Statistics for Duration/Dispersion by Utterance Position (estimates are from statistical models) .................................................................................................................. 77
Table 7 Summary of all Results (collapsed across vowels) ....................................... 85
Table 8 Descriptive Statistics for Rhythm Metric results by Language, with Final Segment Included (estimates are from statistical models) .................................................. 121
Table 9 Descriptive Statistics for Duration by Utterance Position and Vowel with Interaction (estimates are from statistical models) ........................................................................................................ 138
Table 10 Descriptive Statistics for Dispersion by Utterance Position and Vowel with Interaction (estimates are from statistical models) ........................................................................................................ 140

List of Equations
Equation 1 Equation for $\Delta V$ (Ramus et al., 1999) .............................................. 24
Equation 2 Equation for $\Delta C$ (Ramus et al., 1999) .............................................. 24
Equation 3 Equation for $\%V$ (Ramus et al., 1999) .............................................. 24
Equation 4 Equation for VarcoV (Ferragne & Pellegrino, 2004; White & Mattys, 2007a, 2007b) ................................................................. 26
Equation 5 Equation for VarcoV (Dellwo, 2006, 2010) ......................................... 26
Equation 6 The raw Pairwise Variability Index (rPVI) (Low, 1998) ....................... 27
Equation 7 The normalised Pairwise Variability Index (nPVI) (Deterding, 1994; Grabe & Low, 2002) ................................................................. 27
Kriol is an English-lexifier creole language spoken by over 20,000 children and adults in the Northern parts of Australia (ABS, 2010), yet much about the prosody of this language remains unknown. This thesis seeks to describe the rhythm of Barunga Kriol - a variety of Kriol local to Barunga Community, Northern Territory, Australia - and to compare it to that of a relatively standard variety of Australian English. The phonology of Kriol varieties differ from Australian English, at least segmentally, so there is reason to expect that patterns of prosodic prominence may also differ. Segmentally, for example, the Barunga Kriol variety has a more constrained set of vowel phonemes and more consonant variability than Australian English (Jones, Demuth, Li, & Almeida, 2017). This can be seen in the example sentence (1) where the vowel in the word ‘girl’, derived historically from English, is pronounced instead with /e/. This example sentence also displays other grammatical characteristics of Kriol, thought to come from Australian Aboriginal languages. Grammatically, for example, the use of auxiliary ‘bin’ as a past tense marker (<been>, in English) and ‘im’ used to express 3rd person singular (any gender), here marks a possessive. Such patterned differences from Australian English exist at phonological, semantic, and syntactic levels. They are likely due to the process of creolization of Kriol, which historically drew for its lexicon mostly from English, but also shows effects in particular of the traditional Aboriginal languages originally spoken in the area.

(1) Wandei wan lilgel bin pleiplei garra im dog
one.day a little.girl PST playing 3SG dog
‘One day a little girl was playing with her dog’.

North Australian Kriol varieties were initially described in the 1970s and 1980s (e.g. Harris, 1986; Hudson, 1983; Sandefur, 1979, 1986). More recent research on Kriol varieties in the Northern Territory has primarily focused on syntax and morphology (Dickson, 2015; Munro, 2011; Nicholls, 2010; Schultze-Berndt, Meakins, & Angelo, 2013) and although recent research has started to examine phonetics and phonology (e.g. Baker, Bundgaard-Nielsen, & Graetzer, 2014; Bundgaard-Nielsen & Baker, 2015; Jones et al., 2017; Jones, Meakins, & Buchan, 2011), very little is known about prosody, including rhythm and patterns of vowel reduction. This
thesis adds to this growing body of research on Kriol varieties. With the inclusion of a
comparison sample of Australian English, the thesis will also add to the understanding of
suprasegmentals in Australian English, an area less developed than segmental phonology (e.g.
Cox, 1996; Cox & Palethorpe, 2007). Further, the research adds to the developing body of
research on prosodic rhythm found in phonetics and sociolinguistic literatures.

Rhythm can be conceptualised as a patterning of weak and strong events through time: the
“systematic patterning of sound in terms of timing, accent and grouping” (Patel, 2007, p. 96). In
spoken language, patterns are found in the contrasts between prominent (e.g. stressed)
syllables against weak (e.g. unstressed) ones. In the mid 20th century, attempts were made to
class languages into rhythm groups based on the variability in duration of syllables and
placement of stress (the Rhythm Class Hypothesis) (Abercrombie, 1967; Pike, 1945). Although
this work became well known in the literature, some researchers advocate that research focus
is better spent on comparing languages to each other along a spectrum rather than grouping
them together into classes (e.g. Dauer, 1983). In one line of work since the mid 1990s, rhythm
is quantitatively described through the use of rhythm metrics. These are formulae designed to
measure the acoustic correlates associated with variation in prominent and weak syllables,
such as variation in duration (e.g. Low & Grabe, 1999; Ramus, Nespor, & Mehler, 1999). To use
rhythm metrics effectively in cross-linguistic comparison of rhythm, it is essential that a
combination of metrics be used on highly controlled speech data (Arvaniti, 2012b; Arvaniti,
Ross, & Ferjan, 2008; Wiget et al., 2010). Rhythm may well be supported by parameters other
than duration, however. It has been argued that to understand rhythm more comprehensively,
variation beyond duration should be explored, such as variation in F0 and intensity (e.g. Dauer,
1983). Analysis of patterns of vowel reduction is another specific source of variation in
prosodic prominence, which can enrich a description of rhythm.

In this thesis, I provide a preliminary description of the rhythm and patterns of vowel
reduction of Barunga Kriol, and compare it to Australian English. The thesis addresses two
overall research questions:

1. How variable are vocalic and consonantal intervals in Barunga Kriol compared to
   Australian English? How ‘vocalic’ are the two languages? Is Barunga Kriol less
   stress-timed than Australian English, as measured by the rhythm metrics?
Chapter 1: Introduction

2. How is vowel reduction realized in Barunga Kriol compared to Australian English, and to what extent do vowels reduce in unstressed positions in Barunga Kriol, as compared to Australian English? What are the effects on vowel reduction of word type (function word vs. content word) and utterance position? If Barunga Kriol is less stress-timed than Australian English, is this evident as less or different patterns of vowel reduction?

Each research question is addressed in a separate study in this thesis. Both studies use the same dataset: controlled recordings of storytelling by young adult, female speakers from Barunga Community, NT for the Barunga Kriol sample, and from Victoria, for the Standard Australian English sample. Both studies utilise automatic segmentation through webMAUS (Kisler et al., 2016) (with additional manual edits) and scripts through Praat (Boersma & Weenink, 2012), RStudio (R Core Team, 2016) and Matlab (The Mathworks, 2014) to speed up processing and analysis. In Study 1, the Rhythm Metric Study, rhythm metrics are used to measure variation in the duration of vocalic and consonantal intervals, providing quantitative measurements to compare Barunga Kriol with Standard Australian English. Following this, Study 2, the Vowel Reduction Study, delivers a description of the patterns of vowel reduction in both Barunga Kriol and Standard Australian English, obtained through temporal and spectral analysis of vowels. The extent of vowel reduction is measured as a function of both word type and utterance position. These studies present a first description of the prosodic rhythm and vowel reduction in Barunga Kriol, and add to our understanding of Standard Australian English.

This research is vital to furthering the description of Kriol varieties. Outcomes of this project have potential impact and application beyond academia. It is known that cross-linguistically, rhythm is a major contributor to intelligibility: failure to master patterns of timing and stress when learning a language can lead to foreign accent (Faber, 1986; Taylor, 1981). Therefore, this research has the potential to inform teaching practice (e.g. English as Second Language instruction) and teacher professional knowledge. Additionally, as this research is the first of its kind describing adult speech, it provides groundwork needed for more sophisticated research into children’s language development towards adult speech in this language (e.g. Payne, Post, & Vanrell, 2011). Children whose home language is Barunga Kriol are exposed to some English in the years before school but this becomes much more intensive once they enter the classroom, so more linguistic work such as this may raise awareness of subtle language differences that
are educationally relevant, and support teachers and parents in working together to optimize children's educational and life chances.

The next chapter (Chapter 2) provides a review of empirical literature on the focus language, Barunga Kriol, as well as rhythm in varieties of English. The subsequent chapters (Chapter 3 and 4) cover relevant theoretical literature on rhythm and outline the first study in this thesis, using rhythm metrics. Following that (Chapters 5 and 6) a review of relevant literature on vowel reduction is provided as well as the second study in this thesis involving vowel reduction. The final chapter (Chapter 7) offers a discussion of the overall results of this thesis and outlook for further research.
Chapter 2: Kriol

Chapter 2 provides background for the focus language of this thesis, Kriol (also known as north Australian Kriol) and the variety that will be described, Barunga Kriol. Section 2.1 in this chapter outlines the social and linguistic context of Barunga Kriol in particular and Section 2.2 the emergence and historical context of Kriol in general. In Section 2.3 the discussion moves to existing descriptions of varieties of Kriol and research on its phonetics and phonology in Section 2.4. Rhythm, such as patterns of timing and prominence, and aspects of connected speech including vowel reduction are some of the least researched areas of this language. In Section 2.5 towards the end of this chapter, languages relevant to the description of Kriol are briefly outlined, including English and a selection of Australian Aboriginal languages (henceforth, Aboriginal languages) based on regional and heritage considerations.

2.1. Social and Linguistic Context of Barunga Kriol

Information in Section 2.1 is based on conversations with my supervisor, Associate Professor Caroline Jones as well as conversations with long-term researcher Denise Angelo and my own observations on field trips to Barunga, NT and Katherine, NT in 2015 and 2016. Any errors are my own.

Barunga Kriol is the term used in this thesis for the home language of children and adults in Barunga Community (previously, Bamyili), located approximately 100km (40 minutes drive) east of Katherine, NT. Barunga Community members travel frequently into Katherine town (e.g. for shopping and administration) where English is spoken, and to other communities in the area (e.g. to visit relatives, for “family business” and cultural reasons) where Kriol varieties are spoken (e.g. Wugularr/Beswick, Jilkminggan, Ngukurr). So too do they travel to communities where some traditional language is still in use (e.g. Bulman/Weemol and outstations). In a meeting about birds at Barrapunta outstation (Emu Springs, about 90kms from Bulman) in July, 2017, Jawoyn, Mimal, Warddeken, Arafura Swamp and Wardaman rangers were in attendance and there were songs and stories in at least four languages: Kunwinjku/Kune, Dalabon, Rembarrnga and Wagilak, as well Kriol and English (ARC Centre of Excellence for the Dynamics of Language, 2017; Jawoyn Association Northern Territory, 2017).
Chapter 2: Kriol

Within Barunga, Barunga Kriol is the main language, though younger people in the community have exposure to English through English-based education and media. Younger people have varying levels of second language proficiency in Australian English. Written Kriol, however, is not in regular use in the community and literacy in Kriol is uncommon, particularly in younger people. From 1976 – 1992 a bilingual school in English and Kriol ran in Barunga, so Community members who attended through those years do have some level of literacy in Kriol. The school followed a model of transfer/transitional bilingualism where Kriol literacy was taught in the earliest years before being replaced with English literacy. Literature in Kriol is limited to children’s books produced at the school, the Bible and a project named ‘Barunga Books’ is currently under way to produce more books.

In this part of the Northern Territory, Kriol is spoken most often as a first language for at least two generations, and in some communities for four or more generations (Butcher, 2008a). Barunga Kriol emerged towards the end of the first half of the 20th century (Sandefur, 1979, 1986). In Barunga community, it is only the grandparents of the young people who may be speakers of traditional languages, which are generally not part of the everyday experience of young people and children, and there are very few speakers of these languages. Historically until perhaps the mid 20th century, speakers living in the area reportedly spoke several Aboriginal languages, often multilingually. These languages are typically from the Central Arnhem Desert region, such as Jawoyn, Dalabon, Mayali (a dialect of Bininj Gun-wok) (Evans, 2003a) and Rembarrnga (Ponsonnet, 2012). It has also been suggested that others from further afield south of the Roper River areas have been historically spoken in Barunga, such as Mangarrayi and Alawa.

2.2. Emergence and Historical Context of Kriol
Barunga Kriol is a variety of the English-lexifier creole (generically named Kriol) spoken throughout the ‘Top End’ of the Northern Territory, from the Queensland gulf country in the East, to the Kimberleys in the west (e.g. Meakins, 2014; Munro, 2000; Schultze-Berndt et al., 2013). The extent of the short overview in this section is on Kriol only, though one other major creole based on English is spoken in Australia, Torres Strait Creole.

Kriol resulted from the stabilisation of a pidgin believed to have originated in New South Wales due to contact between the first colonizers and original inhabitants of the Sydney area (Simpson, 1996; Troy, 1993; Tryon & Charpentier, 2004). From New South Wales, the pidgin
and then resulting creole spread inland and north to Queensland, and was increasingly used in the northern areas due to the need for communication between Aboriginal people and English-speaking pastoralists on cattle stations, as well as non-English-speaking colonists (e.g. Chinese speaking) (for overview see Meakins, 2014).

A pidgin is a restricted language enabling communication among people who do not share a common language, i.e. a lingua franca (Siegel, 2008). A creole on the other hand, is a language developed from a pidgin and used as first language by a community of speakers, more regular and complex than a pidgin. There are various accounts of how the pidgin ultimately spread and developed into the varieties of Kriol spoken today. Some depict an account of rapid creolization of the pidgin in the early 20th century, where Kriol (the Roper River Kriol variety) emerged to meet the need for a common language by children from various Aboriginal groups at the Anglican Roper River Mission near today’s Ngukurr community and progressively spread out over all of northern Australia (Harris, 1986). Others contest that children were responsible for the formation of creole, arguing that adults were active agents in the creolization, which occurred over many generations (Munro, 2000). Others yet argue for independent genesis and creolization in different areas, which accounts for different varieties (Sandefur & Harris, 1986).

With the changes to traditional settlement brought by colonisation, as well as the increased mobility of Aboriginal people, groups of people who had been speakers of different Aboriginal languages gradually shifted to using Kriol as a lingua franca in many places (Munro, 2000). Though the 2016 ABS census (ABS, 2016) recorded that just under than 4,500 people speak Kriol in the Northern Territory, it is estimated that there are 20 to 30,000 speakers of the multiple varieties of Kriol (e.g. ABS, 2010; Meakins, 2014; Sandefur & Harris, 1986). Census data is widely known to be misrepresentative of true multilingualism in remote Australia, for a variety of reasons such as under-reporting of Kriol in favour of heritage languages, unstandardized naming for Kriol and its varieties versus Aboriginal English, or the fact that Kriol has only relatively recently been recognised as a language (see the following for discussion: Dixon & Angelo, 2014; McConvell & Thieberger, 2001; Simpson, 2008).

2.3. Descriptions of Kriol
The first descriptive grammars of Kriol were written by Summer Institute of Linguistics (SIL) linguists who provided a dictionary and described the Kriol spoken at Ngukurr Community (previously Roper River Mission) and Barunga (Bamyili) (Harris, 1986; Sandefur, 1979, 1986)
as well as in the Kimberleys (Hudson, 1983). They also supported speakers in vernacular literacy and orthography standardization through the Kriol Bible translation project, and did much to recognize and raise the status of Kriol, which until the 1970s, had been “held in low esteem by non-Kriol speakers as well as Kriol speakers themselves” (Sandefur, 1986, p. 146). By naming Kriol in its newly standardized orthography, it was also distinguished as a language in its own right, rather than a variety of English.

Descriptions for Kriol in Roper River and Barunga were later expanded (e.g. Harris, 1991; Munro, 2000, 2011; Ponsonnet, 2011; Rhydwen, 1992, 1996; Sandefur, 1991). In more recent descriptions of Kriol it has been common to delimit the variety of Kriol being described by the place that it is spoken, or where the linguist has worked or collected data.

Descriptions have been written for varieties such as Kimberley Kriol spoken around Fitzroy crossing and Halls Creek, WA (Hudson, 1977, 1983; Moses, 2009), Westside Kriol spoken in the northern Victoria River District, Timber Creek area, NT (Schultze-Berndt et al., 2013), Ngan’giwatyfala or Daly River Kriol (Reid, 1990; Rhydwen, 1996, 2003), Wumpurrarni English or Barkly Kriol spoken in Tennant Creek, NT (Disbray, 2008; Disbray & Simpson, 2005) and Central Australian Aboriginal English (Koch, 2000). Gurindji Kriol has also been described as not a variety of Kriol but instead a ‘mixed language’ strongly influenced by both Kriol and a traditional Pama-Nyungan language, Gurindji (emerged from code-switching) (McConvell & Meakins, 2005).

Kriol is spoken over a large landmass of Australia, so variation is unsurprising. To avoid generalising over all possible varieties in descriptions, it makes sense to delimit the variety of Kriol being described as being a variety local to a particular area. Determining whether Kriol spoken in one area is a different variety to that spoken in another area is a difficult discussion, and one beyond the scope of this thesis. Work is currently underway in the Kriol Project at The University of Queensland, describing variation for Kriol’s ‘dialects’ (rather than ‘varieties’). It has been said at least, that although there are differences in pronunciation, structure and lexicon between Kriol varieties, they are all mutually intelligible (Meakins, 2014, p. 375).

2.4. Phonetics and Phonology of Kriol
Research on Kriol varieties has primarily focused on lexicon, syntax and morphology (e.g. Dickson, 2015; Munro, 2011; Nicholls, 2010; Schultze-Berndt et al., 2013) with little
description of phonetics or phonology (e.g. Sandefur, 1979, 1986). Work in this area is increasing, and as a result of research since the mid 1990s, more is known about the phonological system of varieties in the area (Baker et al., 2014; e.g. Bundgaard-Nielsen & Baker, 2015; Jones, Demuth, German, & Cutfield, 2015; Jones, Demuth et al., 2017; Jones, Meakins et al., 2011). Kriol varieties have slightly different phonemic inventory, phonetics and prosody (e.g. Munro, 2000). Some varieties, such as Roper and Barunga Kriol have received more detailed attention, including acoustic study. This section will provide a general overview of the general phonetics and phonology of Kriol.

The vowel system of Kriol is made up of five short monophthongs and five diphthongs, shown in Figure 1 below (Bundgaard-Nielsen & Baker, 2015; Jones et al., 2017; Schultze-Berndt et al., 2013). Kriol has been recorded as historically having only three vowels (Sharpe, 1975, 1985) or between five and seven monophthongs and three or four diphthongs (Sandefur, 1979, 1986), and it is likely that there may be more vowels in some varieties, as evidenced by variation in vowel orthography as well as noted durational contrasts (Bundgaard-Nielsen & Baker, 2015). Kriol has fewer vowels than Australian English, which has 11 contrastive monophthongs (Cox & Palethorpe, 2007) and is more like the smaller vowel systems typical of many Aboriginal languages (Fletcher & Butcher, 2014).

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Open</td>
<td></td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>Diphthongs:</td>
<td>ei ai ou oi au</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1** Vowel inventory of Kriol, orthographic (Schultze-Berndt et al., 2013, p. 242). The consonant inventory has six places of articulation (some varieties also have an interdental stop), with plosives, fricatives, nasals, laterals, a rhotic trill and three approximants (Figure 2) (Baker et al., 2014; Schultze-Berndt et al., 2013). There is variation regarding fricatives, such as for the labiodental /f/ which may be pronounced instead as labial stop /p/ and for the alveo-palatal /sh/ and alveolar fricative /s/ which may instead be pronounced as palatal stop /c/.
## Chapter 2: Kriol

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Inter-dental</th>
<th>Alveolar</th>
<th>Retroflex</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p/b</td>
<td>th* [t]</td>
<td>t/d</td>
<td>rt [t/d]</td>
<td>j [c]</td>
<td>k/g</td>
<td>h</td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>s</td>
<td>rr [l]</td>
<td>ll [l]</td>
<td>ly [ɛ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>rr [r]</td>
<td>r [r]</td>
<td>y [j]</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill/tap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*In some varieties

**Figure 2 Consonant inventory of Kriol (Schultze-Berndt et al., 2013, p. 243).**

With regards to syllable structure, syllables can be open or closed, though there are only certain consonant cluster combinations allowed in initial position: a plosive followed by a liquid, rhotic or glide, or a combination of alveolar fricative and plosive (e.g. /st/, /sk/). This cluster, however, is often reduced to a sole plosive (Schultze-Berndt et al., 2013), and other clusters may have a vowel inserted between consonants, such as /silip/ ‘sleep’ or /sinek/ ‘snake’ (e.g Sharpe, 1975). It has been claimed that word-final consonant clusters are non-existent (Sandefur, 1979) though /lp/ and /ks/ are possible and may be reduced (Schultze-Berndt et al., 2013). Stress has been described as phonemic and predictable (Sharpe, 1975, p. 6).

Though Kriol has been said to have no voicing distinction (Schultze-Berndt et al., 2013), Numbulwar and Ngukurr speakers of Roper Kriol have been described as having a stronger stop voicing contrast (similar to English), and a constriction duration contrast (similar to some Australian Aboriginal languages), as well as a durational vowel contrast (Baker et al., 2014; Bundgaard-Nielsen & Baker, 2015).

With regards specifically to Barunga Kriol, this variety has been referred to in descriptions as a light Kriol, unlike Roper Kriol, which is often regarded as heavy i.e. phonologically further from standard Australian English (Sandefur, 1979, 1986). There is some disagreement on the use of a continuum model from ‘acrolectal’ (closer to superstrate language) to ‘basilectal’ (closer to superstrate) in order to explain variation in Kriol phonology. For example, it has been disputed on the basis that variation in Kriol varieties can be explained by cross-linguistic transfer from substrate languages into Kriol as a second language (Bundgaard-Nielsen & Baker, 2016). More on substrate and superstrate influences will be outlined in the following section.
In terms of phonological variation due to connected speech processes in Kriol, Jones et al. (2015) have tracked variation in pronunciation of function words in Barunga Kriol, and suggested a process of grammaticalization in progress, as in the speech of younger people they are shorter and their vowels more centralized at faster speech rates, in medial position, and in highly frequent word combinations. Other connected speech processes such as co-articulation, lenition and fortition, as well as phonotactics of stress, timing patterns, prosody and rhythm are yet to be described for any variety of Kriol.

This represents the extent of enquiry into the acoustics of Kriol in this area, and prior to this project, there is almost nothing to be found on the phonetics and phonology of Kriol, let alone specifically for Barunga Kriol. There is some research in the area of creolistics, which inform hypotheses and expected results of the studies in this thesis. This research will be briefly reviewed in the following section.

2.4.1. Superstrate and Substrate Influence
As with other creoles around the world, certain features of Kriol can be understood in terms of transfer from ‘superstrate’ and ‘substrate’ languages (Siegel, 2000, 2008). The contributions of these languages to the creole vary depending on context, but often it is the ‘superstrate’ that contributes vocabulary and the ‘substrate’ that contributes some phonology and grammatical structure to the creole (Siegel, 2008). Generally speaking, Kriol derives words from English (the lexifier), and the influence of Aboriginal languages occurs in the grammar and semantics (e.g. Hudson, 1977; Munro, 2011; Siegel, 2011). Aboriginal languages are often mentioned as influencing phonetics and phonology of Kriol, though this is difficult to test with few speakers of these languages left. In this framework, English is the ‘superstrate’ language, typically understood as that of a colonising group of people. The ‘substrate’ language on the other hand, is typically that of the oppressed or colonised people. In the case of Kriol, the “substrate” language/s could potentially be various Australian Indigenous languages, though as mentioned in Section 3.2, as there is no agreement of the process of creolization of the pidgin into the creole spoken today, the precise nature of substrate Australian languages which may have influenced Kriol remains an open question, unlikely to be resolved (e.g. Harris, 1986; Munro, 2000; Sandefur, 1986). In one account (Munro, 2000), it has been suggested that the different varieties of Kriol reflect influence from different substrate languages, via a mechanism of transfer constraints proposed by Siegel (2008).
Chapter 2: Kriol

The origins of Kriol may lie in languages from the Roper River region (e.g. Harris, 1986; Munro, 2004, 2011; Nicholls, 2010) as they are spoken in some proximity to where Kriol is most likely to have originated, and some older generations of speakers of Kriol may have also spoken these languages. As these languages have been argued to be substrate languages and are often invoked for their influence on the sound system of Kriol, the following section will provide a brief overview of the features of these languages.

