Universität Paderborn

Fakultät für Kulturwissenschaften

Accents of English in English Language Teaching in Germany - A Corpus-Phonological Approach to Textbook Analysis

Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.) im Fach Englische Sprachwissenschaft der Universität Paderborn

Von

Lisa Scheiwe

Betreuerin

Prof. Dr. Ilka Mindt

Acknowledgements

I am incredibly thankful to my supervisor Prof. Dr. Ilka Mindt for taking me on as a PhD student and for her guidance on this project. I would also like to thank my second supervisor Prof. Dr. Christian Langstrof for his support and feedback.

Overall, this dissertation would not have been possible without the support and encouragement of many. I am grateful to my colleagues at the Department of English and American Studies at Paderborn University for their insightful discussions during our *Kolloquium*, their feedback and overall support. I am also grateful to our student assistants Noa, Nicholas, and Marie for their support and insights.

I would like to express my deepest gratitude to my developmental editor, Bailey Lang, for her feedback on the structure and flow of my early drafts. Your support and encouragement helped me get this project over the finish line.

To my family and friends: your patience and encouragement helped me through this journey. At the risk of forgetting someone, I still want to mention a few people specifically. Thank you to my parents and grandparents for always believing in me and cheering me on no matter what.

I would like to thank my *Thrive Squad*: Our frequent online writing sessions and social gatherings helped me get to the finish line. Thank you so much for your support over the last years.

Thank you, Mara, for going on multiple writing retreats with me. These weeks away helped me really focus on this thesis and move it forward.

I am grateful to Charlotte, Rebecca, Stefan and Ben for the constant support, and for the distractions when I needed them.

I am most grateful to Ben and Charlotte. Thank you for everything! I couldn’t have finished this without you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AAVE	African American Vernacular English
CT2	Camden Town 2
CT4	Camden Town 4
EFL	English as a foreign language
EGA2	English G Access 2
EGA4	English G Access 4
EGA5	English G Access 5
ELT	English Language Teaching
F1	first formant
F2	second formant
F3	third formant
GA	General American
GL2	Green Line 2
GL4	Green Line 4
GL5	Green Line 5
Hz	hertz
LPC	linear predictive coding
NRW	North Rhine-Westphalia
R1	rhotic variety
R2	non-rhotic variety
R3	broad non-rhotic variety
RP	Received Pronunciation

1 Introduction

To what extent are different varieties of English addressed or integrated into English language teaching (ELT)? Researchers within the fields of World English studies, applied linguistics, and ELT studies have explored this question from different perspectives. A strong consensus among these different research fields is that it is important to introduce learners of English to the diversity of the language to prepare learners for today's globalised world.

As a part of this discourse, this thesis offers a phonetic analysis of three different varieties of English currently included in ELT textbooks used in German secondary schools to evaluate the representation of these varieties in German ELT teaching materials. By using a corpus-phonological approach to textbook analysis, this research offers a new approach toward analysing the representation of different varieties in ELT teaching materials. Where previous research focused on whether individual varieties of English were present in teaching materials, this study focuses on how varieties of English are presented in these materials. This corpus-phonological approach to analysing the representation of varieties of English in teaching materials positions this study at the intersection of World Englishes research, corpus linguistics, applied linguistics, and ELT research.

Section 1.1 introduces the concept of World Englishes, before elaborating on the implications of the global spread of English on applied linguistics and ELT research. Section 1.2 provides insights into the curricular requirements of North Rhine-Westphalia that were in place at the start of this study and evaluates which role World Englishes play. Section 1.3 presents previous research on the inclusion of different varieties in English teaching materials and presents the corpus-based approach to textbook analysis of this study. The results of different studies on the inclusion of varieties of English in ELT teaching materials will be presented to establish the research gap this thesis aims to fill. Section 1.4 then introduces the research question and provides an overview of the structure of this thesis.

1.1 World Englishes and English Language Teaching

English is often said to be *the* global language as "[it] has become part of the daily lives of many people from diverse linguistic and cultural backgrounds, and this is also true in countries where it is not a primary language but functions as either a second language or has a supranational function" (Galloway and Rose 2015: 11). According to Seargeant (2012: 50),

"close to two billion people – that is, almost a third of the world's population – have a certain competence in English." The vast majority of these are non-native speakers of English (Seargeant 2010: 50). Between 350 and 380 million people have English as their first language (Seargeant 2010: 49). While "[e]stimates of non-native English speakers are even more difficult to determine with anything approaching certainty" (Seargeant 2010: 49), it can be assumed that there are roughly 600 million speakers who have English as a second language. The number of English speakers who learn English as foreign language have been estimated to be around one billion speakers (Jenkins 2015: 11).