2.4.1.1. Aboriginal Languages

Aboriginal languages are generally divided into the Pama-Nyungan or non-Pama-Nyungan typological groups (Evans, 2003b; Hale, 1964). Most languages are held to belong to the Pama-Nyungan group, which constitute a single genetic group originating from one “Proto-Pama-Nyungan” language. The non-Pama-Nyungan group on the other hand, are languages which do not constitute a single genetic group, but are instead various languages grouped together for convenience (Evans, 2003b). The languages thought to be substrate languages of Kriol are all non-Pama-Nyungan languages, and based on genetic descent as well as typological similarities, can be separated into the following two families (Baker, 2004; Evans, 2003b; Munro, 2004, 2011): Gunwinyguan languages, which are Ngalakgan, Nunggubuyu and Ngandi; and Marran languages, which are Marra and Warndarrang. Other languages which have historically been spoken in Barunga Community, and their families are as follows: Alawa, similar to Marran languages and often placed in the Marran family, Jawoyn, Dalabon, Rembarrnga and Mayali, all Gunwinyguan languages, and Mangarrayi, which is unclassified, but may be Gunwinyguan (see Evans, 2003a, 2003b).

Though agreement on Kriol’s substrate languages remains unresolved, research on the phonetics and phonology of Australian Aboriginal languages may provide clues as to what might be expected for the rhythm or prosody of Kriol, as these may have transferred into Kriol. This research however, is limited. In a chapter reviewing research into the sound patterns of Aboriginal Languages, Fletcher and Butcher (2014) made claims about possible substrate languages to Kriol. These claims aid in implicating possible hypotheses about stress, prosody and rhythm of Kriol in this thesis. How these relate to rhythm will be clearer after the literature review on rhythm in Chapter 3.

In general, languages of the “Top End” of Australia have more complex syllable phonotactics than central Australian languages. This could be a possible acoustic correlate of stress-timing, a
rhythmic pattern in which stressed syllables tend to occur at regularly timed intervals, regardless of the number of unstressed syllables between (see Chapter 3). Further possible evidence of stress-timing in these languages is that there are many closed syllables in Bininj Gun-wok and Dalabon, and that low central vowels in connected speech display much variation, as well as F1 in unstressed /a/ vowels, and they may also completely elide in Dalabon (Fletcher & Butcher, 2014). There are also suggestions that prominent syllables are longer and more intense in Bininj Gun-wok (Fletcher & Evans, 2002, 2016) and that pitch and tone changes may be important markers of prominence in Aboriginal languages (Fletcher & Butcher, 2014). However, there is some research suggesting that the picture may be more complicated than claiming these languages as simply stress-timed. For example, vowels are not reduced to schwa in Bininj Gun-work (Fletcher & Butcher, 2014), which may be expected given that they do in English, a stress timed language (see Chapter 3 and 5). In a further complication, it was found that although Bininj Gun-wok has been described as stress-timed, when measured with the nPVI-V (a metric for calculating patterns in timing related to rhythm, see Chapter 3), it actually scored lower than British English (prototypically stress-timed) and closer to syllable-timed languages such as Spanish (Fletcher & Butcher, 2014). Mixed evidence such as the claims discussed here, suggest that more research is needed to understand prosody in these languages, including how they may have influenced the prosody of Kriol.

2.4.1.2. Superstrate language – English in Australia
‘Australian English’ in this thesis is the simplified term to denote the majority dialect of English spoken in Australia, also known as ‘Standard Australian English’ (Cox & Palethorpe, 2007). This dialect is in contrast to the other dialects spoken in Australia: Aboriginal English (discussed below) and Ethnocultural varieties. Though Australian English is less documented than varieties such as American English and British English, its phonological system has been documented in detail (e.g. Cox & Palethorpe, 2007; Harrington, Cox, & Evans, 1997; Mitchell & Delbridge, 1965). The rhythm of English in general has been traditionally been described as “stress-timed”, a classification suggesting, amongst other things, that there is high variability in vowel and consonant intervals, due to vowel reduction and lengthening effects from stress, as well as more consonants and more complex consonant clusters compared to “syllable-timed” languages (Chapter 3). As expected in English, for Australian English variation in vowel pronunciation has been shown in some prosodic patterns (Harrington et al., 1997), that schwa is the most common vowel in unstressed syllables (Cox & Palethorpe, 2007) and that duration of vowels is influenced by accentuation, with stressed syllables showing vowel lengthening,
and unstressed syllables showing vowel reduction (e.g. Fletcher & McVeigh, 1993). There is little in-depth acoustic analysis of rhythm, prosody and vowel reduction in Australian English. This thesis will add to this body of literature.

Another variety, or dialect of English spoken in Australia is what is called ‘Aboriginal English’. There is however, indeterminacy in the way in which the term Aboriginal English is used in the literature, making a review complicated. Although considered problematic both linguistically and politically by many linguists and educators, Aboriginal English has been conceptualised as being on a continuum between English and Kriol, ranging from heavier (basilectal, further away from English) to lighter (acrolectal, more similar to English) varieties (Butcher, 2008a; Fletcher & Butcher, 2014). However, Aboriginal English varies throughout Australia, and some say that rather than unifying it into one type of dialect along a continuum, it should more appropriately be treated as different varieties, or referred to as ‘Aboriginal ways of using English’ to do this variation justice (see Eades, 2013).

Despite the problems with conceptions of Aboriginal English, some aspects of its sound systems have been described in the literature. It has been noted that Aboriginal languages have influenced the sound system of Aboriginal English, though empirical support here is limited. For example, some heavier varieties have been described as having phonology that is almost identical to local languages, with basilectal varieties (although which particular ones are discussed in this work is somewhat unclear) show little to no stop voicing distinction or distinction between stops and fricatives, and speakers voice intervocalic obstruents and devoice word-final ones (Butcher, 2006, 2008a). Like descriptions of Kriol and present in many Australian Aboriginal languages, sibilant consonants are realized as the alveopalatal stop (Butcher, 2006).

Aboriginal English in general, has also been claimed to have a “distinct rhythm” which would suggest it is different to Australian English (Fletcher & Butcher, 2014, p. 128). As an oversimplification, this may be because in the basilectal variety, complex syllable onsets are reduced by dropping a consonant, or simplified by inserting a schwa in between consonants (Butcher, 2008a), possibly because these are mostly absent from Australian languages (R. Dixon, 2002). Similarly, many Australian languages allow only open syllables, which may have transferred into Aboriginal English with complex codas often reducing (Butcher, 2008a). Further, stress patterns in Australian languages often fall on the initial syllable of a word and
this pattern is replicated in Aboriginal English, so that words which would otherwise have final stress in Australian English are pronounced differently (Butcher, 2008a).

2.5. Conclusion
There is little literature on the phonology and phonetics of varieties of Kriol (e.g. Baker et al., 2014; Jones & Meakins, 2013), particularly for rhythm and phonological variability due to connected speech processes and for Australian English, there is little in depth analysis. This project seeks to address this gap by comparing the rhythm of Barunga Kriol and English (Study 1), and a connected speech process connected to rhythm, vowel reduction (Study 2). Please note that from here on in this thesis, unless specified otherwise, when I mention “Kriol”, I am referring to the variety spoken at Barunga, from where the data has been collected. If making claims with regards to Kriol as a language in general, I will refer to “varieties of Kriol”. 

Chapter 2: Kriol
Chapter 3: Rhythm Background

This chapter provides a background to rhythm, discussing research relevant to the first study in this thesis, the Rhythm Metric Study (Chapter 4). Section 3.1 introduces the concept of rhythm, its history and various perspectives in the literature. Section 3.2 discusses the rhythm metrics, the formulae employed for measuring rhythm which will be employed in Study 2. Section 3.3 reviews issues surrounding the use of these metrics before the chapter is concluded in Section 3.4.

3.1. The Concept of Rhythm

The idea that languages can be differentiated from each other by their rhythm is not a new one. As far back as 1940, language rhythm was described impressionistically as being either ‘morse-code’ or ‘machine-gun’ in nature (Lloyd James, 1940), a distinction later termed stress- and syllable-timing (Rhythm Class Hypothesis) (Abercrombie, 1967; Pike, 1945). In the Rhythm Class Hypothesis, a language with stress-timed rhythm (e.g. English) is said to have a discernible, regular beat made up of prominent syllables which are evenly spaced no matter how many syllables are between them. In a language with syllable-timed rhythm (e.g. Spanish), the syllables are themselves evenly spaced no matter where prominent syllables are placed. In this early concept of rhythm, the focus was placed on classifying languages into two groups, based on a distinction related to isochrony, the division of time into equal portions. For stress-timed languages, isochrony is maintained at the foot level, such that the duration of the interval between prominent syllables is kept constant and syllables within this interval are allowed to compress or elongate. For syllable-timed languages it is maintained at the syllable level, so all syllables are kept constant in duration (Abercrombie, 1967). Traditionally, most Germanic and Slavonic languages have been classified as stress-timed, Romance languages as syllable-timed, and later, a third group was added based on isochrony of the mora, a category in which Japanese, and possibly Telugu have been placed (Han, 1962; Murty, Otake, & Cutler, 2007).

Rhythmic classifications - often selected on the basis of intuition on the part of the classifier - have been historically maintained, despite evidence suggesting that isochrony is not useful for separating languages into stress-, syllable- or mora-timed classifications (e.g. Bolinger, 1965; Delattre, 1966; Pointon, 1980; Shen & Peterson, 1962). Some of the most compelling evidence
Chapter 3: Rhythm Background

against isochrony as the basis for classifying languages can be found in seminal studies such as those by Roach (1982) and Dauer (1983). In the original languages classified by Abercrombie (1967) in the formulation of the Rhythm Class Hypothesis – English, Russian and Arabic as stress-timed and French, Telugu and Yoruba as syllable-timed – Roach (1982) searched for isochrony at both the syllable and foot level by calculating the standard deviation of 1) syllable durations and 2) the durations of the intervals between prominent syllables. It was found that standard deviations for both were not different across the two classes, despite the reasoning that syllable-timed languages should have lower standard deviations for syllable durations higher deviation for the intervals between prominent syllables, and stress-timed the opposite (assuming that syllable durations are equal in syllable-timed and more variable in stress-timed). For the intervals between prominent syllables, both Roach (1982) and Dauer (1983) examined their duration with respect to the number of syllables they contained, reasoning that in stress-timed languages the duration should remain equal regardless of the number of syllables, as syllables are allowed to compress and elongate to maintain isochrony. In syllable-timed languages, on the other hand, syllables are kept even, so these intervals should be longer. Both found that for all languages the intervals became longer when they contained more syllables, no matter the classification. Through this research, two important claims were made about the basis of rhythmic classification: that isochrony may lie not in production but instead only in perception, and that rhythm classification must be explained by both a phonetic and phonological account, as both affect how rhythm is perceived, beyond just durational patterns (Dauer, 1983).

These claims represented a change towards a concept of rhythm beyond just a timing distinction in the acoustic signal to one with the intention of exploring the acoustic correlates related to the impression of rhythmic differences. For example, Dauer (1983) showed that in a selection of stress-timed languages, monosyllabic intervals between prominences were on average equal in duration to disyllabic intervals between prominences in syllable-timed languages. It was also found that in stress-timed languages, this interval was also restricted to a small amount of syllables, whereas this was more flexible in syllable-timed languages. These observations were argued to contribute to the impression that these intervals sound more regular in stress-timed languages (Dauer, 1983). No longer being based solely on duration, this account of rhythm related to perception was a change from the way that rhythm had been previously been defined. In this account of rhythm, an accumulation of various features
Chapter 3: Rhythm Background

(acoustic correlates) make a language sound more or less like one classification than the other. This accumulative features approach to rhythm represented not only a shift from a focus on isochrony, but also a shift from discrete classification to a model where languages lie on a continuum or along a rhythmic dimension. In this model, the ends of the continuum are represented by prototypical stress- or syllable-timed languages, determined as a function of the accumulation of properties that promote that rhythm type. Languages can also have intermediate rhythms between the two ends on the continuum (Dauer, 1983, 1987) Two such languages are Polish and Catalan, which had been argued to fit in neither stress- or syllable-timed groups and could now be placed along the continuum (Nespor, 1990).

In this extended account of rhythm, impressions of language rhythm are attributed to a variety of specific phonetic and phonological structures or parameters which cause syllables to sound grouped in different ways (Bertinetto, 1977, 1989; Borzone de Marique & Signorini, 1982; Dauer, 1983). The first phonological structures contributing to the impression of rhythmic differences between classifications were said to be a language’s syllable structures as well as differing patterns of vowel reduction and the role of stress (Dauer, 1983) and more have been added since (Bertinetto, 1989; Dauer, 1987). Dauer (1987) developed a checklist for noting the accumulation of syllable-timed and stress-timed features and placing a given language on this continuum.

The following is a list summarising factors or parameters responsible for the impression of different rhythm, provided by Bertinetto (1989):

1. vowel reduction vs. full articulation in unstressed syllables;
2. relative uncertainty vs. certainty in syllable counting, at least in some cases;
3. tempo acceleration obtained (mainly) through compression of unstressed syllables vs. proportional compression;
4. complex syllable structure, with relatively uncertain syllable boundaries, vs. simple structure and well-defined boundaries;
5. tendency of stress to attract segmental material in order to build up heavy syllables vs. no such tendency;
6. relative flexibility in stress placement [...] vs. comparatively stronger rigidity of prominence [...] ;
Chapter 3: Rhythm Background

7. relative density of secondary stresses, with the corresponding tendency towards short inter-stress interval, and (conversely) relative tolerance for large discrepancies in the extent of the inter-stress interval (Bertinetto, 1989).

Of the various features outlined in this concept of rhythm, it has been said that the most important contributors responsible for the impression of rhythmic differences are 1) differing patterns of vowel reduction and 4) a language’s syllable structures (Bertinetto, 1989; Brakel, 1985; Dasher & Bolinger, 1982; Dauer, 1983; Nespor, 1990). Languages that sound stress-timed are said to be more likely to display more types of syllable structures including complex or heavy syllable consonant clusters, as well as display vowel reduction in unstressed syllables. These two things are argued to cause the impression of greater contrast between stressed and unstressed syllables. They also cause unstressed syllables to sound shorter or smaller than they actually are, so the inter-stress intervals seem more equal. On the other hand, languages sounding syllable-timed are more likely to have simple and more open syllable structures, and less likely to display vowel reduction in unstressed syllables (or display it to a lesser extent). This is argued to cause the impression of a smaller difference between stressed and unstressed syllables, so they seem more similar to each other and there is an impression of more syllable regularity.

A further factor influencing the impression of stress- from syllable-timing in this formulation of rhythm is the tendency of stress-timed languages to concentrate more prosodic effects on stressed syllables. These prosodic effects include durational contrasts (lengthening effect) and stressed syllables as the placement of turning points of intonation. Syllable-timed languages on the other hand, show less or no effect of duration or intonation associated with stress placement. Syllable-timed languages are also more likely to break up complex syllables with addition of vowels (epenthesis) or consonants (liaison). Phonetic and phonological features such as syllable structure, patterns of vowel reduction and prominence are said to accumulate in different patterns towards either end of a rhythmic continuum, contributing to the impression of different rhythms (Dauer, 1987).

There are two problems with classifying rhythm based on a list of phonetic and phonological parameters. Firstly, an accumulation of parameters in a language cannot precisely point to where languages are to be placed along the continuum (Arvaniti, 2009). For example, in one study it was found that Dauer's (1983) features were not readily applicable for classifying
Chapter 3: Rhythm Background

Italian and Bulgarian in comparison to German (Barry, Andreeva, Russo, Dimitrova, & Kostadinova, 2003). The parameters tested were duration (expecting stressed syllables to be longer than unstressed ones in stress-timed languages), syllable complexity (higher complexity in stress-timed languages) and vowel reduction (vowels reduced in unstressed positions in stress-timed languages). German is traditionally classified as stress-timed, and the parameters all matched with this classification. However Italian and Bulgarian have less clear definitions. For Italian, a duration effect was found, but the language displays less complex syllables, and though it has been thought to show no vowel reduction, it can be present in certain situations. For Bulgarian, a durational effect was not convincingly found, and it does have more complex syllables though these are rare, and it also does show vowel reduction. Nespor (1990) suggested that rhythm could be neither dichotomous nor continuous between stress- and syllable-timing because the parameters do not easily predict where these languages should be placed on the continuum. Instead it was put forward that these languages be described as ‘rhythmically intermediate’, referring to languages such as Catalan with low syllable complexity as well as vowel reduction and Polish with high syllable complexity and no vowel reduction. Unless languages represent prototypical languages for rhythm classes, they can only be placed between ends of a continuum rather than in a specific location.

A second problem with the phonological account of rhythm based on parameters presented by Dauer (Dauer, 1983, 1987) is that there is a lack of strong empirical testing of their ability to classify languages for their rhythm (for more on this see Arvaniti, 2009, 2012b; Turk & Shattuck-Hufnagel, 2013). Despite these problems, the parameters have been generally accepted in rhythm research, and later work (i.e. rhythm metrics, to follow), has built on this account of rhythm. Further, despite work suggesting that rhythm is not based on isochrony or even on a timing distinction, and despite suggestions that there is no distinction but instead a continuum, rhythm classifications remain a useful distinction, and a useful shorthand as a stand-in for combinations of rhythm correlates. In this thesis, although “stress- and syllable-timing” will be used, it is understood as tending towards syllable- or stress-timing rather than a distinct classification.

More recently, rhythm has been described using empirical measurements known as the rhythm metrics (e.g. Grabe & Low, 2002; Ramus et al., 1999). These are formulae designed to measure the phonetic correlates of the patterns of phonetic and phonological parameters thought to contribute to the impression of rhythm. As mentioned earlier, stress-timed
languages are argued to have greater varieties of syllable structures, made up of heavier and more complex consonant clusters. They also show vowel reduction in unstressed syllables. Syllable-timed languages on the other hand, tend to have less complex consonant clusters and lower instances of vowel reduction. Metrics account for these differences by measuring durational variability of vocalic and consonantal intervals, unlike the syllables and inter-stress intervals proposed in the Rhythm Class Hypothesis. The basic assumption underlying the metrics is that languages can be differentiated by their rhythm, and that languages towards the stress-timed end of the continuum show higher variability in the durations of their vocalic and consonantal intervals than syllable-timed ones. A vocalic interval is a sequence of vowels uninterrupted by consonants, and a consonantal interval is a sequence of consonants uninterrupted by vowels. Both intervals may cross word boundaries, as they are phonetically based. For example, the sentence ‘she’s extra cool’ (/ʃɛkstrəkʊl/) is made up of the alternation of four vocalic intervals and five consonantal intervals: these are /ʃ/, /i/, /z/, /ɛ/, /kstr/, /ə/, /k/, /u/, /l/. The first two consonantal intervals contain just one consonant, but the third has four consonants, which is more complex and will be longer in duration than the other consonantal intervals.

The metric DeltaC (ΔC) (Ramus et al., 1999), accounts for variability in consonantal duration, and shows a higher score for stress-timed languages and lower for syllable-timed. Vocalic durations also vary more in stress-timed languages than in syllable-based languages due to vowel reduction, as durations of vowels in unstressed syllables are shorter than those in stressed syllables. DeltaV (ΔV) measures this variability, however it is affected by other properties such as vowel lengthening processes at certain parts of an utterance, and vowel duration differences intrinsic to certain languages. Instead the percentage of vocalic intervals in the entire duration of the speech sample (%V) is a more useful measurement for vocalic intervals, with a lower percentage reflecting a stress-timed language (Ramus et al., 1999). A stress-timed language has greater consonantal durations, so a greater percentage of the entire duration is potentially taken up by these, lowering the vocalic percentage. The Pairwise Variability Index (PVI) (Low, 1994, 1998) has also been designed to measure consonantal and vocalic variability. The growing body of research utilizing rhythm metrics will be discussed in the next section.
3.2. The Rhythm Metrics

Rhythm metrics arose to meet a need for surface-based, objective evaluation of language rhythm. This section will present an overview of these metrics, which work at a phonetic level by capturing information from the acoustics of speech. As mentioned in the previous section, metrics measure variation in durations of vocalic or consonantal intervals in a speech sample. They can do this in two ways: at a global level, by capturing across the entire speech sample how all intervals of one type vary in their durations (essentially a standard deviation); or at a local level, by capturing how successive intervals differ from each other, and the variability in these differences across the entire sample. This difference in global and local metrics will be further explained in the following subsections, which will outline some of the most popular and widely used metrics: the global metrics $\Delta V$ and $\Delta C$ by Ramus et al. (1999), with additional speech rate normalization: VarcoV and VarcoC (Dellwo, 2006; Ferragne & Pellegrino, 2004), and local metrics raw and normalised Pairwise Variability Index (rPVI and nPVI) (Deterding, 1994; Low, 1994, 1998; Low & Grabe, 1995).

The metrics presented in this section measure vocalic and consonantal intervals, though it has been argued that using metrics based on syllables (and feet) can also be useful in describing rhythm, particular in situations where rhythms “coexist” at different levels (orthogonal dimension rather than continuum, see Nolan and Asu, 2009). Existing metrics based on syllable durations include the Syllable Ratio (Gut, 2003), as well as modifications of the metrics discussed in this section, such as VarcoS (Rathcke & Smith, 2011) and myriad versions of the PVI (e.g. Variability Index, Ackermann & Hertrich, 1994; Deterding, 1994, 2001; the normalised PVI for syllables and Rhythm Ratio, Gibbon & Gut, 2001; and Yet Another Rhythm Determination, Wagner & Dellwo, 2004).

Metrics involving syllables require the notoriously difficult task of determining syllable boundaries, and controversies regarding placement of syllable boundaries are rife in the literature (for a review see Fuchs, 2015). There is little agreement on decisions such as inclusion of consonants between two syllable nuclei into the following or preceding syllable and treatment of syllable boundaries that overlap over word boundaries. This is further compounded by the difficulty in distinguishing stress and word boundaries in the first place. Judgments determining these boundaries are often highly subjective and almost impossible to predict in situations of fast speech or where the segmenter is a non-native speaker of a language. For these reasons, studies using syllable-based metrics have not been as readily
applied in the literature as those based on vocalic and consonantal intervals. One attempt to solve syllabification issues is the Maximum Onset Principle (MOP) (Pulgram, 1970) however, interpretations of the MOP have not been consistent (e.g. syllables were allowed across word boundaries in Gut (2005) but not in Deterding (2001) and Nolan and Asu (2009), and Deterding (1994) approached the problem completely differently. As segmentation decisions can greatly affect metric results, it is difficult to compare between these studies. Further, this Principle makes use of phonological information, making it prone to segmenter-induced error, while a metric that depends on purely phonetic information is more reliable. In the interests of better comparability with other studies and cross-linguistically, as well as for the sake of simplicity and “starting somewhere”, the most popular metrics - those using more objective vocalic and consonantal intervals - are to be employed in the Rhythm Metric Study in this thesis. To follow, a discussion of these metrics.

3.2.1. Global metrics
Inspired by the role that rhythm plays in first language learning by cueing the acquisition of word boundaries, a necessary precursor to assigning meaning to individual words (e.g. Cutler, Mehler, Norris, & Segui, 1986; Jusczyk & Aslin, 1995; Mehler, Dupoux, Nazi, & Dehaene-Lambertz, 1996) and Nazi, Bertoncini and Mehler’s (1998) work on the discrimination of rhythm classes by infants, Ramus et al. (1999) claimed that “the infant primarily perceives speech as a succession of vowels of variable durations and intensities, alternating with periods of unanalysed noise (i.e., consonants)” (p. 270). This motivated the development of the duration-based metrics, which focused on vocalic and intervocalic (henceforth, consonantal) intervals to measure rhythm. This was a change of focus from syllables, which had previously been the focus of rhythmic theory.

The metrics developed were DeltaV ($\Delta V$, Equation 1) and DeltaC ($\Delta C$, Equation 2), the standard deviation of vocalic and consonantal intervals respectively. Higher scores in these reflect higher variation in the duration of these intervals, which can also be understood as reflecting greater complexity in vocalic and consonantal structures in the speech sample. Also developed alongside these metrics is the proportion of duration made up by vocalic intervals in the total utterance duration (%V, Equation 3), where a higher proportion reflects ‘vocalic-ness’ of the speech sample and a lower proportion reflects greater consonantal complexity such as the presence of larger/heavier consonant clusters. These metrics are regarded as ‘global’ metrics, because they measure variability across an entire speech sample.
Chapter 3: Rhythm Background

$\Delta V$ is the standard deviation of the duration of vocalic intervals within a sentence/utterance:

$$\Delta V = \sqrt{\frac{\sum |x - \mu|^2}{N}}$$

where $\Sigma$ means "sum of", $x$ is the duration of a vocalic interval, $\mu$ is the mean of the durations of all vocalic intervals, and $N$ is the number of vocalic intervals.

Equation 1 Equation for $\Delta V$ (Ramus et al., 1999)

$\Delta C$ is the same as $\Delta V$, but for consonantal intervals:

$$\Delta C = \sqrt{\frac{\sum |x - \mu|^2}{N}}$$

where $\Sigma$ means "sum of", $x$ is the duration of a consonantal (intervocalic) interval, $\mu$ is the mean of the durations of all consonantal intervals, and $N$ is the number of consonantal intervals.

Equation 2 Equation for $\Delta C$ (Ramus et al., 1999)

$%V$ is the sum of vocalic interval duration divided by the total duration of the sentence/utterance, multiplied by 100:

$$%V = \frac{\text{sum of vocalic intervals}}{\text{sum of vocalic and consonantal intervals}} \times 100$$

Equation 3 Equation for $%V$ (Ramus et al., 1999)

Ramus et al. (1999) used these metrics to classify languages into their traditional groups, arguing that this grouping was not due to isochrony, but to the differing patterns in phonetic and phonological properties such as vowel reduction and syllable structure (see Section 3.1). They developed metrics to capture these phonological properties through their effects on the variability of vocalic and intervocalic intervals, which when measured, reveal the rhythm class of a language. Testing the value of their metrics by their success at classifying eight languages into their traditional classes, they found that a combination of $\Delta C$ as well as $%V$ or $\Delta V$ was successful in separating the languages into their traditional rhythm classes, and that when $\Delta C$ and $%V$ were graphed on one plane, three distinct groupings emerged (i.e. stress-, syllable- and
mora-timing) (Figure 3, reproduced from Ramus et al., 1999). This finding was contra to the idea of a rhythmic continuum, because they found that Polish and Catalan fell into distinct classes, which had previously been used to postulate a continuum (Dauer, 1983) or “mixed rhythm” (Nespor, 1990).

![Figure 3 Distribution of languages over the (%V, ∆C) plane (Ramus et al., 1999, p. 273)](image)

Error bars represent ± 1 standard error. Languages are English (EN), Polish (PO), Dutch (DU), Spanish (SP), Italian (IT), French (FR), Catalan (CA) and Japanese (JA).

Later studies using these metrics showed that ∆V and ∆C scores were highly affected by speech rate, such that speech samples of the same language at different rates would produce results comparable to those between different languages (Barry & Andreeva, 2001; Barry et al., 2003; Dellwo & Wagner, 2003). To address this, the rate-normalised metric VarcoV (Equation 4, Ferragne & Pellegrino, 2004; White & Mattys, 2007a, 2007b) and VarcoC (Equation 5, Dellwo, 2006, 2010) were proposed, and both were shown to be more consistent than ∆C and ∆V at discriminating traditional stress-timed languages from syllable-timed ones at all speech rates. VarcoV and VarcoC also calculate standard deviation, but then normalise for speech rate by dividing the standard deviation by the mean duration of vocalic/consonantal intervals across the whole speech sample (multiplying this result by 100, so that the result is not a fraction). However, Dellwo (2010) showed that speech rate affects the measures differently, with %V being least affected by differences in speech rate (Dellwo & Wagner, 2003) and therefore a stronger measure to describe rhythm.
VarcoV is a normalized version of $\Delta V$ where it is divided by the mean of the durations of all vocalic intervals within a sentence/utterance, multiplied by 100:

$$VarcoV = \frac{\Delta V}{\mu} \times 100$$

where $\mu$ is the mean of the durations of all vocalic intervals

Equation 4 Equation for VarcoV (Ferragne & Pellegrino, 2004; White & Mattys, 2007a, 2007b)

VarcoC is the same as VarcoV but for consonantal intervals:

$$VarcoC = \frac{\Delta C}{\mu} \times 100$$

where $\mu$ is the mean of the durations of all consonantal intervals

Equation 5 Equation for VarcoV (Dellwo, 2006, 2010)

3.2.2. Local metrics
The raw Pairwise Variability index (rPVI, Equation 6; Low, 1998; Low & Grabe, 1995) characterises a different way of measuring rhythm. Unlike global measures, the rPVI takes a local measure, capturing durational differences in successive vocalic or consonantal interval pairs. It calculates the difference between each successive interval, takes the absolute value (discarding negative signs), and sums these differences, dividing this by the number of all interval pairs (all intervals minus one). Like the previous metrics, the rPVI has a version for normalization (nPVI, Equation 7; Deterding, 1994; Grabe & Low, 2002). The nPVI calculates the absolute of the difference of each successive pair of intervals, divided by the mean duration of that pair, then divides the absolute by the number of all interval pairs (all intervals minus one), and multiplies this by 100 to produce a non-fractional result. The rPVI is used for consonants rather than nPVI, because they found that consonantal intervals were not as affected by speech rate as vocalic intervals, and they argued that the durational variation of consonantal intervals reflects language structure (i.e. longer consonantal intervals can reflect more complex syllable structures and consonant clusters), so normalising for them removes an important source of language rhythm. For vocalic intervals, they suggested using the nPVI instead as they are affected by speech rate. A ‘–V’ or ‘–C’ is added to the end of the nPVI or rPVI to denote which interval (vocalic or consonantal) is measured.
A higher rPVI or nPVI result reflects greater variability in differences between successive pairs, argued to be the result of phonological phenomena such as vowel reduction, which affects the rhythmic impression of a language (Section 3.1).

**rPVI** is the raw Pairwise Variability Index, which is:

\[
 r_{PVI} = \frac{\sum_{k=1}^{m-1} |d_k - d_{k+1}|}{(m - 1)}
\]

where \(d\) is the duration of the \(k\)th interval and \(m\) is the number of intervals

**nPVI** is the normalised Pairwise Variability Index, which is:

\[
 n_{PVI} = \frac{\left(\sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{2} \right| \right)}{(m - 1)} \times 100
\]

where \(d\) is the duration of the \(k\)th interval and \(m\) is the number of intervals

The reasoning behind a local measure is as follows: a local measure reflects patterns in variation, rather than just variation itself. To conceptualise, Low (1998; also, Low, Grabe, & Nolan, 2000) proposed imagining the results of global and local metrics on hypothetical languages A and B (Figure 4, reproduced from Low et al., 2000). If vowel variability were measured using a global metric such as VarcoV, the score would be high for both languages, reflecting the same amount of variability in vocalic interval durations. A local metric like the nPVI-V on the other hand, as it measures variation between successive pairs of vowels, would be able to distinguish between the languages by their different patterns. The nPVI-V would measure lower for Language A, as only one pair of successive vowels (the middle pair) shows a difference in duration, while Language B would produce a higher result, as all vowel pairs are different durations. The impression of rhythm is particularly affected by differences in duration when they vary successively (i.e. the difference between stressed and unstressed syllables is maximized, see Section 3.1). Local metrics capture this local variation that may be missed by global metrics.
The rPVI and nPVI have been used effectively to measure rhythm for understanding perceived rhythmic differences. For example, Low (1998) compared Singapore English to British English which had been described as having syllable-timing rhythm (Low, 1998; Low et al., 2000). Singapore English showed lower nPVI scores, suggested that it is more syllable-timed than British English. Use of the nPVI was also extended cross-linguistically in Grabe and Low (Grabe & Low, 2002) where both rPVI and nPVI were used on samples of 18 different languages (one speaker per language). They calculated both nPVI-V and the rPVI-C, plotting the languages in a plane defined by those two parameters. Their data “support a weak categorical distinction between stress-timing and syllable-timing … [but] … there is considerable overlap between the stress-timed and the syllable-timed group and hitherto unclassified languages” (Grabe & Low, 2002, p. 538). They thus argued for a rhythmic continuum. because while they did find that prototypically stress- and syllable-timed languages separated into the classes based on metric results, there were many other languages such as Greek, Malay, Romanian, Tamil and Welsh which were located in the middle between the two classes (see Figure 5, reproduced from Grabe & Low, 2002). Also, while Ramus et al. (1999) had placed Japanese in a separate group for mora-timing, Grabe and Low found it to be placed with the syllable-timed set.
Chapter 3: Rhythm Background

Metrics have successfully separated languages into traditional rhythm classes (e.g. Grabe & Low, 2002; Ramus et al., 1999; White & Mattys, 2007a) and been used to measure rhythm for many languages, such as:

- Arabic (Dockendorf, Almubayei, & Benton, 2008; Droua-Hamdani, Alotaibi, Selouani, & Boudraa, 2014; Mairano & Romano, 2011),
- Bulgarian (Barry et al., 2003; Stojanovic, 2013),
- Serbian (Stojanovic, 2013),
- Latvian (Bond, Markus, & Stockmal, 2007; Stockmal, Markus, & Bond, 2005),
- Hawaiian (Stojanovic, 2013),
- Tamil (Keane, 2006),
- Greek (Arvaniti, 2009; Grabe & Low, 2002),
- Mandarin (Lin & Wang, 2007; Mok, 2009),

Figure 5 PVI profiles for eighteen languages (Grabe & Low, 2002, p. 24)
Chapter 3: Rhythm Background

- Czech (Stojanovic, 2013),
- Korean (Arvaniti, 2009; Mok & Lee, 2008) and
- Cantonese (P. Mok, 2009).

For non-prototypical languages results have varied between different studies, with languages being placed in different classes depending on the study, most likely because of unstandardized methodology. Some languages have had mixed results for rhythm metrics, and have not been able to be placed into distinct classes by the metrics (Keane, 2006 for Tamil; Loukina et al., 2011 for Mandarin, Greek and Russian; Stockmal et al., 2005 for Latvian). Others lie in between the traditional classes, supporting the idea of a continuum with languages lying between stress- or syllable-based.

Metrics have been applied to describe differences and distinguish between dialects of the same language, such as Singaporean English and British English (Deterding, 2001; Grabe & Low, 2002; Low et al., 2000), French dialects (Cumming, 2011; Faygal, 2010; Kaminskaïa, 2016; Kaminskaïa, Tennant, & Russell, 2016), European and Brazilian Portuguese (Frota & Vigário, 2001), and various Arabic dialects (Hamdi, Barkat-defradas, Ferragne, & Pellegrino, 2004). They have also been able to distinguish in the same language between formal and colloquial varieties (Keane, 2006), speech of males and females (e.g. Kaminskaïa, 2016). Further, they have been very popular in areas of first language acquisition, such as discerning characteristics of child-directed speech (Bunta & Ingram, 2007; Lee, Kitamura, Burnham, & McAngus Todd, 2014; Payne et al., 2011) and language learning (Bond et al., 2007; Mok & Dellwo, 2008; Stockmal et al., 2005; White & Mattys, 2007a), particularly for studying foreign accent, speech intelligibility and bilingualism in situations of rhythm transfer and language contact (e.g. Bunta & Ingram, 2007; Mok, 2011; Thomas & Carter, 2005, 2006). Beyond all of these, they have even been useful in clinical application such as speech pathology (Liss, Utianski, & Lansford, 2013; Lowit, 2014; Selouani, Dahmani, Amami, & Hamam, 2012), automatic recognition of emotion (Ringeval, Chetouani, & Schuller, 2012) and spoken language identification (Timoshenko & Höge, 2007; Zhang & Glass, 2009). This growing body of research utilizing rhythm metrics suggests that they are valuable and informative for describing and comparing language rhythm.

Despite their widespread application, the usefulness of these metrics has been contested on the basis of problems with their underlying theoretical framework, tenuous links to the
Chapter 3: Rhythm Background

perception of rhythm, and also that they lack a standard methodology (e.g. Barry, Andreeva, & Koreman, 2009; Loukina et al., 2011; Loukina, Rosner, Kochanski, Keane & Shih, 2013; Turk & Shattuck-Hufnagel, 2013; Wiget et al., 2010). Much of this work also argues that metrics oversimplify the concept of rhythm by equating it with only durational patterns, despite a body of research indicating that timing in the acoustic signal is not the entire picture of rhythm but one facet. The following section (Section 3.3) presents a deeper discussion of the issues with rhythm metrics.

3.3. Issues Surrounding the Use of Rhythm Metrics
This section will discuss issues with rhythm metrics in the literature. These issues involve their relationship to rhythm perception, methodological issues and limitations on their explanatory power. The theoretical and practical limitations of metrics will be acknowledged, and discussion and possible solutions to these issues provided by outlining methodological decisions to be made in the Rhythm Metric Study in this thesis.

3.3.1. Rhythm Metrics and Speech Rate
It has been argued that speech rate may be a confounding factor that could be used to differentiate between languages in many studies and its effect is underestimated in rhythm metric literature (Arvaniti & Rodriguez, 2013; Dellwo, 2010). Research has drawn attention to the interaction of speaking rate with other phonetic and phonotactic features that are used to separate languages into rhythm groups (Dellwo, 2010). In this work, it has been proposed that as speaking rate is a perceptually salient feature of speech (p. 111), it may either cue or interact with rhythm correlates to give the impression that languages sound different. For example, it has been shown that faster speech is perceived as more syllable-timed, regardless of rhythm classification of a given language (Dellwo, 2010; Mok & Dellwo, 2008). Germanic languages (typically stress-timed) may also generally have slower speaking rates than Romance languages (typically syllable-timed) (Arvaniti & Rodriguez, 2013), such as Pellegrino et al. (2011) finding that syllable-timed Spanish had a faster speaking rate than stress-timed English. Rhythm metrics have been modified to take speech rate into consideration through normalization, and in the Rhythm Metric Study to follow, only these are used.

3.3.2. Focus on Duration, and Disregard of Rhythm as a Perceptual Entity
Metrics are often contested on the basis that they conflate the entire conception of rhythm with durational variability (Kohler, 2009a, 2009b, Loukina et al., 2011, 2013; Nolan & Asu, 2009; Turk & Shattuck-Hufnagel, 2013) although the phonological account of acoustic rhythm
suggests that it involves an accumulation and combination of many phonetic and phonological features, of which duration is only one part (Bertinetto, 1989; Dauer, 1983). The focus on duration may be the case because the rhythm metrics were originally developed to implement the phonological account of rhythm (Ramus et al., 1999). This basis led to empirically measuring the acoustic correlates of rhythm contributing to the rhythmic impression of a language, identified as syllable structure, vowel reduction and stress (Dauer, 1983). Low (1998) also developed rhythm metrics based on these correlates, in particular the durational variability associated with vowel reduction. Although the metrics have been applied to other acoustic correlates such as acoustic cues of prominence, the original durational-based metrics have been most popular. This may be due to their simplicity, ease of application, and widespread popularity.

Taking into account variation beyond duration can be both feasible and useful in speech rhythm research (see Fuchs, 2015). A growing body of work has explored the perception of rhythm in relation to the accumulation and combination of measurable acoustic properties besides duration, particularly those which cue the perception of prominence: changes in F0 (Barry et al., 2009), intensity (Cumming, 2008, 2011; Keane, 2006; Lee & Todd, 2004), modeled auditory prominence (Lee & Todd, 2004) and rate of spectral change (Kochanski, Loukina, Keane, Shih, & Rosner, 2010). Examples of studies which have applied these acoustic properties within rhythm metrics include, amongst others, intensity (Ferragne, 2008; Low, 1998) loudness (Fuchs, 2015), F0 (Cumming, 2010) and sonority (Galves, Garcia, Duarte, & Galves, 2002).

In one study comparing the rhythm of various dialects of British English, the syllable intensity was applied to the PVI, as well as the more common vocalic and consonantal durations, and it was found that variation in syllable intensity was more successful than vocalic and consonantal durations for differentiating between these dialects (Ferragne, 2008). A large project comparing British English to Indian English examined many rhythm correlates to determine that Indian English is more syllable-timed than British English (Fuchs, 2015). This classification was based on findings that Indian English showed less variability of vocalic durations, higher percentage of voiced durations over total utterance duration and less variation in sonority, intensity and loudness than British English, all argued in the phonological and perceptual account of rhythm to contribute to a perception of more syllable-timed rhythm in Indian English than British English. Another study using sonority measures found that Polish
Chapter 3: Rhythm Background

patterned with stress-timed languages on mean sonority variation but was intermediate between stress- and syllable-timed languages on mean sonority measures (Galves et al., 2002). Traditionally Polish had been classified as mixed, so this extra level of measurement was able to deepen its rhythm classification beyond that provided by durational measurement. Many studies examined rhythm alongside spectral dispersion (Low, 1998; Low et al., 2000; Tilsen & Johnson, 2008). Although the body of work applying rhythm metrics on acoustic phenomenon besides durational variability is fairly small, it has shown that it can be useful and feasible in rhythm research.

For many researchers however, rhythm is defined as a perceptual phenomenon (Couper-Kuhlen, 1993) and an argument against the metrics is that by focusing on the acoustics of speech they “disregard the listener” (Kohler, 2009b, p. 32). When discussing the relationship between rhythm metrics and perception, it is important to distinguish between rhythm as an acoustic entity and as a perceptual entity (Fuchs, 2015, p. 65). As mentioned in the introduction to this thesis (Chapter 1), this focus of this thesis is a description of the acoustics of rhythm in the speech signal, not perceptual rhythm. Regardless, it is still important to acknowledge this argument against the metrics.

Further, there is strong evidence that durational variation is perceived by both infants and adults (Jusczyk & Aslin, 1995; Nazzi et al., 1998; Nazzi & Ramus, 2003; Ramus, Dupoux, & Mehler, 2003; Ramus et al., 1999). Many of these experiments used a filtered speech technique called flat sasa which degrades speech samples so that only relevant phonological cues (e.g. timing variation) are preserved (Nazzi & Ramus, 2003; Ramus et al., 2003). In this technique, all consonantal intervals are replaced by [s] and vocalic intervals by [a], and fundamental frequency F0 modulation is transformed into a flat and slightly declining F0. It has been found that using only this information, listeners are able to discriminate between languages in separate rhythm classes (Ramus et al., 2003, 1999). These experiments have provided evidence that durational variation is both a perceptual phenomenon and also an acoustic one which can be measured through rhythm metrics.

Accent judgment experiments provide further support for the relevance of a connection between rhythm metric results and the perception of rhythm. It has been suggested that failure to master patterns of rhythm, timing and stress when learning a language can lead to foreign accent (e.g. Faber, 1986; Taylor, 1981). In one experiment English speakers were asked to
Chapter 3: Rhythm Background

make judgments of the “foreignness” of English spoken by native speakers of British English, Dutch or Spanish and it was found that the rhythm metrics VarcoV, %V and nPVI-V, as well as speech rate, were significantly correlated with the accent judgments (White & Mattys, 2007b). In a similar study, accent judgments of Korean learners of Japanese correlated with nPVI-V results (Kinoshita & Sheppard, 2011). Although rhythm has been argued to be an important facet of accent and intelligibility, it is not the only factor to account for foreign accent. For example, pitch contours and specific pronunciation of consonant and vowels may also play a role (Polyanskaya, Ordin, & Busa, 2016). Even so, that the metrics significantly correlate with perceptual judgments shows that they are relevant to perception.

As mentioned Chapter 1, this thesis is focused on describing the acoustics of rhythm in the speech signal, not describing perceptual rhythm. This project will focus on metrics which assess durational variability, as although metrics may be applied to a wide range of acoustic phenomena, this would be difficult to do with no knowledge of, or research into the correlates of prominence and stress in Kriol, i.e. in terms of speech production and acoustics. The decision in this project to focus on acoustic rhythm rather than perceptual rhythm, and to limit measurement to durational variation has two implications: firstly, the arguments against metrics on the grounds of lacking a perspective of perception are not currently relevant, and beyond the scope of this project to explore. Future work could investigate perception of rhythm in this variety (see thesis discussion Chapter 7). Secondly, as durational metrics are the most commonly employed type of metric in the rhythm literature, a study using these will contribute to the wider discussion around rhythm metrics as a tool for measuring rhythm, and allow comparability of rhythm with other studies. Though durational variability only represents one facet of rhythm, measuring it is a “necessary prerequisite to understanding rhythm in all its possible meanings” (Turk & Shattuck-Hufnagel, 2013, p. 94), a useful starting point for analysing an under-described language and sets the ground for future research.

3.3.3. Methodological Issues

A body of work draws attention to the fact that classifications of languages based on metric results often contradict each other, especially for non-prototypical languages (Arvaniti, 2009), and it has been widely argued that metrics cannot unequivocally classify these languages (Keane, 2006; Mok & Lee, 2008; Stockmal et al., 2005). For example, using metrics, Greek has been described as syllable-timed (Tsiartsioni, 2003), but also as having mixed rhythm (Baltazani, 2007), and by others as being unclassifiable (Grabe & Low, 2002). It has also been
Chapter 3: Rhythm Background

noted that metrics results can vary widely even for the same language, with Wiget and White (2010) pointing out that in a selection of studies using metrics (Grabe & Low, 2002; Ramus et al., 1999; White & Mattys, 2007a), Castilian Spanish measured a range from 30 to 42 for nPVI-V and Standard Southern British English from 63 to 73.

Variation in metric scores is likely to be due to a lack of standard methodology. Studies using metrics vary widely in respect to methodological choices, presenting many sources of variation. These sources include: elicitation methods, type of linguistic materials spoken, the number of speakers measured, and protocols followed for segmenting intervals. A selection of studies have assessed the impact of varying each source on metric results (e.g. Arvaniti, 2009, 2012a; Arvaniti et al., 2008; Wiget et al., 2010). The following is an outline of findings in this area, which has on the whole, shown that metrics are highly sensitive to variation in these sources, explaining variation in results and classifications.

It has been proposed that metrics are not robust to variation in type of linguistic materials, such as whether sentences measured vary in syllabic composition or whether they are naturalistic or elicited (Arvaniti, 2009, 2012b). Arvaniti (2009) tested the performance of ΔC, %V, PVIs and Varcos for on English, German, Greek, Korean and Spanish, with eight participants per language. Three elicitation methods were used: spontaneous speech, story reading and elicitation of carefully selected sentences. To test variation in linguistic material, there were three types of sentences chosen: one ‘typical’ set chosen from original works in each language, and two sets carefully selected to have more or less syllable structure complexity, which they called stress-timed and syllable-timed sets respectively. The study found three important results with regards to variation in elicitation method and linguistic material: firstly, when all elicitation methods were pooled, metric results were inconsistent with their traditional classes. Secondly, scores between languages were not statistically significant, even between those that are prototypical representations of different classes, such as between English and Spanish. Thirdly, the effects of elicitation method and sentence type were as large as the effect of language itself. In other words, every language could be classified as either stress- or syllable-timing, depending on the materials selected. They proposed that this was a reflection of metrics being too sensitive to linguistic materials, as a prototypical “stress-timed” language should measure this way using metrics regardless of materials used. That metrics were sensitive to linguistic material was also shown in Wiget et al. (2010) where the greatest variation in rhythm metric scores was accounted for by the sentences measured.
Chapter 3: Rhythm Background

In this study however, the sentences were not carefully selected but a pseudo-random set of five naturalistic sentences also used by White and Mattys (2007a).

The reasoning behind metric sensitivity to sentence type being a problem is somewhat unclear. The sentences/linguistic material in Arvaniti’s (2009) study were carefully chosen to reflect different rhythm types, therefore the fact that the metrics were able to reflect these choices suggests that they are indeed sensitive to rhythmic differences. This is what the metrics were developed for: to capture acoustic differences between languages to reflect phonological differences such as syllable complexity. In order to control or manage this sensitivity, studies that use metrics need to control the speech they measure so that it is representative of that language. In one such study, sentences used to represent English, Spanish and Catalan were controlled for phonological differences such as syllable complexity and when those differences were controlled, it was found that rhythm metrics were able to successfully discriminate between languages (Prieto, Vanrell, Astruc, Payne, & Post, 2012). Metric results for studies which have used spontaneous speech (e.g. Deterding, 2001; Thomas & Carter, 2006) should be compared to studies using elicited speech with caution.