Researchers have created many linguistic models to categorise the spread of English worldwide, but one of the most popular models is the Three Circles Model by Braj B. Kachru (1985). This model represents "the types of spread, the patterns of acquisition and the functional domains in which English is used across cultures and languages" (Kachru 1985: 12). The acquisition and use of English can be visualised in terms of three concentric circles consisting of the *Inner Circle,* the *Outer Circle* and the *Expanding Circle* (Kachru 1985).

The Inner Circle consists of countries in which English is the first language for the majority of people, so the US, the UK, Australia, Canada, and New Zealand belong to the Inner Circle (Kachru 1985: 12). The different Inner Circle varieties of English are "typically endonormative in orientation and norm-providing since they serve as models for the Expanding Circle and partly also for the Outer Circle" (Buschfeld and Kautzsch 2020: 54).

In the Outer Circle, English is used as an official language in "countries like India, Kenya, and Singapore" (Buschfeld and Kautzsch 2020: 54) and is often used in areas such as education and administration (Kachru 1985: 13).

The Expanding Circle encompasses all countries that "do not necessarily have a history of colonization by the users of the [Inner Circle]" (Kachru 1985: 13). Countries such as Germany, France, or Russia belong to the Expanding Circle. English is taught as a foreign language in these countries (Kachru 1992: 3) and they often "lay no claim to an indigenous variety of English and accept an exonormative standard" (D'Souza 1999: 272).

Many scholars have pointed out the limitations of this model (cf. Bruthiaux 2003; Modiano 1999; Nelson 2011: 17-21; Seargeant 2012: 152-153). However, Kachru's (1985) Three Circles Model "still offers the most convenient framework (…) for thinking about different kinds of English use" (Jenkins 2015: 15) and is therefore also used in this study.

In the field of ELT, there has traditionally been a significant emphasis on displaying English as the language of the Inner Circle, as 'native' English has often been maintained as the 'standard' in ELT (Galloway, 2017: x; Matsuda and Matsuda, 2018: 65). The

standardized varieties of American English and British English have commonly been adopted as the instructional models and target varieties in ELT (Matsuda, 2013: 1; Matsuda and Matsuda, 2018: 66), especially in contexts where English is used as a second or foreign language (Bayyurt, 2018: 412; Matsuda and Friedrich, 2012: 23).

There are some practical reasons for this strong focus on Inner Circle varieties in general, and British and American English in particular, in ELT. The standardized varieties of British and American English are codified varieties of English (Matsuda, 2013: 1; Matsuda and Friedrich, 2012: 22; Seargeant, 2012: 67), which means that, for example, normative grammar and pronunciation rules exist, and grammar books and dictionaries of these varieties are available. Since Great Britain and the United States of America both have large ELT industries, there is a high availability of teaching materials using either British or American English as the target variety and instructional model (Bhowmik, 2015: 143; Seargeant, 2012: 67).

Several ideological reasons also contribute to this focus on Inner Circle varieties. Both reference varieties, British and American English, enjoy a high societal prestige and international currency (Matsuda and Friedrich, 2012: 23; Seargeant 2012: 66) resulting in favourable attitudes towards these varieties. Numerous attitude studies with key stakeholders in ELT—for instance educators, teacher trainees, or language learners—have consistently revealed that Inner Circle varieties of English, particularly British and American English, are often preferred to Outer or even Expanding Circle varieties of English (cf. Davydova 2015; Hartmann 2021; Meer, Hartmann, and Rumlich 2021).

1.2 The Role of World Englishes in the Secondary School Curriculum in Germany

This section examines the status of World Englishes in the German secondary school curriculum, focusing on the curriculum for the lower level (*Sekundarstufe I*) at German grammar schools (*Gymnasium*) in the state of North Rhine-Westphalia (NRW)—the relevant curriculum for the textbooks analysed in this study. The curriculum presents competences students should achieve at the end of year 6, the end of year 8, and the end of year 9.

Meer (2021) investigates how World Englishes are incorporated in the curricula for German grammar schools (in 2018, the time of data collection) in all 16 German states. He focuses on the explicit mention of specific varieties of English and English-speaking countries in the curricula. He also analyses how World Englishes are relevant in the development

of competencies required by the curriculum, such as communicative competence. The results relevant to this study are presented in the following paragraphs. In instances when Meer (2021) does not differentiate the results with regard to the individual states, the curriculum for NRW (MSW NRW, 2007) is consulted to elaborate on the relevant results.