Studies employing rhythm metrics also vary widely in the amount of speakers and material they measure, and it has been shown that the effect of interspeaker variability can often account for variability in rhythm scores beyond the languages themselves (Arvaniti, 2012b; Wiget et al., 2010). This has led to the suggestion that rhythm metrics can only pick up on differences between languages if used on a very large corpus with many speakers (Loukina et al., 2011). Studies measuring rhythm have ranges of speakers from as low as one or two per language (Barry et al., 2009; Grabe & Low, 2002) to as high as 40 to 50 (Benton, 2010; Dockendorf et al., 2008) and as much variation in between (12 in Arvaniti, 2012b; 10 in Low et al., 2000; four in Ramus et al., 1999; and six in White & Mattys, 2007a). One reason that speakers could introduce variability is because of variation in speech rate, mentioned earlier. Although normalized metrics like VarcoV, VarcoC and nPVI-V are unaffected by speech rate, it is important to keep in mind and also to measure as many speakers as possible.

Lastly, studies can introduce variation through the segmentation criteria of the data prior to running the rhythm metrics. For example, whether glides are counted as vowels, consonants or excluded altogether, and whether utterance-final segments are included in the measure (see
Chapter 3: Rhythm Background

Chapter 4 for further discussion). The protocols followed for segmenting data can greatly affect the metrics results (see Wiget et al., 2010 for a review).

The lack of standard methodology in the use of rhythm metric research has lead to contradictory results and difficulty comparing between studies. To address this, the present project will be as transparent as possible in outlining methodological choices employed, and also consider closely recommendations made by researchers in this area. The following seven “best practice” recommendations for research using metrics was presented by Wiget et al. (2010), based on a wide review of the literature:

1. %V, VarcoV, and nPVI-V are robust to variation in articulation rate and are effective at discriminating between language varieties previously held to differ in terms of contrastive rhythm. However, as all rhythm metrics have limitations, it is safest to use %V in combination with either VarcoV or nPVI-V rather than rely on a single metric.

2. Results for non-rate-normalized metrics ∆V, ∆C, and rPVI-C are difficult to interpret and not reliably discriminative, likewise all metrics of consonantal interval variation VarcoC, in addition to the non-rate-normalized ∆C and rPVI-C. Furthermore, they show relatively poor consistency between speakers, sentences, and measurers.

3. Single-speaker or low-N studies should absolutely be avoided where speakers are intended to be representative of a particular linguistic group. Contrastive rhythm metrics may be useful in single-speaker longitudinal studies, however.

4. Rhythm scores are strongly affected by the particular linguistic materials used. Either a large sample of sentences should be used or materials should be constructed to be representative of the relevant phonological and metrical properties of the language under study.

5. Where several measurers are used, they should work according to an agreed protocol for the identification of segment boundaries. Furthermore, discussion and comparison of difficult cases between measurers, which were avoided in the current study, should help minimize variation in rhythm scores

6. Contrastive rhythm scores obtained through automatic alignment show good agreement with those obtained from human measurers, assuming sufficient training data are available for the language in question. Of course, use of automated methods would allow much larger sampling of speakers and sentences.

7. Do not rely too heavily on contrastive rhythm metrics or over-interpret the results of studies that use them. They merely provide an approximate indication of the degree of temporal stress contrast in a language and are susceptible to extraneous variation from multiple sources (Wiget et al., 2010, p. 1566).

3.3.4. Issues of Interpretation

The last issue with rhythm metrics discussed in this chapter is that metric results may be limited in their explanatory power, making them difficult to interpret in certain situations
Chapter 3: Rhythm Background

(Arvaniti et al., 2008; Loukina et al., 2011; Turk & Shattuck-Hufnagel, 2014). For example, there may be a situation where two different languages score the same nPVI-V measure. However, inspection of the actual variability may show that it is produced from different reasons, such as vowel reduction in one language and syllable lengthening related to prosodic structure in the other, or even variability due to things such as speech type, sentence type, segmentation protocol or speaker variability. The metrics do not alone provide direct explanation as to how they reach their result. It has been suggested throughout the literature that a description of rhythm in a language must not rely solely on metric results (Arvaniti, 2012b; Wiget et al., 2010), and it has been proposed that additional, further studies are required to explore more deeply the relationship between rhythm metric results and language variability by relating measurements to grammatical features of the languages and testing specific hypotheses (Turk & Shattuck-Hufnagel, 2013, p. 109). The current project seeks to do this with a second study, the Vowel Reduction Study, which builds on the description of rhythm in Barunga Kriol beyond the rhythm metrics, by exploring vowel reduction patterns in this variety.

3.4. Conclusion

This chapter has outlined literature relevant to describing rhythm at the acoustic level, outlining a variety of perspectives on the concept of rhythm and culminating in a discussion of the development and use of rhythm metrics. Despite criticism of the metrics, they are widely applied in linguistic research for describing and distinguishing between rhythm of languages, understanding differences in dialects and extending description of rhythm in previously undescribed languages or languages in which their rhythm had been controversial. With careful theoretical and methodological considerations, rhythm metrics can be useful and informative as a first step to describe the rhythm of a language.

To follow, Chapter 4 will outline the methods, results and conclusion of the Rhythm Metrics Study, the first study in this thesis.
Chapter 3 provided the theoretical framework and background for the first study in this thesis, the Rhythm Metric Study. Chapter 4 presents this study: a comparison of rhythm in Kriol and English using rhythm metrics. In Section 4.1 hypotheses are outlined, Sections 4.2-4.3 present the study method and results methods and 4.5 a discussion of limitations and directions for future research.

4.1. Research Hypotheses
By describing the rhythm of Barunga Kriol and Australian English using rhythm metrics, this study followed literature suggesting that durational rhythm metrics are a practical starting point for describing rhythm in an under-described language. This study compared the durational variability of vocalic and consonantal intervals as well as the overall ‘vocalic-ness’ of Kriol and English. The aim was to investigate whether Kriol tends towards syllable- or stress-timing on the rhythmic continuum in relation to English, which is traditionally classified as stress-timed.

In Chapter 2 it was suggested that possible substrate languages of Kriol show correlates of stress-timing, including vowel reduction. In work by Fletcher and Butcher (2014) on the rhythm of Bininj Gun-wok (historically spoken in some proximity to Barunga, and Barunga Community members have historically spoken Mayali, a dialect of Bininj Gun-wok), it was found that though this language had previously been informally described as stress-timed, it actually had lower nPVI-V scores than British English for variability. This suggested that it was more syllable-timed than English. As Kriol’s historically English lexicon is understood to be mixed with grammatical, semantic and phonological structures thought to come from substrate languages (and Bininj Gun-wok is likely to be one of these, based on geographical location), it is possible that rhythm may also have transferred. Therefore, it is expected that Kriol would be more syllable-timed than Australian English.

In order to add to the literature on rhythm in both Kriol and English, the following research questions were explored using rhythm metrics on storytelling data.
Chapter 4: Rhythm Metric Study

4.1.1 Research Question 1
How variable are vowel and consonant intervals in Kriol compared to English?

The hypothesis for this research question was based on the phonological account of rhythm where ‘stress-timed’ languages show more vowel reduction and a greater number and more complex consonant clusters than ‘syllable-timed’ languages (see Chapter 3). Using the rhythm metrics VarcoV, VarcoC, and nPVI-V, presence of temporal vowel reduction and more consonant clusters would be reflected in higher scores, i.e. more variability in vowel and consonant durations. As English is traditionally regarded as having ‘stress-timed’ rhythm, Kriol was expected to be more ‘syllable-timed’, with lower durational variability for both vowel and consonant intervals than English. Research Question 1 was answered by measuring vocalic variability using VarcoV and nPVI-V, and consonantal variability using VarcoC, on speech data from both Kriol and English.

4.1.2. Research Question 2
How ‘vocalic’ is Kriol compared to English?

Because stress-timed languages are expected to have vowel reduction and greater amount of consonant clusters, these languages are also expected to have less vocalic material overall than syllable-timed languages. It was expected then, that Kriol would have higher vocalic-ness than English, measured using %V. Research Question 2 was answered by measuring the ‘vocalic-ness’ (%V) of Kriol and English.

Justification for the choice of rhythm metrics used in this project was provided in Chapter 3, Section 3.3, where it was suggested that using a combination of both global and local metrics can provide greatest explanatory power to the results, and that rate-normalised vocalic metrics (VarcoV and nPVI-V) are most robust to speaking rate, offering the greatest discrimination between languages perceived to have different rhythm. For English in particular, stress is known to cause variation in vowel duration (e.g. Klatt, 1975; Sluijter & van Heuven, 1996), so vocalic metrics are most appropriate. %V is included because it is useful for differentiating between languages with different constraints on syllable complexity (Ramus et al., 1999). This measure reflects this relationship, as the higher complexity and amount of consonant clusters, the lower the proportion of vocalic duration there is in the speech sample. It also reflects vowel reduction patterns, with languages having more reduction having a lower
Chapter 4: Rhythm Metric Study

proportion of vocalic material. %V has also been reported to be robust, with little correlation to speech rate (Dellwo & Wagner, 2003; Wiget et al., 2010).

The consonantal metric (VarcoC) was also used, despite reports that it can be unreliable and difficult to interpret (e.g. Wiget et al., 2010). There were two reasons for this inclusion: firstly, seminal studies measuring rhythm have used consonantal metrics (e.g. Grabe & Low, 2002; Ramus et al., 1999; White & Mattys, 2007a) so this will allow cross-linguistic comparison in future research. Secondly, the phonological account of rhythm (Bertinetto, 1989; Dauer, 1983, 1987) outlined variation in consonants cluster complexity as being linked to perceived rhythmic differences, so it makes sense to measure for these too. Consonantal metrics therefore, were used though their results will be treated with caution. The local measure of consonantal variability (rPVI-C) was not used as it is not robust to speech rate, following Wiget et al., (2010) recommendation not to use it (see Chapter 3).

The method is outlined in Section 4.2 below.

4.2. Method

4.2.1. Participants
The participants were ten female speakers within the ages of 20-34 years. These were made up of two language groups: five speakers of Barunga Kriol (residing in Barunga Community, NT) and five speakers of Standard Australian English (residing in, and born or lived in Victoria most of their lives). The English speakers were all monolingual, while the Kriol speakers are fluent speakers of Kriol and English (acquisition history unknown).

4.2.2. Data Collection
The Kriol dataset used in this project is a subset taken from the ‘Conversational Corpus of Barunga Kriol’, a corpus of Kriol collected in 2014-2017 with speakers at Barunga, NT. This subset consisted of spontaneous storytelling, in which participants were asked to “tell the story” using as stimulus two word-less picture books. These books were The Monster Story (O’Shannessy, 2004) and Frog, Where Are You? (Mayer, 1969). Storytelling data is a suitable type of data for studies of rhythm (Ladefoged, 1997).

Participants were given time to familiarize themselves with the books prior to recording, in order to avoid disfluencies. During the recording, if a participant stalled in their storytelling, they were prompted to continue. Otherwise, they were left to speak naturally and for as long as
they wished. The order of the books was not controlled. A new dataset was recorded for the English comparison group, matching the speaking style and session conditions as closely as possible.

4.2.3. Recordings
Recordings were made using an Olympus LS-14 linear PCM recorder with lapel microphone, at 44.1kHz and 16-bit. The Kriol recording sessions were conducted in quiet, outdoor field conditions by a familiar local Aboriginal peer from Barunga Community. The English sessions were conducted by the researcher in quiet conditions at the researcher’s home and in a recording studio. Recordings were transferred from SIM card to a laptop computer for transcription, processing and analysis.

4.2.4. Data Preparation
The recordings were orthographically transcribed in ELAN, the English recordings by the researcher and the Kriol recordings by non-native speaker linguists familiar with the language, and checked by a native speaker. These transcriptions were then exported to Praat TextGrids (Boersma & Weenink, 2012) and run through web-based Munich Automatic Segmentation (Kisler et al., 2016), a forced alignment tool that produced Praat TextGrids with tiers for time-aligned word and phoneme boundaries. This tool was created by training acoustic models on large corpora and lexicons of major languages. As input it takes a sound recording and an orthographic transcription of the speech, from which it estimates the most likely pronunciations and produces phonetic labeling and segmentation in TextGrid format. For the English data in this project, there was a MAUS model trained on Australian English that was available for use. For Kriol on the other hand, there was no Kriol-trained model available and so a model trained on Italian was used instead. The Italian model was more robust at segmentation than either the Australian English model or the language-independent model, which combines training on many languages.

For the English and Kriol data which had been automatically aligned, the portion of speech data pertaining to storytelling (excluding introductions or other dialogue, usually about 20 seconds into the recording) was manually checked. Phoneme boundaries which had been incorrectly placed by MAUS (e.g. due to background noise, fast speech, inaccurate transcription, etc.) were shifted manually in Praat. Particular attention was placed on boundaries between vowel and consonant intervals. The final phonetic labels were
automatically produced by MAUS (based on grapheme-to-phoneme rules within the language model), and only corrected if a vowel was mislabeled as a consonant, or a consonant as a vowel; if labels themselves were not correct but within the correct grouping (consonant or vowel) they were left as found.

The work was done by primarily visually inspecting the spectrograms and speech waveforms in Praat, with reference made to standard linguistic labeling criteria (Peterson & Lehiste, 1960). For the boundaries between consonants and vowels, attention was predominantly paid to the onset and offset of regular pitch periods in the waveform as well as vowel related intensity in higher frequencies as judged in the spectrograph. In the interests of transparency, consistency and reproducibility, detailed segmentation criteria followed with regards to these boundaries as well as those surrounding various types of consonants can be found in Appendix 1.

In the case of ambiguity in the visual inspection, and to make sure that boundaries chosen were not completely erroneous, the segment was checked. In the case of gross ambiguity unresolved by listening, the rule of thumb was to remain conservative and leave the forced alignment output produced by WebMAUS as it was found.

Silent pauses between and within utterances were excluded from all analysis, with vowel-vowel or consonant-consonant segments on either side of a pause summed into one interval (as in Grabe & Low, 2002; White & Mattys, 2007a; Wiget et al., 2010). The purpose of this was to avoid subjectivity, particularly when it was unclear whether these periods were part of the closure of voiceless plosives. Lastly, due to the nature of using a stimulus for spontaneous speech, occasionally there were filled pauses present, such as when speakers were searching for words to name a pictured object or continue their story. These filled pauses were excluded, as speech with unnaturally lengthened durations could influence the rhythm metric results in ways that do not reflect its rhythmic nature (as in Arvaniti, 2012c; Deterding, 2001; Tan & Low, 2014).

4.2.5. Analysis
Using a script performed in R Version 1.0.153 (R Core Team, 2016), the manually corrected phonetic segmentation (both labels and durations) was then extracted and the phoneme labels were transformed into vocalic (“V”) and consonantal (“C”) intervals. Where there were two or more adjacent vowels, the boundaries between them were dissolved into one interval, as done
in other rhythm metric studies (Grabe & Low, 2002; White & Mattys, 2007a; Wiget et al., 2010). The same was done for consonant clusters. Using the same script, the speech was then divided into utterances and each utterance was saved as a smaller file. An utterance was defined as an amount of speech bordered by a boundary pause (defined as silence or non-speech noise of 150ms or more) on either side. Earlier studies have used various other methods for defining utterances, including divisions based on duration such as seconds and minutes (Arvaniti, 2012b), prosodic units (Faygal, 2010), syllable number (Tan & Low, 2014) or full sentences (P. Mok & Dellwo, 2008). The decision in the present study to use a reproducible definition was made after Grabe and Low’s (Grabe & Low, 2002, p. 252) to take as few “subjective and intuitive decisions as possible when taking measurements”. For each speaker, the first 4 utterances were excluded, allowing for a speaker’s period of adjustment to the recording situation.

Some studies (e.g. White & Mattys, 2007a) excluded glides and approximants /w/, /l/, /j/, and /r/ as their status as vowels or consonants can be ambiguous. Others prioritized acoustic criteria over phonetic or phonological criteria, with glides and approximants treated as consonants if they could be distinguished by clear changes in formant structure or intensity in the spectrogram, or as vowels if the formant movements were seamless (no friction) and no different to any adjacent vowels (Arvaniti, 2012b; Grabe & Low, 2002; Payne et al., 2011; Ramus et al., 1999). Mostly, these parameters meant that glides in prevocalic position were considered consonants, and post-vocically as vowels, and rather than allow ambiguity in these decisions, this study followed a rule and labeled all glides as consonant or vocalic based on these positions (as in Payne et al., 2011; Ramus et al., 1999).

To achieve regularity and consistency, utterance-initial consonants were excluded (those following pauses greater than 150ms) (White & Mattys, 2007a), as often it was not possible to ascertain where the closure was first made. Some studies (Fuchs, 2015; Gabriel & Kireva, 2014) added pre-determined durations ranging from 30ms-50ms to the start of these consonants based on the average length of consonants in their languages. I chose not to do this, because of limited information about utterance-initial consonants in Kriol and because consonant length may not be equal in the two languages. In this study, these intervals were excluded altogether.
Chapter 4: Rhythm Metric Study

With regards to utterance-final (i.e. pre-pausal) C or V segments there are differences of opinion in the literature. Due to possible lengthening effects in these segments (e.g. Klatt, 1976), some studies argue that prosodic lengthening processes such as accentual lengthening as well as variation in the type and duration of lengthening may be language-specific and contribute to the perception of their different rhythms (Grabe & Low, 2002; Mok & Dellwo, 2008; Payne et al., 2011; White & Mattys, 2007a). Thus they included these segments in their analysis. Other studies (Bertinetto, 1989; Bertinetto & Bertini, 2008; Gibbon & Gut, 2001) excluded these segments, arguing, “this portion has an entirely different rhythmic behaviour, that should best be analyzed on its own” (Bertinetto & Bertini, 2008, p. 2). Indeed, if rhythm varies within utterances, and with the final rhythm noted as different to the rest of the utterance, then it makes sense to analyse the final part of an utterance separately.

For the present study, it was decided to exclude this segment, for various reasons. Firstly, it was noticeable to the researcher on listening to the recordings that some speakers would lengthen this segment considerably, while they thought of the next thing to say. This was more like a hesitation and filled pause rather than language-specific rhythm, and could be specific to the storytelling style of the speech. Also, due to the spontaneous nature of the speech, there were uncontrollable speech quirks and background sounds that were difficult to segment. It was found during processing that many utterance-final consonants were voiceless and unreleased, despite sounding like they were present. Thus, the final boundary was often impossible to determine and it was difficult to write boundary protocols for this situation. In the interests of providing both an accurate reflection of language-specific rhythm rather than that influenced by storytelling in particular, as well as segmentation reliability (as final boundaries were otherwise inconsistent), it was decided that these segments (final C or V) would be excluded for all speakers and utterances.

One further methodological consideration pertaining to utterance-final segments was relevant. Loukina et al. (Loukina et al., 2011) ran a study measuring the effects of including or excluding this segment on the rhythm metric results. Finding that there was no difference in results for either methodological choice, this seemed like a good reason to believe that in the present study it would not matter whether I included or excluded this segment. Despite this finding, the present study considered that this methodological choice could have an effect on results, because of the reasons presented above, and also because the languages in this research are actually different languages and not dialects of the same language (see Chapter 2). So,
although it was decided to exclude these segments (previous paragraph), to be sure that they were not a confounding/correlated factor in any differences found between the languages (i.e. to be sure that the final segment was not super important in differentiating between the rhythm of each language), this present study ran the metrics on both data types (exclude and include). The results in Section 4.4 of this chapter will present results for data with utterance-final segments excluded, and for the reader’s interest and for consideration in future research, the results for data with the final segment is contained in Appendix 2.

Following all exclusions, there were 233 utterances in Kriol and 252 in English available for analysis. An R script extracted vocalic and consonantal interval durations from each utterance and used these to calculate the following per utterance (see Chapter 3 Section 3.2 for the formula for each measurement):

1. Global measures of variability in vocalic and consonantal durations: VarcoV and VarcoC
2. Local measure of variability in vocalic durations: nPVI-V
3. Measure of ‘vocalic-ness’: %V

A last note: This study calculated a metric value for each utterance per speaker and for plotting purposes took the mean of all utterances to calculate one value per speaker, and the mean of all speakers to calculate one value per language. The purpose of dividing the speech into utterances and calculating a metric value per utterance was twofold: firstly, it is the most common method of calculating rhythm metrics. Most of these studies reviewed however, used elicited sentences rather than spontaneous or natural speech as in this study, and kept utterances or sentences of similar lengths or types. In this study as utterances were divided by pauses, not all utterances were complete sentences or intonational phrases, and they also ranged in size. The rhythm metrics have been shown to be sensitive to sentence type and length. As utterance sizes couldn’t be controlled, major deviations in size from the mean utterance size may cause the metric results to be very different. Other studies in this situation chose utterances of roughly comparable length (Arvaniti et al., 2008; Ramus et al., 1999), but because speakers have different speech rates in spontaneous speech this method not possible. Tan and Low (2014) chose similar utterances consisting of eight syllables. In the present study, utterance length ranged from 10 to 64 for total C and V segments (mean of 19), so selecting utterances of a particular size may have resulted in a large loss of useable data. As such, very
small utterances, with less than 4 tokens were excluded. To avoid further weighting of the metric scores due to different utterance sizes, the speaker means were calculated from all the utterance scores for that speaker. Language scores were then calculated from the mean of all speaker scores.

4.4. Results

Statistical analysis was applied to the data to test whether any rhythm metric scores between English and Kriol were significantly different. A linear mixed effects model with random effects (Baayen, Davidson, & Bates, 2008) in the package lme4 (Bates, Maechler, Bolker, & Walker, 2015) was performed for each rhythm metric (dependent variable: VarcoV, VarcoC, nPVI-V or %V). The model included as fixed effects (IV): Language (Kriol vs. English) and as random effects: Speaker and Storybook.

To test for significance of each effect, an ANOVA was conducted to compare two linear mixed effects models: a null model without the factor of interest (i.e. Language) against a complete model with the factor. This section will present results for data excluding utterance-final segments, as mentioned in Section 4.2.5 above. For data with the inclusion of utterance-final segments, see Appendix 2.

Residual plots of all full models were visually inspected for deviations from homoscedasticity or normality and no violations of model assumptions were found.

The results of lmerTest on a linear mixed effects model with fixed effects of Language and dependent variable as metric (VarcoV, VarcoC, nPVI-V or %V) are shown in Table 1 for both languages. The mean (M) and standard error (SE) here and in all descriptive statistics tables in this chapter are estimates of the class means that would be expected had the groups been of equal size, established by ‘least squares means’ (Harvey, 1960) using the lmerTest package (Kuznetsova, 2017). Significances (p values less than 0.05) are in bold in all tables.

Boxplots for speaker variation are included in Appendix 3 but besides being included as a random factor in the linear mixed effects models, are not analyzed for their contribution this study.
Chapter 4: Rhythm Metric Study

<table>
<thead>
<tr>
<th>Measure</th>
<th>Language</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kriol (N=233)</td>
<td>English (N=252)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td>VarcoV</td>
<td>81.594</td>
<td>2.353</td>
<td>79.627</td>
<td>2.291</td>
</tr>
<tr>
<td>VarcoC</td>
<td>61.010</td>
<td>2.003</td>
<td>63.428</td>
<td>1.961</td>
</tr>
<tr>
<td>nPVI-V</td>
<td>61.306</td>
<td>2.033</td>
<td>71.617</td>
<td>1.997</td>
</tr>
<tr>
<td>%V</td>
<td>48.279</td>
<td>0.757</td>
<td>43.259</td>
<td>0.735</td>
</tr>
</tbody>
</table>

Table 1 Descriptive Statistics for Rhythm Metric Scores by Language (estimates from statistical models)

Significance codes: ‘****’ is p<0.001, ‘**’ is p<0.01 and ‘*’ is p<0.05.

4.4.1. Variability of Vocalic and Consonantal Intervals in Kriol and English

It was expected that English would show higher variability in vocalic and consonantal intervals than Kriol, indicated by higher measures in VarcoV, VarcoC and nPVI-V.

4.4.1.1. VarcoV

The expected hypothesis for VarcoV was not supported: VarcoV scores were not higher for English than Kriol. Figure 6 is a box plot for VarcoV scores by Language, showing that the English and Kriol results were the same.

There was no effect of Language on VarcoV results ($\chi^2(1) = 0.3507$, p=0.5537). This was established by an ANOVA comparison, which showed that a full model including Language as a fixed effect and VarcoV as a dependent variable did not fit the data better than a null one without Language.

Figure 6. VarcoV Scores for Kriol and English
Chapter 4: Rhythm Metric Study

4.4.1.2. VarcoC

The expected hypothesis for VarcoC was not supported: VarcoC scores were not higher for English than Kriol. Figure 7 is a box plot for VarcoC scores by Language, showing that the English and Kriol results were the same. There was no effect of Language on VarcoV results ($\chi^2(1) = 0.7207, p=0.3959$).

![Varco scores for Kriol and English](image)

Figure 7 VarcoC scores for Kriol and English

4.4.1.3. nPVI-V

The expected hypothesis for nPVI-V was supported: nPVI-V scores were higher for English than Kriol. Figure 8 is a box plot for nPVI-V scores by Language, showing that the English results were higher than Kriol. There was an effect of Language on nPVI-V results ($\chi^2(1) = 8.2479, p=0.00408$). In Table 1 above it can be seen that English was higher than Kriol in nPVI-V score by 10.31 ($t(9.5)=-3.62, p<0.01$).
4.4.2. Vocalic-ness in Kriol and English - %V

The expected hypothesis for %V was supported: %V was higher for Kriol than English. Figure 9 is a box plot for nPVI-V scores by Language, showing that the %V for Kriol was higher than English. There was an effect of Language on %V ($\chi^2(1) = 11.381$, $p=0.0007419$). In Table 1 above it can be seen that Kriol was higher than English for %V by 5.21% ($t(8.4)=4.76$, $p<0.01$).
Chapter 4: Rhythm Metric Study

%V for Barunga Kriol and Australian English

Figure 9 %V for Kriol and English

4.5. Discussion and Conclusion
In this study, rhythm metrics were calculated on storytelling data to compare Kriol and English. The following two research questions were answered:

RQ1: How variable are vowel and consonant intervals in Kriol compared to English?

RQ2: How ‘vocalic’ is Kriol compared to English?

These research questions were based on a body of literature suggesting that language rhythm can be determined using durational-based rhythm metrics, and languages can be placed on a rhythmic continuum between syllable- and stress-timing depending on the behaviour of their vocalic and consonantal intervals (see Chapter 3). In this theory, stress-timed languages show greater variability and less vocalic-ness than syllable timed languages.

RQ1 was answered using rhythm metrics VarcoV, VarcoC and npVI-V. Kriol was expected to be more syllable-timed than English. Therefore, it was expected that English would score higher than Kriol on all measures, reflecting more variability in the vocalic and consonantal intervals at a global level, and vocalic intervals at a local level. Kriol and English were not found to be significantly different for VarcoV and VarcoC, suggesting that they are not different in the
Chapter 4: Rhythm Metric Study

variability of vocalic and consonantal intervals at a global level. However, English had a significantly higher score for nPVI-V than Kriol, suggesting more variability in the durations of vocalic durations than Kriol at a local level. This result suggests that Kriol is more syllable-timed than English on a local level.

RQ2 measured ‘vocalic-ness’ (%V), which was expected to be lower in stress-timed languages due to vowel reduction in unstressed syllables, as well as more complex consonant clusters causing consonants to take up a greater proportion of speech. %V was found to be significantly higher for Kriol than English, suggesting that Kriol has a greater proportion taken up by vowels and that it might be more syllable-timed than English. %V is a strong metric for discriminating between languages (Chapter 3), and while %V score alone is not enough to classify a language on the stress- and syllable-timing continuum (see Wiget recommendations about combinations of metrics, in Chapter 3), in conjunction with the significant difference found for nPVI-V, there is some evidence that Kriol may be more syllable-timed than English.

Table 2 is a summary of results found in the rhythm metric study. In this table, areas marked with the colour green and letter ‘Y’ denote that significance was found for the effect of that independent variable on dependent variable (i.e. = comparing null vs. full model was significant). Areas marked with the colour red and letter ‘N’ denote no significance found for the effect of that independent variable on dependent variable. The direction of the effect is noted after the ‘Y’ or ‘N’ with the following letters: ‘E’ for ‘English’ and ‘K’ for ‘Kriol’.

<table>
<thead>
<tr>
<th>RQ1: Variability of Vocalic and Consonantal Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>VarcoV</td>
</tr>
<tr>
<td>VarcoC</td>
</tr>
<tr>
<td>nPVI-V</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RQ2: Vocalic-ness</th>
</tr>
</thead>
<tbody>
<tr>
<td>%V</td>
</tr>
</tbody>
</table>

Although Kriol and English were not significantly different for VarcoV and VarcoC, the findings suggest that Kriol may be more syllable-timed than English, due to the fact that English has a higher variability between successive vocalic intervals (nPVI-V) and that Kriol has a higher proportion of vocalic material (%V). The strong discriminative ability of %V has been
supported in studies as by White and Mattys (White & Mattys, 2007a) who found that %V was able to successfully discriminate between two stress-timed languages, Dutch and English (see Chapter 3 for more examples).

There was no difference found between the languages for VarcoV or VarcoC, which did not match my expectation. As mentioned in Chapter 3, Section 3.2.2, the local metric nPVI-V was developed as a response to the likelihood that VarcoV and VarcoC would inadequately capture variation differences between languages when there were differences present. By localizing variation at successive intervals rather than taking a normalized standard deviation of variation across an entire speech sample, it is a stronger measure of variability. Studies such as Low (1998) found that the PVI was a more successful rhythm metric than VarcoV for distinguishing between stress-timed British English and more syllable-timed Singapore-English. That VarcoC was not successful at discriminating between languages was also found in White and Mattys (2007b) who noted that the rate normalization incorporated into the measure may have eliminated the distinctions between their languages. It has also been recommended that VarcoC not be used it is difficult to interpret, not reliability discriminative and show relatively poor consistency between speakers, sentences and measurers (Wiget et al., 2010). Therefore, it is possible that in this study VarcoV and C did not capture rhythmic differences that may have been present between Kriol and English, but it may also be possible that these differences do not exist at a global level (despite the %V result).

A possible limitation of the methodology and analysis of this study is that outliers in the data were not inspected. The results for rhythm metrics were pooled across all utterances, without analysis undertaken for the variation across utterances or speakers. While 'Speaker' was included as a random effect in the linear mixed effects models as a possible source of variation, variation from particular utterance characteristics such as number of words or segments was not considered. It is also possible that the rhythm measurements varied across each speech sample in a systematic way, such as if speakers spoke slower at the start of recording and faster at the end when they may have been more comfortable with the recording situation. For more homogenous results and perhaps greater discriminative validity, outliers could have been investigated, and utterance characteristics included in the models.

Another possible limitation with the methodology is that speech rate was neither controlled nor measured. As mentioned in Chapter 3, has been argued that speech rate can be used to
differentiate between language rhythm as it can interact with phonetic and phonotactics features in a language to make them sound different, and this would make it a confounding factor (e.g. Dellwo, 2010). For example, it has been shown that faster speech is perceived as more syllable-timed, regardless of rhythm classification of a given language (Dellwo, 2010). Germanic languages (typically stress-timed) may also generally have slower speaking rates than Romance languages (typically syllable-timed) (Arvaniti & Rodrriquez, 2013), such as Pellegrino et al. (2011) finding that syllable-timed Spanish had a faster speaking rate than stress-timed English. Compared to stress-timed languages, syllable-timed languages have also been found to have a higher number of phonemes per syllables (Fenk-Oczlon & Fenk, 2006). In the planning of this study, speech rate was considered by using only rhythm metrics that incorporate speech rate normalisation. Despite findings that VarcoV and VarcoC were not different, Kriol seems, on listening, to be spoken at a much faster speech rate, causing it to sound more syllable timed. This impression may be explained by the “Gabbling Foreigner Illusion” (Cutler, 2012, p. 338). As the perception of rhythm interacts with speech rate, measuring speech rate would benefit future studies, and might give a clearer picture of differences found by rhythm metrics.

As mentioned in Chapter 3, metric results can be limited in explanatory power as they can suggest differences in rhythm, but are not able to provide any explanation to the cause of variability (Arvaniti, 2009; Turk & Shattuck-Hufnagel, 2013). For example, although vowel reduction was a main factor in the development of rhythm metrics (they sought to measure rhythm as a function of vowel reduction), the metrics don’t necessarily indicate that vowel reduction is present. It is possible that variability in durations may be caused by factors other than vowel reduction, such as syllable lengthening or consonant reduction. It has been recommended that descriptions of rhythm should not rely on metric results alone but instead should relate metric results to language features by testing specific hypotheses (Turk & Shattuck-Hufnagel, 2013). As rhythm is more than just temporal variability, the Study 2 in this thesis, the “Vowel Reduction Study” (Chapter 6), builds on the Rhythm Metric Study by exploring patterns of vowel reduction in Kriol and English.
Chapter 5: Vowel Reduction Background

In the Rhythm Metric Study (Chapter 4) some rhythmic differences between Kriol and English were found, and Study 2 explores whether patterns of vowel reduction might be behind these differences. In the phonological account of rhythm, vowel reduction plays a role in the rhythmic perception of a language, providing a means for unstressed syllables to be distinguished from stressed syllables (see Chapter 3). As the Rhythm Metric Study found that Kriol might show a tendency towards syllable-timing compared to stress-timed English, it is possible that Kriol may have different patterns of vowel reduction to English.

This chapter provides a background to vowel reduction, discussing research relevant to the second study in this thesis, the Vowel Reduction Study (Chapter 6). Section 5.1 provides a definition of vowel reduction and outlines the acoustic correlates of vowel reduction. Section 5.2 continues by describing where vowel reduction occurs. Finally, the chapter is concluded in Section 5.3.

5.1. Definition of Vowel Reduction and Associated Acoustic Correlates
Reduced vowels are termed ‘reduced’ because they are not realized to their full extent. Reduced vowels occur in unstressed positions and are less perceptually salient than full (stressed) vowels. Reduced vowels are correlated with reduced duration and shifts in spectral characteristics towards a mid-central, neutral point in the F1 by F2 vowel space (Aylett & Turk, 2004; Bell et al., 2003; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; de Graaf & Koopmans-van Beinum, 1984; Koopmans-van Beinum, 1980; Lindblom, 1963a, 1990; Moon & Lindblom, 1994; Wright, 2004). This central position in the vowel space is also known as schwa, and unstressed vowels realized towards this position in the vowel space are often the product of reduced subglottal pressure, decreased muscular energy and lowered coordination of gestures (Browman & Goldstein, 1990, 1992; Harrington, Cassidy, Fletcher, & McVeigh, 1993; Lindblom, 1963a; Olive, Greenwood, & Coleman, 1993). Full vowels on the other hand, are produced on the periphery of the vowel space, dispersed from each other and associated with greater intelligibility and clearer speech (Bradlow, Torretta, & Pisoni, 1996a).

Figure 10 below is a graphical representation of an F1 and F2 vowel space with hypothetical full and reduced vowels (motivated by Koopmans-van Beinum, 1980; reproduced from Low,
Chapter 5: Vowel Reduction Background

1998). Full, target vowels in Figure 10 are the white circles and are fully realized, found in positions dispersed around the outside of a vowel space. The central point of the vowel space however, reflects the ‘neutral’ schwa position (Shearme & Holmes, 1962), and it is towards this point that unstressed vowels, when reduced, are shifted. Reduced vowels are represented in Figure 10 by the black circles.

![Figure 10 Spectral patterns for full and potentially reduced vowels in British English (reproduced from Low, 1998, p. 57)](image)

5.2. Where Reduced Vowels Occur and Factors Associated with Vowel Reduction

There are many factors associated with vowel reduction, and style and stress have long been regarded as important influences. Spontaneous and conversational speech, for example, are more likely to show vowel reduction than elicited speech, which is clearer and more careful (Aylett & Turk, 2006; Barry & Andreeva, 2001; Gahl, Yao, & Johnson, 2012; Meunier & Espesser, 2011; Smiljanic & Bradlow, 2005). Read prose is also more likely to show vowel reduction than read wordlists, or isolated vowels (Koopmans-van Beinum, 1980; Shearme & Holmes, 1962).

In an early comparative study of isolated vowels, stressed and unstressed vowels from read speech in English, Tiffany (1959) found that when plotting on an F1 by F2 vowels space, they were located on the outside periphery, whereas they were closer to the central point when in connected, read speech. In read sentences, the unstressed vowels were even closer to the central point than the stressed vowels. This was also found in Shearme and Holmes (1962) with vowels from read text being closer to the central point than from isolated monosyllables. Koopmans van Beinum (1980) investigated Dutch vowels in an extensive study on isolated vowels, from isolated words, read speech, in a retold short story and in free conversation. In that study vowels were shortest in duration for words in conversational speech then longer (in order) for retold speech, read speech and words in isolation. They were longest for vowels
produced in isolation. They also found a strong correlation between vowel duration and the Euclidean distance of vowels from the central point for each speaker’s vowel space, with vowels being more centralized and less dispersed as they became shorter.

That vowels become shorter and move towards a central point in the vowel space when unstressed and in more conversational or spontaneous speech has since been replicated in many cross-linguistic studies (e.g. Bell et al., 2003; de Graaf & Koopmans-van Beinum, 1984; DiCanio, Nam, Amith, Garcia, & Whalen, 2015; Gendrot & Adda-Decker, 2005; Jurafsky, Bell, Gregory, & Raymond, 2001).

Other factors involved in vowel reduction include syntactic word type, and position within an utterance. These factors are outlined in the next sections and explored in the Vowel Reduction Study (Chapter 6).

5.2.1. Word Type, function vs. content
In many languages, function words and content words represent distinct syntactic word classes. Content words such as nouns, verbs, adjectives and adverbs tend to hold more semantic information than function words. Function words such as auxiliaries, propositions and conjunctions are usually used to signal relationships between content words (Bolinger, 1965; Couper-Kuhlen, 1986). Function words are prototypically more frequent than content words (Bybee, 2007).

There is a large body of research suggesting that in connected speech, function words in English tend to be reduced, with shorter duration and more centralization than content words (e.g. Aylett & Turk, 2006; Bell et al., 2003, 2009; Bolinger, 1965; Jurafsky et al., 2001; Kohler, 1990; Shi, Gick, Kanwischer, & Wilson, 2005). This effect has also been found in other languages, such as by Meunier and Espesser (2011) who found that the function/content word distinction was a contributor to duration and centralization of vowels in conversational French, with monosyllabic function words being shorter than monosyllabic content words, and the /a/ vowel centralizing for function words.

Much work has sought to explain why function words are more susceptible to reduction than content words. The answer remains unclear (and it is also not the focus of this study). It has been argued that function words are more susceptible to reduction because they tend to be unstressed, they are more frequent than context words, are highly predictable from their
context and that they hold less information and semantic weight (Aylett & Turk, 2004, 2006, Bolinger, 1972, 1985; Jurafsky et al., 2001). Corpus studies have found correlations between duration and frequency of function words, with more frequent function words showing more reduction than less frequent ones (e.g. Bell et al., 2003, 2009; Gahl, 2008; Jurafsky et al., 2001). Vowel reduction has also been found to correlate with differences in predictability or redundancy, with vowels from words with high probability being more likely to be shorter and more centralized (Aylett & Turk, 2004, 2006; Jurafsky et al., 2001). In a regression study of conversational speech, Bell et al. (2009) found function words were shorter than content words, with higher frequency, contextual predictability and repetition all being correlated with greater vowel reduction. In another study, Wright (2004) separated words into ‘hard’ and ‘easy’ categories, with ‘hard’ words having a high number of phonologically similar words (i.e. high neighbourhood density) and lower relative frequency, and ‘easy’ words having the opposite (see Luce & Pisoni, 1998; Luce, Pisoni, & Goldinger, 1990). They showed that vowels from hard words were more dispersed than easy words on a vowel plot, with those from easy words being closer to the central point. This effect has been replicated (Gahl & Strand, 2016; Gahl et al., 2012).

It is not been easy to extricate from the literature how individual factors are involved in reduction of function words as they are all interrelated and experimentally tested in different ways/methodologies. In one attempt to look at the effect of only frequency on reduction, Shi et al. (2005) tested whether function words as a distinct syntactic category are different (more likely to be reduced and assimilated) from content words phonetically/phonologically, or if it is only a frequency effect (because function words are repeatedly more frequently than content words). They isolated the factors of frequency and word type (controlling speech rate and stress) and found that word type had a larger effect on vowel reduction (duration and intensity) than frequency, which had a secondary within-category effect (for duration). Others have even argued that function words are not necessarily reduced, but that content words are more likely to be stressed because they hold more information content and semantic weight and stress is correlated with longer vowels and more clear pronunciation (more vowel dispersal) (Aylett & Turk, 2004; Bolinger, 1972, 1985).

In the Vowel Reduction Study, to follow (Chapter 6), I compared vowel reduction in function words and content words in Australian English and Kriol. Because function words tend to be
unstressed and are likely sites of vowel reduction in English, it was decided to replicate these findings in English and see how Kriol compared.

5.2.2. Utterance Position
Another location where vowel reduction is likely to occur is in words that are in medial position in an utterance. The location of a word in a prosodic domain has been known to play a role in reduction. For example, as mentioned in the Rhythm Metric Study, there is final lengthening in words in final position (Crystal & House, 1990; Klatt, 1975; Ladd & Campbell, 1991). Final lengthening in vowels and syllables has been found in Australian English by Fletcher and McVeigh (1993). Also, words in initial position may be strengthened and pronounced more clearly (Fougeron & Keating, 1997). Bell et al. (2003) showed that vowels from words in utterance initial and utterance final words were longer in duration and less likely to be reduced than those in medial position. Their measure of reduction in that study however, relied on perceptual coding of just saying whether the vowel was full or reduced.

In the Vowel Reduction Study, I sought to replicate this finding in Australian English, and to see if it was also present in Kriol. I also wished to do a more fine-grained analysis using duration measurements as well as a dispersion measurement from F1 and F2.

5.3. Conclusion
This chapter has outlined literature relevant to vowel reduction. For the Vowel Reduction Study (Chapter 6, to follow), I will concentrate on vowel variation as a result of being in unstressed contexts where reduction usually occurs in English. For example, in fast connected speech (our storytelling data), in function words and in medial position. I will measure duration and dispersion (centralisation) and compare findings for Australian English and Barunga Kriol.
Chapter 6: Vowel Reduction Study

Chapter 5 provided the theoretical framework and background for the second study in this project. Chapter 6 presents the second study: a comparison of vowel reduction in Kriol and English. In Section 6.1 hypotheses are described, Section 6.2 outlines methods and Sections 6.3-6.4 present results and a discussion of limitations and directions for future research.

6.1. Research Questions and Hypotheses
This study extends beyond findings of the rhythm metric study by comparing patterns of vowel reduction in Barunga Kriol and Australian English using temporal and spectral analyses. The rhythm metric study had found some differences in rhythm between the two languages based on temporal patterns, and this study seeks to investigate whether these differences could lie in patterns of vowel reduction.

In the phonological account of rhythm, vowel reduction is a factor determining rhythmic differences between languages, and rhythm metrics were developed based on this theory (Chapter 3). Due to the presence of vowel reduction in stress-timed languages, it is expected that rhythm metrics will show that stress-timed languages have higher variability in durations of their vowels compared to syllable-timed languages, in which vowel reduction are expected to be absent. Further, the absence of vowel reduction in syllable-timed languages leads to a higher %V - the percentage of speech made up by vocalic material - as full vowels supposedly take up more temporal space in these languages. Stress-timed languages on the other hand, show lower %V due to reduction in vocalic material caused by vowel reduction.

The rhythm metric study found that at a local level, durational variability was lower in Kriol than English, and that %V was higher in Kriol than English. This could be interpreted as possible evidence that Kriol is more syllable-timed than English; could this be due to less vowel reduction in Kriol than English?

Lower metric results for Kriol than English could be due to transfer effects from possible substrate languages (e.g. Dalabon, Bininj Gun-wok, see Chapter 2). In developing predictions for the vowel reduction study, the effect of substrate languages was again considered. Considerable variation has been found in the height of low central vowels, as well as F1 measures of unstressed /a/ vowels in the connected speech of Bininj Gun-wok and Dalabon (Fletcher & Butcher, 2014), and in Dalabon stressed vowels show much longer duration and more intensity than unstressed vowels (Fletcher & Evans, 2002). These findings suggest that to some extent, vowel reduction may be present in these languages,
Chapter 6: Vowel Reduction Study

though the reduction may be different in nature. For example, in Bininj Gun-work, unstressed vowels are not reduced to schwa (Fletcher & Butcher, 2014), vowel quality is not a significant correlate of stress in either Kundedjnjenghmi (a dialect of Bininj Gun-wok), or Dalabon (Fletcher & Evans, 2002), and in Dalabon vowels in unstressed syllables are more likely to elide than reduce (Fletcher & Butcher, 2014). Due to transfer effects from these possible substrate languages (e.g. Siegel, 2008), vowel reduction was expected to be present but perhaps different in nature in Barunga Kriol.

Although literature on the phonetics and phonology of Kriol itself is limited, some preliminary work suggests that function words in Barunga Kriol are shortening and changing in quality compared to content words, in the speech of younger speakers and may be evidence of Grammaticalization in progress (Jones et al., 2015). This may be evidence of vowel reduction in function words. It is important to note that this work used the same corpus as the current project, so it was likely that similar results would be found in this work. A review article (Butcher, 2008b) on 'Australian Aboriginal English' (encompassing the continuum of Aboriginal ways of using English through to varieties of Kriol in northern Australia, see Chapter 2) states that vowel reduction is present in these varieties, as it is in some traditional languages which are arguably substrate languages. No previous work, however, has examined in any depth for a specific Kriol variety the acoustic extent of vowel reduction, in durational and/or spectral terms.

In Australian English, unstressed vowels are expected to reduce to schwa-like quality (Cox & Palethorpe, 2007) and reduce in duration (Fletcher & McVeigh, 1993). Vowel reduction is a characteristic common to English and languages with heavy stress such as German (Lindblom, 1963b). Few studies have looked at vowel reduction patterns in Australia English with reference to word type or utterance position, though Ingram (1989) in their study stated that vowel reduction was confined exclusively to function words. In the present study it was expected that findings would replicate work for English suggesting that function words were more reduced than content words, and in medial position (see Chapter 5).

In sum when considering the evidence for vowel reduction in Kriol, its possible substrate languages, and in Aboriginal English (Chapter 2), together with the findings of the rhythm metric study (Chapter 4), it was predicted for the vowel reduction study that vowel reduction would be present in both Barunga Kriol and Australian English, though the vowel reduction might differ (in extent and/or in nature) to that found in Australian English. It was expected that vowels in unstressed syllables would be shorter in duration than those in stressed syllables, and that they would centralize or become less dispersed. It was not possible to specify the nature of vowel reduction (i.e. to construct an expectation
Chapter 6: Vowel Reduction Study

of the spectral changes) in Kriol based on substrate languages as there currently is no detailed phonetic analysis of those languages (and not many speakers left).

In order to add to the literature on Kriol and to expand on the description of rhythm beyond the first study, the present study compared the effects of the following variables on durational and spectral measures of vowel reduction in both Kriol and English: 1) word type and 2) utterance position. Research has shown that in English, vowel reduction can be seen in syllables in function words and unstressed medial position (Bell et al., 2003). Does vowel reduction also occur in these contexts in Kriol, and if so what is its nature? By choosing specific contexts which are relatively well understood in studies of English vowel reduction, I hoped to advance the understanding of Kriol vowel reduction, via a methodology that did not rely on (non-native speaker) intuitions of what might be 'stressed' or 'unstressed' syllables in Kriol.

6.1.1. Research Question 1

What are the effects of word type (function vs. content) on vowel reduction (duration and dispersion)?

The hypothesis for this question was based on assumptions that cross-linguistically, function words are generally more frequent, more predictable and hold less information than content words and are therefore more likely to be unstressed, whereas content words are generally stressed (Aylett & Turk, 2004; Bolinger, 1972, 1985; Bybee, 2007). It was expected that vowels from function words would be reduced compared to content words in both Kriol and English. Function words were expected to have shorter durations and that spectrally they would shift in the F1 and F2 measures towards a more central position, becoming less dispersed.

6.1.2. Research Question 2

What are the effects of utterance position (initial or final word in utterance, or medial) on vowel reduction?

The hypothesis for this question was based on findings in English that words in medial position more likely to undergo vowel reduction compared to words in initial or final position (Bell et al., 2003). It was expected for both languages that vowels from words in medial position would be most reduced, with shorter duration and central shifts (less dispersion) in F1 and F2. It was expected that the effect would be similar for both languages. It was expected that vowels in final position would be longest in duration, an effect found in the literature (Crystal & House, 1990; Klatt, 1975; Ladd & Campbell, 1991).
Chapter 6: Vowel Reduction Study

The research questions were answered by measuring duration and F1 and F2 (combining them into a normalised dispersion measure) from selected vowels and contexts, from both Kriol and English. The method is outlined in Section 6.3 below.

6.2. Method

6.2.1. Data Preparation

This study used the same spontaneous storytelling data as the rhythm metric study (same participants, data collection, recordings, data cleaning), with the following preparation.

All vowel segments in the speech sample were extracted using a Matlab script which generated measures for duration (in seconds), F1 and F2 (taken at the vowel midpoint), and preceding and following segment labels. Utterance position was also coded, with utterance boundaries defined as pauses of 150 ms or longer (as per rhythm metric study) and vowels coded as ‘initial’ if taken from the first word in the utterance, ‘final’ if taken from the last, or ‘medial’ if neither initial nor final (as in Bell et al., 2003). Each vowel was coded manually for preceding and following environment (place of articulation), and for word type (whether the vowel was taken from a function or content word). As per the rhythm metric study, vowels from words bordered by pauses on both sides were excluded (i.e. they were not regarded true ‘utterances’).

The data was then filtered so that only vowel tokens from single syllable words were included for analysis. This was done to avoid the need to identify word stress perceptually in Kriol, which could be problematic as a non-native speaker as expectations for English would influence the coding. Due to the fact that the speech data was spontaneous and vocabulary uncontrolled, following extraction it was found that there were most tokens available for the vowels /ɪ/, /ɛ/ and /ʌ/. A decision was made to focus on these vowels for this study, as a starting point for comparing patterns of vowel reduction. The labeling of vowels here was not based on the automatic MAUS phonetic labeling but instead on a combination of acoustic and perceptual coding based on the author’s familiarity with the English and Kriol lexicon.

Of the selection of function words extracted, those with most tokens per word available were chosen for analysis. Function words containing vowels with variation in production, such as ‘the’ which can be pronounced /ði/ or /ðʌ/ in English were excluded from analysis. As there was more variety in unique content words present, and less token numbers for each, a random selection of content words were chosen from each vowel until there was minimum 30 tokens of each vowel category. For some vowels, this represented as many tokens available. Following this step, there were 1361 tokens
Chapter 6: Vowel Reduction Study

available for analysis. See Appendix 4 for the list of specific function and content words analysed in this study.

The final step in data preparation for this study involved manually checking all tokens. Automatic formant tracking for F1 and F2 was checked for all tokens, and manual edits were made to 271 of the 1361 tokens. Tokens were excluded if they were:

- pronounced atypically or incorrectly (14 tokens were excluded for these reasons)
- too glottalised, breathy or formant extraction was obscured by laughter or singing (7)
- unnaturally long (13)
- near hesitations or filled pauses (7)
- disrupted by noise (e.g. background noise) (6)
- completely elided or deleted, or not present enough to be able to track formants (33)

As mentioned earlier, tokens from words bordered by two silences were also excluded (23). Following exclusions (103 in total), 1258 tokens were available for analysis across all vowels and word types.

To answer RQ1, a further exclusion of tokens in final position left 1085 tokens. This final exclusion was a methodological choice due to tokens in final position being longer than in other positions and less reflective of typical patterns of vowel reduction (following other vowel reduction research, e.g. Bell et al., 2003; Gahl et al., 2012). As RQ2 investigated the effect of utterance position on vowel reduction, tokens in final position were kept in the analysis.

6.2.2. Analysis

To calculate dispersion measurements, vowel plot central points were calculated for each language and the distance for each token from this point was measured. These were calculated first for each speaker by extracting the F1 of up to 10 tokens (as many as were available) of the vowels /ɪ/, /ʌ/ and /ʊ/ (representing peripheral vowels on the F1/F2 vowel space) and then taking the mean of all three values. This was repeated for F2 values. To calculate the central point for each language, the mean of all individual speaker central points was taken.

An R script was used to calculate the Euclidean distance of each vowel token from that speaker’s central point (adapted from Pythagoras theorem) (Equation 1). This is illustrated in Figure 11, which is an F1 by F2 vowel plot showing a hypothetical speaker’s central point (marked by ‘x’) and the rationale for calculating a Euclidean distance for an arbitrary token, marked by a black filled circle. These distances were then normalised (by vowel and speaker) into z-scores following Gahl, Yao and Johnson (2012), to take into account vowel spaces of different sizes (Equation 2).
Chapter 6: Vowel Reduction Study

For calculating Euclidean Distance for a token from the central point in the vowel space:

\[ c^2 = a^2 + b^2 \]

\[ c = \sqrt{|F_{1(token)} - F_{1(central point)}|^2 + |F_{2(token)} - F_{2(central point)}|^2} \]

where ‘c’ is the distance of some token from the central point in the vowel space. ‘a’ and ‘b’ are distances for two sides of the right-angled triangle (see Figure 11). F1 and F2 are measurements in Hertz for that token and central point.

**Equation 9 Equation for Calculating Euclidean Distance**

![Diagram for calculating Euclidean distance for an arbitrary token.](image)

The token is marked with a black filled circle. The ‘X’ is that speaker’s central point. ‘a’, ‘b’ and ‘c’ are distances for the sides of a hypothetical right-angled triangle.

**Equation for Calculating Dispersion Measure**, which is a normalised version of the Euclidean Distance in Equation 1.

\[ C_{(normalised)} = \frac{c - \text{mean distance of all of that vowel type from central point}}{\text{standard deviation of distance of all of that vowel type from central point}} \]

where ‘c’ is the distance of some token from the central point in the vowel space.

**Equation 10 Equation for Calculating Dispersion Measure**

6.3. Results

Analysis was performed in R Version 1.0.153 using linear mixed effects models in the package *lme4*. These models are most appropriate for the data used in this study as they allow comparison between groups of unequal sample sizes. By including random effects, they also control for differences across speakers and tokens. Further, they allow analysis of interactions using multiple predictors. Four series of mixed effects models were used in this study, two for each research question. As random effects, all four models had intercepts for speaker, word (from which the token was taken), and preceding and following place of articulation (as in DiCanio et al., 2015). As data was not controlled with respect to place of articulation of adjacent consonant and it is a known factor that can affect how
Chapter 6: Vowel Reduction Study

A vowel is realized (e.g., co-articulation), place of articulation was included as a random effect. Fixed effects and dependent variables differed for each model depending on research question:

For RQ1, two models were used to analyse the effect of word type on vowel reduction (measured with two dependent variables: duration and dispersion). Each model included as fixed effects Word Type (categorical variable of function or content word) and Vowel (/ɪ/, /ɛ/ and /ʌ/) and one of two dependent variables: 1) Duration (continuous variable in seconds, log transformed) or 2) Dispersion (continuous variable in dispersion units).

For RQ2, the same logic was applied but for the effect of utterance position on vowel reduction. Again two models were used, each model including as fixed effects Utterance Position (categorical variable of Initial, Medial or Final positions) and Vowel (/ɪ/, /ɛ/ and /ʌ/) and one of two dependent variables: 1) Duration (continuous variable in seconds, log transformed) or 2) Dispersion (continuous variable in dispersion units).

To test for significance of each fixed effect on vowel reduction, p-values were obtained by conducting an ANOVA to compare a null model including only one fixed effect for Vowel against a full model which included both Vowel as well as the fixed effect of interest (i.e., Word Type or Utterance Position). If inclusion of the fixed effect of interest fit the data significantly better than one without, then a model including the interaction of that effect with Vowel was subsequently tested in the same manner. If that model including the interaction fit the data significantly better than one without, it was included in the analysis. Note that this method of establishing significance approximates only a lower bound on degree of freedom, and it is not possible to obtain the upper bound (Baayen et al., 2008).

The following is an example of the R notation used for the null and two full models in RQ1:

```r
NullModel <- lmer(Duration ~ Vowel + (1|Speaker) + (1|Word) + (1|PrecedingPoA) + (1|FollowingPoA), data=data, REML=FALSE))
FullModel <- lmer(Duration ~ WordType + Vowel + (1|Speaker) + (1|Word) + (1|PrecedingPoA) + (1|FollowingPoA), data=data, REML=FALSE))
FullInteractionModel <- lmer(Duration ~ WordType*Vowel + (1|Speaker) + (1|Word) + (1|PrecedingPoA) + (1|FollowingPoA), data=data, REML=FALSE))
```

To obtain means, 95% confidence intervals, degrees of freedom and p-value estimates, the package `lmerTest` was used. This package relies on the Satterthwaite’s method-of-moment approximation to calculate degrees of freedom and produce a least squares means table. It also produces a table for differences of least squares means to test for significant differences between variables. For all tests, p-values lower than 0.05 were considered significant.

Residual plots of all models were visually inspected for deviations from homoscedasticity or normality and those including the dependent variable of Duration were found to be right skewed.
violating the assumptions of linear mixed effects models. A log transform on the duration data solved this issue.

The following sections present results for each research question, first RQ1 with the effects of word type on vowel reduction (duration then dispersion) and second RQ2 with the effects of utterance position on vowel reduction (duration then dispersion). The languages were analysed separately. Kriol results will be presented first, followed by English.

6.3.1. Word Type and Vowel Reduction
It was expected for both languages that vowels from function words would be reduced compared to vowels from content words, with shorter duration and less dispersion. It was expected however, that vowel reduction would occur in Kriol to a lesser or different extent than English.

The results of lmerTest on a linear mixed effects model with fixed effects of Word Type and Vowel and dependent variable Duration or Dispersion, are shown in Table 3 for both languages. The mean (M) and standard error (SE) here and in all descriptive statistics tables in this chapter are estimates of the class means that would be expected had they been equal size, established by ‘least squares means’ using the lmerTest package. Decimal points are all rounded up to three points and significances are denoted in bold in all tables.
Chapter 6: Vowel Reduction Study

<table>
<thead>
<tr>
<th>Group</th>
<th>Kriol (N=767)</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content (N=44)</td>
<td>Function (N=723)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>-2.314</td>
<td>-2.787</td>
<td>0.117</td>
<td>0.230, 0.715</td>
</tr>
<tr>
<td>SE</td>
<td>0.128</td>
<td>0.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration (Log Seconds)</td>
<td>1.154</td>
<td>0.299</td>
<td>0.249</td>
<td>0.275</td>
</tr>
<tr>
<td>Dispersion (Dispersion Units)</td>
<td>0.372</td>
<td>0.171</td>
<td>-0.238</td>
<td>0.158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>English (N=318)</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content (N=66)</td>
<td>Function (N=252)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>-2.751</td>
<td>-2.836</td>
<td>0.096</td>
<td>-0.092, 0.263</td>
</tr>
<tr>
<td>SE</td>
<td>0.096</td>
<td>0.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duration (Log Seconds)</td>
<td>0.372</td>
<td>0.171</td>
<td>-0.238</td>
<td>0.158</td>
</tr>
<tr>
<td>Dispersion (Dispersion Units)</td>
<td>0.372</td>
<td>0.171</td>
<td>-0.238</td>
<td>0.158</td>
</tr>
</tbody>
</table>

Table 3 Descriptive Statistics for Duration/Dispersion by Word Type (estimates are from statistical models)

Significance codes: ‘***’ is p< 0.001, ‘**’ is p<0.01 and ‘*’ is p<0.05.

6.3.1.1. Word Type and Duration

Overall, the expected hypothesis was supported only for Kriol: vowels from function words were temporally reduced compared to those from content words.

Figure 12 below, which shows the overall duration in log seconds for content and function words, collapsed across all vowels for both Kriol and English. The raw data equivalent of all figures in this chapter using transformed durational measures can be found in Appendix 5. For Kriol, the shorter duration for function words compared to content words indicates temporal reduction/the effect of word type on duration.
For Kriol, there was an effect of Word Type on Duration ($\chi^2(1) = 12.354$, $p < 0.001$). This was established by an ANOVA comparison, which showed that a full model including vowel and word type fit the data better than one including only vowel. In Table 3 above it can be seen that vowels from function words were significantly shorter than those from content words by $-0.47283$ log seconds, which is equivalent to $0.03725$ seconds (after back-transforming the means and calculating the difference) ($t(12.3) = 4.21$, $p < 0.001$).

For English, there was no effect found for Word Type on Duration ($\chi^2(1) = 0.9112$, $p = 0.3398$). This can be seen in Figure 12, where there is no difference between function and content words.

Figure 13 below shows duration of vowels by word type for each vowel individually, showing that the difference between content and function word varies by vowel. In this plot and all following, the vowel /ʌ/ is denoted by ‘a’, /ɛ/ by ‘e’ and /ɪ/ by ‘i’.
Chapter 6: Vowel Reduction Study

Duration (Log Transformed) by Word Type

Figure 13 Vowel Duration (Log Transformed) by Word Type and Vowel (Kriol and English)

Fitting a model with an interaction for the fixed effects of Duration, Vowel and Word Type did not significantly improve the fit for either Kriol ($\chi^2(2) = 5.921, p = 0.0518$) or English ($\chi^2(2) = 0.425, p = 0.8086$). However, the more complex interaction model did provide information as to how the effect of word type varied between vowels. Table 4 presents the results using lmerTest on the interaction models, for both languages. Significant differences are denoted in bold.
### Chapter 6: Vowel Reduction Study

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Kriol (N=767)</th>
<th>95% CI for Mean Difference</th>
<th>Df</th>
<th>t value</th>
<th>English (N=318)</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Group Content</td>
<td>Function</td>
<td>N</td>
<td>M</td>
<td>SE</td>
<td>N</td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td>All Vowels</td>
<td>-2.321</td>
<td>0.129</td>
<td>44</td>
<td>-2.803</td>
<td>0.118</td>
<td>723</td>
<td>0.257</td>
<td>0.708</td>
</tr>
<tr>
<td>/ʌ/ (Kriol N=167, English N=54)</td>
<td>-2.023</td>
<td>0.165</td>
<td>13</td>
<td>-2.781</td>
<td>0.151</td>
<td>154</td>
<td>0.381</td>
<td>1.134</td>
</tr>
<tr>
<td>/ɛ/ (Kriol N=379, English N=47)</td>
<td>-2.199</td>
<td>0.196</td>
<td>9</td>
<td>-2.730</td>
<td>0.140</td>
<td>370</td>
<td>0.089</td>
<td>0.993</td>
</tr>
<tr>
<td>/ɪ/ (Kriol N=221, English N=217)</td>
<td>-2.740</td>
<td>0.169</td>
<td>22</td>
<td>-2.898</td>
<td>0.147</td>
<td>199</td>
<td>-0.222</td>
<td>0.539</td>
</tr>
</tbody>
</table>

Table 4 Descriptive Statistics for Duration by Word Type and Vowel with Interaction (estimates are from statistical models)

Significance codes: ‘***’ is p< 0.001, ‘**’ is p<0.01 and ‘*’ is p<0.05.
Chapter 6: Vowel Reduction Study
When considering vowels separately, the interaction model showed that the difference between function and content words was significant for Kriol vowels /ʌ/, with vowels from function words being significantly shorter than from content words by 0.757585 log seconds (t(16.5)=4.256, p<0.001) and /ɛ/ by 0.531060 log seconds (t(23.6)=2.474, p<0.05) but not /ɪ/. For English, there were no significant differences between function and content words for any vowels.

6.3.1.2. Word Type and Dispersion
Overall, the expected hypothesis was found for both Kriol and English: that vowels from function words were less dispersed than those from content words (closer to the central point). This can be seen in the movements of F1 and F2 in Figure 14 below. Figure 14 is an F1 by F2 vowel plot showing the mean values for each vowel (collapsed across all speakers), for both Kriol and English. Vowels from content words are plotted with red dashed lines and function words in green full lines. The ‘x’ represents the central point for each language, calculated as the mean F1 and F2 of a selection of /ʌ/, /ɪ/ and /ʊ/ tokens in that language (see Section 6.2.2). Compared to content words, function words have shifted towards the central point for all vowels in both languages (except /ɛ/ in English).

![Function vs. Content, with Central Point](image)

Figure 1 F1 by F2 Vowel Plot: Vowels /ʌ/, /ɛ/ and /ɪ/ by Word Type (Kriol and English).

X marks the central point, which is the average of all the speaker central points.
Chapter 6: Vowel Reduction Study

In Figure 15 below, the F1 and F2 measurements have been normalised and converted into a dispersion measure (as explained in method section above). This boxplot shows the overall dispersion from the central point for vowels from content vs. function words, collapsed across all speakers and vowels. It indicates that for both Kriol and English, function words are less dispersed than content words. This difference in dispersion represents less distance from a central point in the vowel space for function words compared to content words, indicating spectral reduction in function words.

**Figure 15 Dispersion by Word Type (Kriol and English)**

For Kriol, there was a significant effect of word type on dispersion ($\chi^2(1) = 11.388$, $p = 0.0007391$). Table 4 shows that collapsed across all vowels, those from function words were significantly less dispersed than content words by 0.904 dispersion units ($t(28.7) = 3.692$, $p<0.001$).

For English, this effect was also significant ($\chi^2(1) = 7.4143$, $p = 0.006471$), with vowels from function words being less dispersed than content words by 0.610134 dispersion units ($t(41.4) = 3.097$, $p<0.01$).

Figure 16 below shows dispersion of vowels by word type for each vowel individually, showing that the difference between content and function word varies by vowel.
Fitting an interaction model with Dispersion and Vowel against one without an interaction (using ANOVA, likelihood ratio) did not significantly improve the fit for Kriol ($\chi^2(2) = 2.1898$, $p = 0.3346$) or English ($\chi^2(2) = 2.9067$, $p = 0.2338$). However, the more complex interaction model did provide information as to how the effect of word type on dispersion varied between vowels. Table 5 presents the results using *lmerTest* on the interaction models, for both languages. Significant differences are in bold.
### Chapter 6: Vowel Reduction Study

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Content (N=66)</th>
<th>Function (N=252)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td>All vowels</td>
<td>1.206</td>
<td>0.294</td>
</tr>
<tr>
<td>/ʌ/ (Kriol: N=167, English: N=54)</td>
<td>0.608</td>
<td>0.373</td>
</tr>
<tr>
<td>/ɛ/ (Kriol: N=379, English: N=47)</td>
<td>1.767</td>
<td>0.443</td>
</tr>
<tr>
<td>/i/ (Kriol: N=221, English: N=217)</td>
<td>1.243</td>
<td>0.383</td>
</tr>
</tbody>
</table>

*Table 5 Descriptive Statistics for Dispersion by Word Type and Vowel with Interaction (estimates are from statistical models)*

Significance codes: *** is p< 0.001, ** is p<0.01 and * is p<0.05.
**Chapter 6: Vowel Reduction Study**

The interaction model showed that for Kriol the effect was significant for vowels /ɛ/, with vowels from function words being significantly less dispersed than those from content words by 1.414 dispersion units ($t(41.4)=3.202, p<0.01$) and /ɪ/ by 0.885 dispersion units ($t(24.8)=2.255, p<0.05$). Vowels from /ʌ/ were not significantly different.

For English, the effect was significant for vowels /ʌ/, with vowels from function words being significantly less dispersed than those from content words by 0.917 dispersion units ($t(39.1)=2.367, p<0.05$) and /ɪ/ by 0.744 dispersion units ($t(41.9)=2.605, p<0.05$). Vowels from /ɛ/ were not significantly different.

In summary, for Research Question 1, significant effects of Word Type on Duration in Kriol and Dispersion in both Kriol and English were found, with vowels from function words being shorter in duration and less dispersed than content words. The effects varied by vowel for some main effects.

6.3.2. **Utterance Position and Vowel Reduction**

It was expected for both languages that vowels from words in medial position would be reduced compared to vowels from initial and final positions, with shorter duration and less dispersion.

The results of *lmerTest* on linear mixed effects models with fixed effects of Utterance Position and Vowel and dependent variable Duration or Dispersion are shown in Table 6 for Kriol and English. Significances are denoted in bold.
## Chapter 6: Vowel Reduction Study

### Kriol (N=869)

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial (N=185)</th>
<th>Medial (N=582)</th>
<th>Final (N=102)</th>
<th>Comparison</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>M SE</td>
<td>M SE</td>
<td>M SE</td>
<td>Initial - Medial</td>
<td>-0.264, -0.123</td>
<td>787.1</td>
<td>-5.371 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>(Log Seconds)</td>
<td>-2.661 0.127</td>
<td>-2.468 0.125</td>
<td>-2.135 0.128</td>
<td>Final - Initial</td>
<td>0.423, 0.629</td>
<td>850.0</td>
<td>10.005 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Final - Medial</td>
<td>0.248, 0.417</td>
<td>837.9</td>
<td>7.759 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>Dispersion</td>
<td>0.657 0.241</td>
<td>0.640 0.237</td>
<td>0.802 0.244</td>
<td>Initial - Medial</td>
<td>-0.142, 0.178</td>
<td>806.2</td>
<td>0.2208 (p=0.825)</td>
</tr>
<tr>
<td>(Dispersion Units)</td>
<td></td>
<td></td>
<td></td>
<td>Final - Initial</td>
<td>-0.088, 0.376</td>
<td>787.9</td>
<td>1.218 (p=0.223)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Final - Medial</td>
<td>-0.026, 0.350</td>
<td>657.8</td>
<td>1.689 (p=0.092)</td>
</tr>
</tbody>
</table>

### English (N=389)

<table>
<thead>
<tr>
<th>Group</th>
<th>Initial (N=44)</th>
<th>Medial (N=274)</th>
<th>Final (N=71)</th>
<th>Comparison</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>M SE</td>
<td>M SE</td>
<td>M SE</td>
<td>Initial - Medial</td>
<td>-0.094, 0.187</td>
<td>232.3</td>
<td>0.658 (p=0.511)</td>
</tr>
<tr>
<td>(Log Seconds)</td>
<td>-2.774 0.111</td>
<td>-2.821 0.094</td>
<td>-2.240 0.102</td>
<td>Final - Initial</td>
<td>0.356, 0.712</td>
<td>336.8</td>
<td>5.900 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Final - Medial</td>
<td>0.455, 0.706</td>
<td>384.6</td>
<td>9.099 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>Dispersion</td>
<td>-0.269 0.297</td>
<td>-0.113 0.272</td>
<td>0.032 0.280</td>
<td>Initial - Medial</td>
<td>-0.444, 0.134</td>
<td>377.9</td>
<td>-1.056 (p=0.292)</td>
</tr>
<tr>
<td>(Dispersion Units)</td>
<td></td>
<td></td>
<td></td>
<td>Final - Initial</td>
<td>-0.063, 0.664</td>
<td>388.1</td>
<td>1.626 (p=0.105)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Final - Medial</td>
<td>-0.109, 0.400</td>
<td>382.9</td>
<td>1.122 (p=0.263)</td>
</tr>
</tbody>
</table>

Table 6 Descriptive Statistics for Duration/Dispersion by Utterance Position (estimates are from statistical models).

Significance codes: ‘***’ is p< 0.001, ‘**’ is p<0.01 and ‘*’ is p<0.05
6.3.2.1. Utterance Position and Duration
Overall, the expected hypothesis was somewhat found for both Kriol and English: that vowels from words in medial position were most reduced (shorter duration) compared to those from final words. Final position was also longest in both languages.

Figure 17 below is a box plot showing duration by utterance position, collapsed across all vowels, for Kriol and English. It is clear that vowels in both initial and medial positions are shorter than those in final position for both languages. It seems that initial and medial positions are not different to each other.

For Kriol the effect of utterance position on duration was significant ($\chi^2(2) = 93.534, p<0.001$). Vowels from words in medial position were significantly shorter than those in final position by 0.332 log seconds (equivalent to 0.033 seconds) ($t(837.9) = 7.759, p<0.001$). Vowels from initial position were also significantly shorter than those in final position, by 0.526 log seconds (equivalent to 0.048 seconds) ($t(850.0) = 10.005, p<0.001$). Vowels from words in medial
position were found to be significantly longer than those in initial position, by 0.193 log seconds (equivalent to 0.015 seconds) \( (t(787.1)=-5.371, p<0.001) \).

The effect of utterance position on duration was also significant for English \( (\chi^2(2) = 73.223, p<0.001) \). A similar pattern to Kriol was found in English, with vowels from words in medial position being significantly shorter than those in final position by 0.580823 log seconds (equivalent to 0.047 seconds) \( (t(384.6)=9.099, p<0.001) \). Vowels from initial position were also significantly shorter than those in final position, by 0.534 log seconds (equivalent to 0.044 seconds) \( (t(336.8)=5.900, p<0.001) \). However, initial and medial positions were not significantly different to each other in English.

Figure 18 is a box plot for Utterance Position and Duration separated by Vowel. It shows how the effect of utterance position on duration varied by vowel. It can be seen that the pattern seems to hold across vowels. It seems that for all vowels, medial position was shorter than final position, and initial and medial weren’t particularly different.
Chapter 6: Vowel Reduction Study

For Kriol, fitting an interaction model with fixed effects of Utterance Position and Vowel against one without an interaction (using ANOVA, likelihood ratio) significantly improved the fit ($\chi^2(4) = 19.934, p = 0.0005146$). For English, fitting an interaction model against one without the interaction did not significantly improve the fit ($\chi^2(4) = 7.1203, p = 0.1297$).

Separating tokens into vowel groups for each utterance position yielded groups with very low numbers, leaving the interaction model with low power. The results of *lmerTest* on the linear mixed effects models with fixed effects of Utterance Position and Vowel are shown in Appendix 6, in Table 9 for Kriol and English. However, due to low power issues they are not considered in the main analysis for this study.

**6.3.2.2. Utterance Position and Dispersion**

Overall, the expected hypothesis was not found for both Kriol and English. Vowels from words in medial position were not less dispersed compared to those from initial and final words.

Figure 19 is a boxplot showing the effect of utterance position on vowel dispersion, collapsed across all speakers and vowels. From the boxplot it might seem that for Kriol, vowels in medial position are less dispersed than initial and final positions. No differences are apparent in English.
Figure 19 Dispersion by Utterance Position (Kriol and English)

There was no effect of Utterance Position on Dispersion for Kriol ($\chi^2(2) = 2.8207, p=0.2441$) or English ($\chi^2(2) = 2.6293, p=0.2686$). In other words, a model with fixed effects for Vowel and Utterance Position did not fit the data better than one with just Vowel. Using the more complex model, no significances were found for differences in dispersion by utterance position in Kriol or English (Table 6).

Figure 20 is an F1 by F2 vowel plot showing vowel means in each position, compared to the central point, for each language. Vowels from initial position are shown in red, from medial position in green and final position in blue. From these vowel plots it might seem as though /ʌ/ in final position and /ɪ/ in medial position were closest to the central point in Kriol and /ɛ/ vowels in initial position for English.
Figure 20 F1 by F2 Vowel Plot: vowels /i/, /ɛ/ and /ʌ/ by Utterance Position (Kriol and English)

Initial is in blue, medial in red, final in yellow.

Figure 21 shows dispersion for each vowel by utterance position. From this boxplot there are no immediate patterns for utterance position and dispersion. Perhaps, /ʌ/ and /ɛ/ in both languages go from least dispersed in initial position, then slightly more dispersed in medial, and most dispersed in final.
Chapter 6: Vowel Reduction Study

Individual Vowels: Dispersion by Vowel and Utterance Position

As in the duration analysis, separating tokens into separate vowel groups for each utterance position yielded groups with very low numbers, leaving the interaction model results with low power. The \textit{lmerTest} results on the linear mixed effects models with fixed effects of Utterance Position and Vowel and dependent variable Duration or Dispersion are shown in Appendix 7, in Table 10 for Kriol and English. However, they are not considered in the main analysis for this study.

In summary for Research Questions 2, significant effects of Utterance Position on Duration but not Dispersion were found, in both Kriol and English. Vowels from words in medial position were shorter in duration than those from final position. They were longer than those in initial position in Kriol, but no different in English. The effects varied by vowel.

6.4. Discussion and Conclusion

In this study, patterns of vowel reduction in Kriol and English were explored. The following two research questions for this study were answered through analyses of vowels from Kriol and English storytelling data (specifically /ʌ/, /ɛ/ and /i/ vowels):
Chapter 6: Vowel Reduction Study

RQ1: What are the effects of word type (function vs. content) on vowel reduction, in terms of duration and dispersion?

RQ2: What are the effects of utterance position (initial, medial or final) on vowel reduction, in terms of duration and dispersion?

Research Question 1 was based on a body of research suggesting that function words in English tend to be reduced. Therefore, it was expected that word type would have an effect on vowel reduction, with function words being temporally reduced and less spectrally dispersed than content words. For the duration measures, an effect of word type was found for Kriol but not English. In Kriol, vowels from function words were shorter in duration than content words. Although I did not have specific hypotheses for how effects varied by vowels, it was notable that the effect held for vowels /ʌ/ and /ɛ/ but not /ɪ/ in Kriol. For dispersion measures, an effect of word type was found for both Kriol and English, with vowels from function words being less dispersed and closer to the central point on a vowel plot than vowels from content words. In terms of individual vowels, the effect again varied. For Kriol, it was only present in vowels /ɛ/ and /ɪ/ but not /ʌ/. In English, it was present in /ʌ/ and /ɪ/ but not /ɛ/. It is interesting that in Kriol the vowel /ɪ/ was only reduced spectrally and not temporally, /ʌ/ was reduced temporally but not spectrally, and both effects were present for /ɛ/. For English, while /ʌ/ and /ɪ/ were found to be spectrally reduced, the effect of temporal reduction was not found for any vowels.

Research Question 2 was based on literature suggesting that words in initial and final positions are longer and less likely to be reduced than those in medial position in English (Bell et al., 2003). Thus I expected that utterance position would have an effect on vowel reduction, with vowels from words in medial position would show temporal and spectral reduction (shorter duration and less dispersion).

For both languages, utterance position had some effect on duration, but neither showed an effect of spectral reduction. For duration, it was found for both languages that vowels from words in medial position were shorter than those from final, but not initial position. In Kriol, vowels from words in initial position were shorter than medial position, which was not an expected effect. In English, there was no difference found between initial and medial positions. This was also not expected. The results for RQ2 were not analysed for how they varied.
between vowels, because the group sizes were very small. Those results are included in Appendix 6 and Appendix 7 but not considered in the overall analysis. Instead, boxplots and vowel plots were provided to visualise the varying effect of utterance position on the vowels in an exploratory fashion. From the boxplot in Figure 21 for the effect of utterance duration on duration across the vowels, it seems that the pattern found when all vowels were collapsed together is also held for individual vowels. For all vowels in both languages, it can be seen that those from medial position were shorter than final position and mostly no different to initial position. For the effect of utterance position on dispersion, there was no effect found for either language, which was not expected.

Table 7 is a summary of results found in the vowel reduction study, collapsed across vowels. In this table, areas marked with the colour green and letter ‘Y’ denote that significance was found for the effect of that independent variable on dependent variable (i.e. = comparing null vs. full model was significant). Areas marked with the colour red and letter ‘N’ denote no significance found for the effect of that independent variable on dependent variable. The direction of the effect is noted after the ‘Y’ or ‘N’ with the following letters. For RQ1 these are ‘F’ for ‘function word’ and ‘C’ for ‘content word’ and for RQ2 these are ‘I’ for ‘initial’, ‘M’ for ‘medial’ and ‘F’ for ‘final’.

<table>
<thead>
<tr>
<th>RQ1: Word Type</th>
<th>Kriol</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>Y, F&lt;C</td>
<td>N</td>
</tr>
<tr>
<td>Dispersion</td>
<td>Y, F&lt;C</td>
<td>Y, F&lt;C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RQ2: Utterance Position</th>
<th>Kriol</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>N</td>
<td>N</td>
</tr>
</tbody>
</table>

Table 7 Summary of all Results (collapsed across vowels).

Green and ‘Y’ means that significances were found for the effect of that IV on DV (comparing null vs. full model), and then the significant direction of results are noted. Red and ‘N’ means no significances found for effect of IV on DV.
Comparing my findings to previous research, the English patterns of vowel reduction found in this study provide some support of previous findings on vowel reduction in English (as discussed in Chapter 5). However, for both research questions there were unexpected results. In the word type study, I had expected that there would be an effect on both duration and dispersion measures, but I found that vowels from function words in English reduced spectrally but not temporally (unlike what would have been expected from literature on vowel reduction in English, chapter 5). In the utterance position study, I had expected that medial position would be shorter and less dispersed than both initial and final. I found that words in both medial and initial position were shorter than final position, and that there was no effect of utterance position on dispersion. For the Kriol results, there is little research on vowel reduction in Kriol to compare to these findings. One study (Jones et al., 2015) found function words in Barunga Kriol were shorter and had more centralised vowels in initial and medial position compared to final. The data used for that study was from the same corpus as the data used in this study, so its not surprising that the results supported those findings. However, the present study extended beyond them by comparing function words to content words, and finding that function words were shorter and more centralised than content words. My study also looked at utterance position on vowel reduction in Kriol, finding that vowels in medial position were shorter than those in final position, and initial position shorter than medial position.

In light of the fact that the results from the English data were not as expected based on vowel reduction literature in English (discussion to follow), the Kriol data in this project is compared to expected English results rather than the ones found. In which case, it can be reasoned that Kriol showed similar patterns to English. For the word type study, I found that function words were reduced both temporally and spectrally compared to content words, as would have been expected in English. For the utterance position study, I found that vowels from words in medial position were reduced temporally compared to those in final position, as expected in English. However, that vowels from words in initial position were shorter than those in medial position differs from results found for English. Further, that utterance position had no effect on dispersion of vowels in Kriol also contrasts to vowel reduction patterns found for English (e.g. Bell et al., 2003). These results for Kriol show that vowel reduction is acoustically similar in both Kriol and English, with changes in duration and dispersion. The Kriol results also show that Kriol behaves similarly to what is found in the English literature in terms of patterns of
vowel reduction (e.g. for function words at least) and in environments where vowel reduction occurs. As a first pass analysis of Kriol, this study has provided information about vowel reduction in this language, which can also point to directions for future research (e.g. further analysis of vowel reduction and particular vowels).

In order to understand why this study only provided limited support previous findings in English, and to put the results into a broader perspective, the shortcomings and limitations of the methodology and analysis will now be discussed. These should be noted for consideration in future research.

One issue apparent in this study is low sample sizes, particularly the fact that the English data had much lower token numbers than the Kriol data. Although not ideal, this was an inevitable result of the data collection method. Being storytelling speech of spontaneous nature, content was not controlled and speakers were given free reign to speak about the stimulus for as long or as little as they wanted. Although at first it seemed like the data would be balanced across languages, once tokens were extracted, it was found that English had much lower token numbers. It is possible that the results for English did not support previous findings in English because of low power issues caused by insufficient sample sizes. Because of the low sample sizes, I also chose not to include Language as an effect in the linear mixed effects models, as this would likely overload them. Instead, I ran the analyses separately for each language. In future, a larger sample size could be collected, which would allow other statistical approaches to be used, such as one comparing the two languages directly. Note however, that if the researcher wished to compare languages within one lmer model in future, they should not include the random effect of Word (as done in analysis here). As the orthography of words in both languages is similar, the lmer algorithm will treat certain words as from the same population, which may obscure differences across languages.

Another methodological choice which may have had unanticipated effects in this project was the choice in the word type study to exclude vowels from words in final position. This followed previous studies (e.g. Bell et al., 2003; Gahl et al., 2012) where tokens in final position were excluded due to being naturally longer than in other positions and less reflective of typical patterns of vowel reduction. The decision to remove these tokens may have had the unanticipated effect of concealing a possible main effect of word type on duration in English. Although I did not find an effect of word type on duration in the analysis, it is possible that this
occurred because there were different proportions of the types of final tokens excluded for the Kriol and English data. Firstly, of the final tokens excluded, these represented a larger percentage of the total data for English than Kriol (18.25% of data was finals in English compared to 11.74% in Kriol). Also, of those final tokens excluded, for English a larger proportion of the tokens were content words (53.52% in English compared to 26.47% in Kriol). As most tokens of longer duration in the sample seem to be in final position (see Figure 17), and most of final tokens were content words in the English data, then it makes sense that the exclusion may have affected the results for each language differently. One possible effect of this difference is exemplified in Figure 22 below. Figure 22 shows boxplots for word type and duration for Kriol and English, with a comparison of data which excluded the final segment (top) and for data which did not exclude the final segment (bottom). While in Kriol the effect of word type on duration seems present in the data with the exclusion and without, in English the effect seems to be evident (at least visually) in the data without the exclusion. In this boxplot of English data without the exclusion, there is a discernable difference between duration of function and content words, with function words being shorter duration than content words. In future work, the decision to remove final tokens should be under more scrutiny, data could be semi-controlled for more balanced groups.

A reason for the unexpected finding of no significant results in dispersion measures for utterance position may lie with how the utterances themselves were demarcated. Utterances were defined as speech between boundaries of 150ms and it is possible that utterances define in way this eradicated true intonational or prosodic patterns in real life utterances. The decision to define utterances in this way was to maintain objectivity, but perhaps future studies could examine this decision and see whether there is a more ecologically valid method of defining utterances. Further, in this project utterance position was coded using an R script, but future studies could redo this work using different software with hierarchical databases and more fine-grained processing such as through EMU-R. Doing so may yield different results to those in this project.

In terms of the actual dispersion measure used in this project, perhaps in future work it would be better to look at F1 and F2 measures separately rather than combined into one measure. It is possible that by combining and normalising, some variability was lost for utterance position. Intelligibility may also be a factor to explore, as this may have also had an effect on dispersion results of utterance position. In the data used in this study, it was my impression that the
Chapter 6: Vowel Reduction Study

English speech sounded more formal and less casual than the Kriol speech, though this was not explored. Bradlow et al. (1996b) found a correlation between dispersion and intelligibility scores, and determined that measures of F1 were responsible for this correlation rather than F2. This lends further support to the idea that a future study might benefit from measuring dispersal separately rather than combined.
The hypothesis for the effect of utterance position on vowel reduction was based on Bell et al. (2003) who found that vowels from words in medial position were shorter than those from initial and final positions. However, in their discussion they claimed that their results might not give a valid indication of the effect of utterance position on duration as other factors might be systematically linked to certain utterance positions. For example, they cited Shriberg (1994) who found that words in initial position are more likely to occur in the context of disfluencies, and as disfluencies cause words to be longer, words in initial are also likely to be longer in duration. They also thought that utterance initial words might be longer and more full because of low predictability, which might mask the effect of utterance position. Bell et al., (2003) found that when they controlled for (semantic) predictability, the effect was no longer significant (words in initial position were not significantly longer than those in non-initial position). In the present study, disfluencies were excluded, so they did not have an impact on vowels in initial position, which I found to be shorter than medial in Kriol and not different to medial in English. For English, the results of this study therefore support the findings of Bell et al., (2003).
A final limitation of this study is that the effect of speech rate was not considered, as in Study 1. Speech rate is likely to be a correlated with vowel reduction as there is a body of research which has associated faster speech with vowel reduction (e.g. Fourakis, 1991; Gay, 1978; Lindblom, 1963a; van Son & Pols, 1992). Bell et al. (2003) for example, found that all forms of reduction were affected by speech rate, with words being more reduced at higher speech rates. In a way, this makes sense as vowel duration is inextricably linked to vowel dispersion in that vowels of shorter duration are more likely to centralise (Lindblom, 1963b). The possibility of a correlation between the vowel dispersion results and the duration of vowel tokens was not explored in this study. Figure 23 below is a scatterplot of duration and dispersion measures for Kriol and English, with a regression line for a linear model with duration as a fixed effect and dispersion as a dependent variable. This figure shows a weak correlation for both Kriol and English, suggesting that ‘duration’ may have been a good addition to the random effects in the dispersion models as it may have helped to explain variation in the data. Future study of vowel reduction would benefit from exploring the relationship between speech rate and vowel reduction.
In the next chapter the implications of the vowel reduction study in light of the rhythm metrics study are discussed, as well as limitations of the entire project and steps for future research on the rhythm of Kriol.
Chapter 7: Discussion and Conclusion

Chapter 7 concludes this thesis. The findings and implications of the two studies are discussed, limitations acknowledged and directions for future research suggested.

The first of its kind, this project compared rhythm and patterns of vowel reduction in Barunga Kriol and Australian English. Two studies were conducted using storytelling speech. In the first study, the Rhythm Metric Study (Chapter 4), speech data was segmented into utterances and rhythm was measured using durational-based rhythm metrics. The aim of this study was two-fold: to compare the variability in duration of vocalic and consonantal intervals in Kriol and English (RQ1), and to establish which language had greater ‘vocalic-ness’ (amount of speech taken up by vowels) (RQ2). The research questions were based on a body of literature suggesting that language rhythm can be determined using durational-based rhythm metrics, and languages can be placed on a rhythmic continuum between syllable- and stress-timing depending on the behaviour of their vocalic and consonantal intervals. Stress-timed languages typically show greater variability in the duration of vocalic and consonantal intervals, and less vocalic-ness than syllable timed languages (see Chapter 3 for background on rhythmic classification).

To answer RQ1, three metrics were used: VarcoV and VarcoC which measure variability at a global level by averaging it out across a whole speech sample, and nPVI-V which measures it at a local level by averaging out the variability between successive intervals (Chapter 3). Based on a traditional classification of English as stress-timed, some evidence that possible substrate languages to Barunga Kriol may have been more syllable-timed, or at least express vowel reduction in a different way to English, and that some varieties of Aboriginal English have been noted to have different rhythm to English, it was expected that Kriol would be more syllable-timed than English. Higher variability of vocalic and consonantal intervals in English than Kriol were anticipated. For VarcoV and VarcoC no significant differences were found between Kriol and English. This suggests that at a global level, variability of vocalic and consonantal intervals (when averaged across whole speech data) is not different for the languages. For nPVI-V however, English was found to measure significantly higher than Kriol, suggesting that at a local level (between pairs of successive intervals) vocalic intervals in English vary more than...
Kriol. The nPVI-V based results therefore support the hypothesis of more syllable-timedness in Barunga Kriol than English.

To answer the second RQ of the Rhythm Metric Study, the vocalic-ness of the two languages was compared with a measure of %V (the percentage of speech taken up by vocalic material). In the stress- and syllable-timing continuum, stress-timed languages are thought to have more vowel reduction, meaning that there is a lower proportion of speech taken up by vocalic material. On the other hand, syllable-timed languages are thought to have fuller vowels and less vowel reduction. As Kriol was expected to be more syllable-timed than English, it was expected to have higher %V. This hypothesis was supported; Barunga Kriol was found to have significantly higher %V than English, suggesting that for this variety of Kriol, at least, a greater proportional of speech material is vocalic.

The findings in the Rhythm Metric Study suggested that Barunga Kriol may be more syllable-timed than English, due to the fact that English has a higher variability between successive vocalic intervals (nPVI-V) (suggesting possibly greater vowel reduction in English) and that Kriol has a higher proportion of vocalic material (suggesting possibly less vowel reduction in Kriol).

As mentioned in Chapter 3, metric results can be limited in explanatory power as they can suggest differences in rhythm, but are not able to provide any explanation as to from where variability arises (e.g. Arvaniti, 2009, 2012c; Turk & Shattuck-Hufnagel, 2013). For example, although vowel reduction was a main factor in the development of rhythm metrics (they sought to measure rhythm as a function of vowel reduction in terms of duration), the metrics do not necessarily indicate that vowel reduction is present. It is possible that variability in durations may be caused by factors other than vowel reduction, such as syllable lengthening or consonant reduction (see Chapter 3 for discussion). It has been recommended that descriptions of rhythm should not rely on metric results alone but instead should relate metric results to language features by testing specific hypotheses. The second study in this project, the Vowel Reduction Study (Chapter 6), built on the first study by exploring patterns of vowel reduction in Kriol and English.

The Vowel Reduction Study sought to compare durational and spectral characteristics of vowel reduction in Kriol and English, and to establish whether reduction occurred in vowels in
similar positions and in similar ways as in English. It had two aims. The first was to compare reduction as an effect of (syntactic) word type, whether it occurred differently in function words compared to English words (RQ1). The second was to compare reduction as an effect of utterance position, whether it occurred differently in vowels from words occurring in initial, medial or final positions in an utterance (RQ2). These research questions were based on studies in English suggesting that vowels are more likely to be reduced in function words and in words in medial position in an utterance compared to initial and final (see Chapter 5). In this second study, vowel reduction was defined as vowels being shorter in duration and spectrally closer towards a central point (less dispersed) on an F1 by F2 vowel plot.

To explore the effect of word type on vowel reduction (RQ1), the temporal and spectral characteristics of vowels extracted from a selection of single syllable function and content words were compared. It was expected that Kriol would show reduction in similar patterns to English (but perhaps to a lesser extent, considering results from the first study and expectation that Kriol would be more syllable timed). It was found that vowels in function words were shorter than in content words only in Kriol and not English. In both languages, however, vowels were spectrally less dispersed in function words than content words.

To explore the effect of utterance position (RQ2), the temporal and spectral characteristics were again compared, this time in vowels from single-syllable words in initial, medial and final positions in utterances. It was expected that vowels in medial position would show greater vowel reduction than vowels in initial and final position. It was found that in both Kriol and English, vowels in medial position were shorter than those in final position, but not shorter than vowels in initial position. In Kriol vowels in initial position were shorter than medial position, but in English they were not different. Neither language showed any effect of utterance position on dispersion (i.e. spectral measures of vowel reduction).

The findings of the Vowel Reduction Study suggest that Kriol shows similar patterns of vowel reduction to English, when vowel reduction is examined in detail in temporal and spectral terms, for function versus content words, and by utterance position. I found some effect of word type on vowel reduction, with spectral effects in both languages and durational effects in Kriol. I also found a durational effect of utterance position on vowel reduction, with vowels from words in medial position being reduced compared to final position.
In light of the Vowel Study results, some questions raised by the Rhythm Metric Study are now addressed. The findings of the Rhythm Metric Study were that the rhythm of Kriol was different to English in two out of the four rhythm metrics (nPVI-V and %V), suggesting that Kriol may be more syllable-timed than English. But I was not able to conclusively say that (because there was no difference for VarcoV and VarcoC). I was also not sure whether the variability found on local level by nPVI-V and on a global level by %V arose from vowel reduction or from something else? The second study sought to help answer these questions by seeing if vowel reduction was present in Kriol and if so, whether it was realized in an acoustically similar way (temporally and spectrally) and in the same patterns.

The Vowel Reduction Study found that reduction was indeed present in Kriol and realized more similarly than expected for the two languages. Particularly relevant to the rhythm metric results, it was found that for both Kriol and English, temporal vowel reduction was realized in typically unstressed locations (function words and medial position). That temporal vowel reduction was found in both languages supports the findings of the Rhythm Metric Study that there was no difference in VarcoV and VarcoC. Why then, was a difference found for nPVI-V and %V? As mentioned in the discussion of the Rhythm Metric Study (Chapter 4) the answer may lie in the methodology. In the Rhythm Metric Study, the final segment (final V token) of each utterance was excluded. When the metric results were compared to measurements with this segment being included (Appendix 2), there was no longer a significant difference for %V (all other metrics had same results). This meant that when the final segment was included, three out of the four metrics resulted in no difference between the languages. It’s possible that this final segment is longer in English than Kriol, and removed the %V difference that was originally there. Although I know from looking at the effect of utterance position on duration in The Vowel Reduction Study that vowels in final position were longer than in medial position, I don’t know how the durations compare between the languages. If it were the case that final position was longer in English, it is also likely this might drive the nPVI-V to be greater in English than Kriol when that segment was included. Why then, was nPVI-V significantly different between the two languages, even with exclusion of the final segment? A limitation of The Vowel Reduction Study discussed in Chapter 6 was low token numbers for English compared to Kriol. This issue was not explored in The Rhythm Metric Study, but perhaps may be a factor in the finding of different results in nPVI-V where there may not be any. Considering no durational difference was found for vowel reduction in the Vowel Reduction Study, it
remains an open question as to why the rhythm metrics picked up on vowel durational differences in the Rhythm Metric Study, and whether these results are valid. On the whole, the difficulty in interpreting findings in this study has been noted as a limitation of rhythm metrics in general (see Chapter 3). Future studies would do well to repeat this study with more data to see if similar results are found.

For many researchers rhythm is defined as a perceptual phenomenon (Couper-Kuhlen, 1993) and as this thesis focused solely on acoustic correlates of rhythm, one possible limitation could be that it has neglected investigating rhythm from a perceptual perspective. The decision not to look at rhythm perception in this study was discussed in the background to rhythm (Chapter 3) as being beyond the scope of this study and difficult to do with little research of correlates of prominence in Kriol, i.e. in terms of speech production and acoustics. Now that some groundwork has been laid in this thesis, future work could do further investigation into the perception of rhythm (e.g. perception analyses) in this variety and how it might differ to English.

In a model proposed by Kohler (2009b) for incorporating perception into a rhythm study rather than solely focusing on durations of segments, syllables or feet, one must first study how prominence is perceived in a given language, such as how prominence is signaled by the interaction of many phonetic factors including syllable duration, F0 changes, syllabic energy and spectral dynamics (all of which have been correlated with stress). Then the contributions of all of these factors to prominence in that language are scaled (Kohler, 2009b for how to do this), and different kinds of rhythm are simulated for perceptual evaluation. Kohler argued that this method eliminates subjectivity when comparing languages rhythm. Another example of work involving rhythm and perception discrimination is a model that combines traditional rhythm metrics with metrics based on auditory cues that mimic ear acoustics (Selouani, et al., 2015). Their model was found to be more successful than rhythm metrics alone at discriminating between native and non-native speech in Arabic. It is possible that the rhythm of Kriol and English are indeed different despite what was found in the acoustic analyses in this thesis (which suggested they were more similar than different). It would be interesting to see in future research if there may be a difference in stress patterns or other factors related to prominence.
Despite some limitations, this project is the first of its kind and there exist no other detailed acoustic analyses of rhythm using rhythm metrics or of patterns of vowel reduction in any variety of Kriol. As a first-pass analysis, this project therefore sets an adult baseline for metric results and patterns of vowel reduction for Barunga Kriol, e.g. for the analysis of vowel production in child speakers, laying the groundwork for further study of the speech rhythm of in this variety and others. It also adds to the research on English in Australia. Further, this thesis joins the conversation on the use of rhythm metrics in general, showing how they can feasibly be used alongside deeper acoustic analysis to provide insight into language rhythm and aid in cross-linguistic comparison.

The findings of this project open up many questions and interesting topics for further research. For example, it would be interesting to see if Kriol is converging towards English or away in terms of rhythm. For a study looking at this, one might measure the rhythm of Barunga Kriol across different generations, and perhaps also a local variety of English such as in Katherine, NT (the nearest town where speakers of Barunga Kriol visit frequently).

One factor which might influence speakers to sound more like English rhythm in their Kriol would be bilingualism or multilingualism with English. There is work in bilingual rhythm suggesting that rhythm can transfer from a speaker’s L1 to L2 (Yuan, 2010) (e.g. Benson, 2002; Yuan, 2010). The English speakers in this study were monolingual (controlled), but the Kriol speakers were also fluent speakers of English. The extent to which this has an effect on the rhythm of Kriol measured in this study remains unknown. If Kriol speakers also speak English, it would be interesting to see if L2 English showed a rhythm different to the rhythm of English of monolingual English speakers. In such a study, the speech of bilingual Kriol-English speakers could be recorded and measured using rhythm metrics, and compared to monolingual English speakers. If the rhythm of English spoken by bilingual Kriol-English speakers was found to be different to monolingual English, this would also suggest that Kriol does indeed have different rhythm to English, and may shed light on some of the findings in the present study.

And lastly, it would be interesting to measure the rhythm of different varieties of Kriol, to compare them and to see if there is an effect of rhythm from possible different substrate languages in other varieties. If speech data for the possible substrate languages were to be available, this could be measured too.
Outcomes of this study have potential applications beyond academia. It is known that cross-linguistically, rhythm is a major contributor to intelligibility: failure to master patterns of timing and stress when learning a language can lead to foreign accent (Faber, 1986; Taylor, 1981). Therefore, this research has the potential to inform teaching practice (e.g. English as Second Language instruction) and teacher professional knowledge.

Additionally, as this research is the first of its kind describing rhythm in Barunga Kriol in adult speech, it provides groundwork needed for more sophisticated research into children’s language development towards adult speech in this language (e.g. Payne et al., 2011). Children whose home language is Barunga Kriol are exposed to some English in the years before school but this becomes much more intensive once they start school, so more linguistic work on Barunga Kriol may raise awareness of subtle language differences that are educationally relevant, and support teachers and parents in working together to optimise children’s educational and life chances.
References


References


References


References


References


Gabriel, C., & Kireva, E. (2014). Prosodic Transfer in Learner and Contact Varieties: Speech


References

*Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.


Koopmans-van Beinum, F. . (1980). *Vowel contrast reduction: an acoustical and perceptual*
References

study of Dutch vowels in various speech conditions. University of Amsterdam.


America, 129(5), 3258–3270.


References


D. Crystal (Ed.), *Linguistic Controversies* (pp. 73–79). London: Edward Arnold.


Siegel, J. 2011. Substrate reinforcement and the retention of Pan-Pacific Pidgin features in modern contact varieties In C. Lefebvre (ed.), *Creoles, Their Substrates, and Language Typology.* (pp. 531-556). Amsterdam/Philadelphia: Benjamins,


APPENDICES

Appendix 1: Coding Protocols

At the data segmentation stage, adherence was made to the following criteria:

- Vowel onset was marked at the zero crossing at the start of the first pitch period associated with clear vocalic formant structure, as well as an increase in intensity at higher frequencies, as judged in the spectrogram. Where a vowel followed a pause, if it was slightly glottalised at the start (identified by changes in the shape of pitch periods such as attenuation and lengthening), the onset boundary of the vowel included this period of irregularity, at the zero crossing at the start of the first pitch period.

- Vowel offset was marked at the zero crossing at the end of the last pitch period associated with the break in regular formant structure of the vowel, as well as a drop in intensity at high frequencies. Where a pause followed a vowel and there was slight glottalisation before the pause, this was also included within the vocalic interval.

- Where a vowel followed a fricative, the vowel onset was marked at the zero crossing following the end of frication, and the appearance of formant structure associated with the vowel. Where a vowel preceded a fricative, the vowel offset was marked at the zero crossing prior to the onset of frication, even if there was still some formant structure.

- Where a vowel followed a nasal, the vowel onset was marked at the zero crossing prior to an increase in intensity at higher frequencies and a change in formant structure associated with the vowel. Where a vowel preceded a nasal, the vowel offset was marked at the zero crossing prior to a decrease in intensity at higher frequencies and a change in formant structure associated with the nasal.

- Where a vowel followed a liquid or glide, if the transition between the vowel and liquid/glide was smooth, the vowel onset was marked in the middle of the transition. Where a vowel preceded a liquid or glide, the vowel offset was marked in the middle of the transition.
• Where a stop released into a vowel, the vowel onset was marked after the release burst, at the zero crossing at the start of the first pitch period associated with vocalic formant structure and an increase in intensity at higher frequencies.

• Where a stop followed a pause, the onset of the stop was marked at the zero crossing immediately preceding the burst, or it was just left at the point where it had been automatically aligned, as utterance initial consonants were excluded in final analysis (see Chapter 3, Section 3.2.5). Where a stop preceded a pause, aspiration following release was always included within the consonantal interval, with the boundary being placed in the end of the burst release, where no more friction could be detected. In the case of multiple releases, the boundary was placed after the last burst.

• Silent pauses denoted by <mp:> in the transcription were excluded from measured intervals and interval durations on either side were summed (Grabe and Low, 2002; White and Mattys, 2007). Any glottalisation (denoted by <gp:>) between vowels were treated as silent pauses and omitted (White and Mattys, 2007).
Appendix 2: Rhythm Metric Results Including Final Segment

As mentioned in Chapter 3, there is no standard methodology for using rhythm metrics. Following a review of the literature, one methodological decision that was made for this study was to exclude the final segment of each utterance. For the sake of transparency and considering in future research, this appendix presents the rhythm metric results with the final segment included.

Overall, similar results were found to the main analysis, with no significant differences between Kriol and English for VarcoV and VarcoC. A difference for nPVI-V measurements was also found, as in the main analysis. The only result that changed was for %V, where the significant different in the main analysis was lost when the final segment was included.

The results of *lmerTest* on a linear mixed effects model with fixed effect of Language (Kriol vs. English), dependent variable as metric (VarcoV, VarcoC, nPVI-V or %V) and random effects, Speaker and Storybook, are shown in Table 8 for both languages. The mean (M) and standard error (SE) here and in all descriptive statistics tables in this thesis are estimates of the class means that would be expected had the groups been of equal size, established by ‘least squares means’ (Harvey, 1960) using the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2013). Significances (p values less than 0.05) are in bold in all tables.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Language</th>
<th>Kriol (N=261)</th>
<th>English (N=282)</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VarcoV</td>
<td></td>
<td>64.148</td>
<td>69.137</td>
<td>-0.209, 10.187</td>
<td>8.8</td>
<td>2.18 (p=0.058)</td>
</tr>
<tr>
<td>VarcoC</td>
<td></td>
<td>59.064</td>
<td>58.830</td>
<td>-5.588, 5.120</td>
<td>10.6</td>
<td>-0.097 (p=0.925)</td>
</tr>
<tr>
<td>nPVI-V</td>
<td></td>
<td>57.571</td>
<td>69.828</td>
<td>6.845, 17.668</td>
<td>9.6</td>
<td>5.074 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>%V</td>
<td></td>
<td>43.862</td>
<td>41.768</td>
<td>-4.478, 0.291</td>
<td>7.9</td>
<td>-2.027 (p=0.077)</td>
</tr>
</tbody>
</table>

Table 8 Descriptive Statistics for Rhythm Metric results by Language, with Final Segment Included (estimates are from statistical models)

Significance codes: ‘***’ is p< 0.001, ‘**’ is p<0.01 and ‘*’ is p<0.05.

VarcoV

Figure 24 is a box plot for VarcoV scores by Language with final segment included, showing that the English and Kriol results were not different, even with final segment included in analysis. There was no effect of Language on VarcoV results when final segment was included ($\chi^2(1) = 3.6502$, p=0.05606).
Figure 25 is a box plot for VarcoC scores by Language, showing that the English and Kriol results were the same, even with final segment included in analysis. There was no effect of Language on VarcoC results, when final segment was included ($\chi^2(1) = 0.0093, p=0.923$).
Figure 25 VarcoC Scores for Kriol and English (Final Segment Included)

nPVI-V

Figure 26 is a box plot for nPVI-V scores by Language, showing that the English results were higher than Kriol. There was an effect of Language on nPVI-V results ($\chi^2(1) = 12.506, p<0.001$). In Table 8 above it can be seen that with inclusion of final segment, English was still higher than Kriol in nPVI-V score, by slightly more this time, 12.2567 ($t(9.6)=5.0735, p<0.001$).
Figure 26 nPVI-V scores for Kriol and English (Final Segment Included)

%V

Figure 27 is a box plot for %V by Language, showing that the English and Kriol results were the same, with final segment included in analysis. There was no effect of Language on %V, when final segment was included ($\chi^2(1) = 3.3501, p=0.0672$).
Figure 27 %V for Kriol and English (Final Segment Included)
Appendix 3: Speaker Variation for Rhythm Metric Results

This appendix presents boxplots of rhythm metric measurements for each speaker. 'Speaker' was included as a random factor in the linear mixed effects models in the main analysis. However, individual speaker behaviour was not investigated.

Figures 28 and 29 show boxplots for VarcoV and VarcoC, respectively. It can be seen that mostly, all speakers from Kriol and English have similar results. The same could be said for Figure 30 and 31.

Figure 28 VarcoV Scores for Speakers of Kriol and English
Figure 29 VarcoV Scores for Speakers of Kriol and English
Figure 30 nPVI-V Scores for Speakers of Kriol and English
Figure 31 %V for Speakers of Kriol and English
# Appendix 4: Function and Content Words Analysed in the Vowel Reduction Study

<table>
<thead>
<tr>
<th>Word</th>
<th>Language</th>
<th>Function or content</th>
<th>Vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>then</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>them</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>get</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>gets</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>else</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>next</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>when</td>
<td>english</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>thet</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>en</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>det</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>den</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>dem</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>wen</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>then</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>them</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>an</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>and</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>des</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>get</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>ged</td>
<td>kriol</td>
<td>function</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>head</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>bed</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>end</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>pen</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>pet</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>lets</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>rest</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>said</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>well</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>yes</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>help</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>legs</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>set</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>slept</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>tell</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>tells</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>went</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>yells</td>
<td>english</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>hed</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>fren</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>bed</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>end</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>Word</td>
<td>Language</td>
<td>Type</td>
<td>Pronunciation</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>---------------</td>
</tr>
<tr>
<td>nes</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>pet</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>ded</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>beng</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>bek</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>bred</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>med</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>men</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>set</td>
<td>kriol</td>
<td>content</td>
<td>/ɛ/</td>
</tr>
<tr>
<td>big</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>picks</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>cliff</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>pick</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>bit</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>kids</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>kid</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>licks</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>sit</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>sits</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>sticks</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>sting</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>thing</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>things</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>think</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>tips</td>
<td>english</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>slip</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>kid</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>stil</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>ting</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>big</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>stik</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>stili</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>tim</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>silip</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>itj</td>
<td>kriol</td>
<td>content</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>is</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>in</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>with</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>his</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>its</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>it</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>him</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>this</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>still</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>if</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>will</td>
<td>english</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>jis</td>
<td>kriol</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>bin</td>
<td>kriol</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>im</td>
<td>kriol</td>
<td>function</td>
<td>/ɪ/</td>
</tr>
<tr>
<td>Kriol Function</td>
<td>English Content</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>dij</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>kriol</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>til</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ji</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>is</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>mi</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>dis</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>this</em></td>
<td>/ɪ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>up</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>but</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>some</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>just</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>much</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>last</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>na</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ba</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>la</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>bat</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>bla</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>comes</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>come</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>runs</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>fun</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>glass</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>mum</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>fast</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>lungs</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>past</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>stuff</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>trunk</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ja</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>ran</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>wel</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>bas</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>lat</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>stat</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>dal</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>staf</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>das</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>fas</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>gats</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>jamp</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>kam</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>lak</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>lantj</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>pas</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>skrab</em></td>
<td>/ʌ/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 5: Vowel Reduction Study Figures using Raw Durational Data

Duration by Word Type (Raw Durations)

Kriol

English

Figure 32 Duration (Raw) by Word Type
Figure 35 Duration (Raw) by Utterance Position and Vowel
Appendix 6: Vowel Reduction Study Results of Duration and Vowel Interaction Model

Some significant effects were found in terms of individual vowels, which were not found as main effects across all vowels. The group sizes however, were small. In the interests of transparency and in case future researchers would like to explore similar research questions, the analyses of utterance position, vowel type and duration interaction are included in this appendix. The results of lmerTest on a linear mixed effects model with fixed effects of Utterance Position and Vowel and dependent variable Duration is shown in Table 9 below. Speaker, Word, Preceding and Following Place of Articulation were included as random effects. Significances are denoted in bold.

For Kriol, all vowels from medial position were significantly shorter than from final position. For /ʌ/, vowels in medial position were significantly shorter than final by 0.5755422 log seconds (equivalent to seconds) \( t(54.7) = 4.1466, p<0.001 \). Initial /ʌ/ were also significantly shorter than final, by 0.5827914 log seconds (equivalent to seconds) \( t(81.0) = 3.6785, p<0.001 \). Initial and medial /ʌ/ were not significantly different.

For /ɛ/, vowels in medial position were significantly shorter than final by 0.1515765 log seconds (equivalent to seconds) \( t(831.4) = 2.3145, p<0.05 \). Initial /ɛ/ was also significantly shorter than final, by 0.4169456 log seconds (equivalent to seconds) \( t(842.5) = 5.8535, p<0.001 \). Medial /ɛ/ was significantly longer than initial, by 0.2653691 log seconds (equivalent to seconds) \( t(771.0) = -6.3965, p<0.001 \).

For /ɪ/, vowels in medial position were significantly shorter than final by 0.4468413 log seconds (equivalent to seconds) \( t(836.5) = 7.5928, p<0.001 \). Initial /ɪ/ was also significantly shorter than final, by 0.5664365 log seconds (equivalent to seconds) \( t(855.4) = 4.6057, p<0.001 \). Medial and Initial /ɪ/ were not significantly different.

Similar patterns to Kriol were held across vowels in English, with all in medial and initial positions being significantly shorter than final positions. No vowels showed any differences between initial and medial positions. For /ʌ/, vowels in medial and initial positions were significantly shorter than those in final positions, by 0.416668 log seconds (equivalent to seconds) \( t(837.0) = 4.16668, p<0.001 \).
seconds) \((t(380.5)=2.9637, p<0.01)\) and \(0.569038 \log\) seconds \((t(382.6)=2.6831, p<0.01)\) respectively.

For /ɛ/, vowels in medial and initial positions were significantly shorter than those in final positions, by \(0.654416 \log\) seconds (equivalent to seconds) \((t(356.8)=6.4012, p<0.001)\) and \(0.886917 \log\) seconds \((t(348.7)=4.5731, p<0.001)\) respectively. For /ɪ/, vowels in medial and initial positions were significantly shorter than those in final positions, by \(0.574846 \log\) seconds (equivalent to seconds) \((t(370.0)=5.9069, p<0.001)\) and \(0.416826 \log\) seconds \((t(350.9)=3.3940, p<0.001)\) respectively.
<table>
<thead>
<tr>
<th>Group</th>
<th>Comparison</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All vowels (N=869)</td>
<td>Initial - Medial</td>
<td>-0.228,-0.034</td>
<td>864.7</td>
<td>-2.646 ** (p=0.008)</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>0.3819, 0.663</td>
<td>215.3</td>
<td>7.296 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>0.281, 0.501</td>
<td>99.4</td>
<td>7.066 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>/a/ (N=185)</td>
<td>Initial - Medial</td>
<td>-0.170, 0.155</td>
<td>844.5</td>
<td>-0.0878 (p=0.9301936)</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>0.268, 0.898</td>
<td>81.0</td>
<td>3.679 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>0.297, 0.854</td>
<td>54.7</td>
<td>4.147 *** (p&lt;0.001)</td>
</tr>
<tr>
<td>/ɛ/ (N=413)</td>
<td>Initial - Medial</td>
<td>-0.347, 0.184</td>
<td>771.0</td>
<td>-6.397 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>0.277, 0.557</td>
<td>842.5</td>
<td>5.854 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>0.023, 0.280</td>
<td>831.4</td>
<td>2.315 * (p=0.021)</td>
</tr>
<tr>
<td>/i/ (N=271)</td>
<td>Initial - Medial</td>
<td>-0.342, 0.103</td>
<td>854.0</td>
<td>1.057 (p=0.291)</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>0.325, 0.808</td>
<td>855.4</td>
<td>4.606 *** (p&lt;0.001)</td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>0.331, 0.562</td>
<td>836.5</td>
<td>7.593 *** (p&lt;0.001)</td>
</tr>
</tbody>
</table>

Table 9 Descriptive Statistics for Duration by Utterance Position and Vowel with Interaction (estimates are from statistical models)

Significance codes: *** is p<0.001, ** is p<0.01 and * is p<0.05
Appendix 7: Vowel Reduction Study Results for Dispersion and Vowel Interaction Model

The results of *lmerTest* on a linear mixed effects model with fixed effects of Utterance Position and Vowel and dependent variable Dispersion is shown in Table 10 on following page. Speaker, Word, Preceding and Following Place of Articulation were included as random effects. Significances are denoted in bold.
Table 10 Descriptive Statistics for Dispersion by Utterance Position and Vowel with Interaction (estimates are from statistical models)

<table>
<thead>
<tr>
<th>Kriol (N=869)</th>
<th>Group</th>
<th>Initial (N=185)</th>
<th>Medial (N=582)</th>
<th>Final (N=102)</th>
<th>Comparison</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
<td>N</td>
<td>M</td>
<td>SE</td>
<td>N</td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.251, 0.384</td>
<td>106.7</td>
<td>0.416 (p=0.678)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.176, 0.319</td>
<td>43.0</td>
<td>0.582 (p=0.564)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All vowels (N=869)</td>
<td>/ʌ/ (N=185)</td>
<td>0.332</td>
<td>0.340</td>
<td>28</td>
<td>0.423</td>
<td>0.317</td>
<td>139</td>
<td>0.300</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.742, 0.679</td>
<td>34.3</td>
<td>-0.089 (p=0.929)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.755, 0.510</td>
<td>22.2</td>
<td>-0.402 (p=0.692)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɛ/ (N=413)</td>
<td>0.872</td>
<td>0.313</td>
<td>145</td>
<td>0.859</td>
<td>0.308</td>
<td>234</td>
<td>0.794</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.397, 0.240</td>
<td>837.0</td>
<td>-0.484 (p=0.629)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.358, 0.228</td>
<td>824.7</td>
<td>-0.437 (p=0.662)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɪ/ (N=271)</td>
<td>0.740</td>
<td>0.309</td>
<td>12</td>
<td>0.832</td>
<td>0.380</td>
<td>209</td>
<td>1.142</td>
<td>0.320</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.240, 0.860</td>
<td>850.7</td>
<td>1.106 (p=0.269)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>0.139, 0.665</td>
<td>634.4</td>
<td>3.000 ** (p=0.003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>English (N=389)</th>
<th>Group</th>
<th>Initial</th>
<th>Medial</th>
<th>Final</th>
<th>Comparison</th>
<th>95% CI for Mean Difference</th>
<th>df</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
<td>N</td>
<td>M</td>
<td>SE</td>
<td>N</td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>0.107, 0.952</td>
<td>383.0</td>
<td>2.465 * (p=0.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.126, 0.406</td>
<td>382.6</td>
<td>1.036 (p=0.301)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Vowels (N=389)</td>
<td>/ʌ/ (N=65)</td>
<td>-0.492</td>
<td>0.453</td>
<td>7</td>
<td>0.135</td>
<td>0.324</td>
<td>47</td>
<td>0.343</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.011, 1.682</td>
<td>380.0</td>
<td>1.941 (p=0.053)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.350, 0.766</td>
<td>379.2</td>
<td>0.734 (p=0.464)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɛ/ (N=82)</td>
<td>-0.862</td>
<td>0.460</td>
<td>6</td>
<td>-0.281</td>
<td>0.326</td>
<td>41</td>
<td>-0.270</td>
<td>0.335</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.199, 1.382</td>
<td>346.0</td>
<td>1.472 (p=0.142)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.402, 0.424</td>
<td>374.8</td>
<td>0.051 (p=0.960)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ɪ/ (N=242)</td>
<td>-0.114</td>
<td>0.314</td>
<td>31</td>
<td>-0.152</td>
<td>0.285</td>
<td>186</td>
<td>0.0488</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td>Final - Initial</td>
<td>-0.330, 0.655</td>
<td>375.0</td>
<td>0.648 (p=0.517)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Final - Medial</td>
<td>-0.190, 0.592</td>
<td>354.7</td>
<td>1.010 (p=0.313)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
From the interaction models some significances appeared which were not present in the non-interaction models. For Kriol, /ɪ/ vowels in medial position were significantly less dispersed than final position, by 0.310 dispersion units (t(834.4)=3.0001, p<0.05). These were visible in Figures 10 and 11, where there was a large difference in dispersion for medial /ɪ/ compared to final /ɪ/.

For English, main effects became significant where they had not been in the simpler model. Collapsed across vowels, all from initial position were significantly less dispersed than those from both final and medial positions, by 0.887 dispersion units (t(383.0)=2.4647, p<0.05) and 0.390 dispersion units (t(382.4)=2.1003, p<0.05) respectively. However, these were not particularly apparent in any visualisations.