The only varieties of English explicitly mentioned in the curriculum are British and American English (Meer, 2021: 92). The curriculum refers to three English-speaking countries: Great Britain, the United States, and Ireland (MSW NRW, 2007: 16; 22), while also noting that one further English-speaking country should be covered, but without specifying which country this should be (MSW NRW, 2007: 22). Meer (2021: 90) observes that "[w]hile these references [to English-speaking countries] do not necessarily imply that students will encounter specific varieties of English spoken in these countries as such, they provide an overall idea of which Englishes students may potentially come in contact with." Therefore, German students of English at the lower level of NRW grammar schools are likely only encountering Inner Circle varieties of English with a strong focus on British and American English.

Meer (2021) also investigates specific examples of varieties of English or English-speaking countries in the curriculum and considers how World Englishes implicitly contribute to or are mentioned in the descriptions of individual competencies. Overall, Meer (2021: 95-97) establishes that World Englishes predominantly play a role in acquiring communicative competencies in audio- and audio-visual comprehension, as well as speaking.

In the category of audio- and audio-visual comprehension at the end of year 6, students must be able to understand "easy standard language" (MSW NRW, 2007: 23, transl.). At the end of year 8, they also have to be able to understand more heterogeneous speech in the form of "easily recognizable pronunciation variants" (MSW NRW, 2007: 29, transl.). In the sub-competence of speaking, students are required to be able to talk to both native- and non-native speakers of English. Although these communicative encounters can at first occur with prepared topics at the end of year 8 (MSW NRW, 2007: 29), students eventually have to be able to communicate with both native- and non-native speakers even when unprepared at the end of year 9 (MSW NRW, 2007: 36).

World Englishes are also represented in the competence of availability of linguistic resources and language correctness within the subcategory of pronunciation and intonation. The curriculum does not specify a language target for this competence for year 6 (MSW NRW, 2007: 26), but in year 8, students should be able to understand and identify typical pronunciation variants of British- and American English (MSW NRW, 2007: 32). No

explicit language target is mentioned for this competence in year 9; instead, students should demonstrate correct pronunciation and intonation patterns (MSW NRW, 2007: 40). While British- and American English are not always explicitly mentioned here, they still “tend to function as the main target varieties in this context” (Meer, 2021: 96).

Overall, the curriculum displays a strong focus on Inner Circle varieties in general, and British and American English in particular. This trend is perpetuated in teaching materials and observable in textbooks used at NRW grammar schools and world-wide.

1.3 Varieties of English in ELT Materials: a Corpus-Phonological Approach to Textbook Analysis

The central role textbooks play in language teaching and learning is indisputable (Matsuda 2012: 168; Nordlund 2016: 48). Textbooks and the respective audio materials “may provide the major source of English language input” (Richards 2014: 19) for learners (see also: Matsuda 2012: 168). Quality teaching materials are indispensable, especially in the context of exposing learners to different varieties of English. “In order to effectively incorporate linguistic and cultural diversity into English classrooms, the availability of well-designed teaching materials is critical" (Matsuda 2013: 3).

The representation of World Englishes in ELT textbooks has been investigated for many specific EFL (English as a foreign Language) contexts ranging from Japan (Matsuda 2002) and Hong Kong (Chan 2020) to European contexts (Syrbe and Rose 2018; Schildhauer, Schulte and Zehne 2020 for Germany, Kopperoinen 2011 for Finland, Vettorel and Lopriore 2013 for Italy, and Lindqvist and Soler 2022 for Sweden), but also from a more international perspective (Naji Meidani and Pishghadam 2013).

One of the aspects that has been intensively investigated in these studies is in how far different varieties of English are included in the respective textbooks. Overall, a strong focus on Inner Circle varieties of English and especially British and American English can be observed in several studies. Kopperoinen (2011) showed that the audio materials of two Finnish textbooks used at upper secondary schools strongly focus on Received Pronunciation and General American and only devote little space to both Outer and Expanding Circle accents of English.

A diachronic investigation of ten textbooks used in Italian secondary schools between 2008 and 2013 revealed that all textbooks show “more or less explicit references to

both British ad [sic!] American English" (Vettorel and Lopriore 2013: 493), and the more recent textbook editions also include other Inner Circle varieties of English (ibid.: 494).

Overall, a focus on Inner Circle accents in ELT textbook audio materials is quite common in EFL contexts. Naji Meidani and Pishghadam (2013) explored four internationally used textbooks that were published between 1994 and 2006. Overall, Outer- or Expanding Circle accents of English were used, if at all, in less than 20% of the audio materials, revealing a strong focus on Inner Circle accents in these teaching materials. This focus on Inner Circle accents is also reflected in more local contexts. In his investigation of eight textbooks used in Hong Kong secondary schools from two different publishers spanning a timeframe from 1975 to 2012, Chan (2020) revealed a similar pattern of focusing on Inner Circle accents of English: "[D]espite the recommendation of exposing students to different English varieties (…), [native speaker] English is still the unquestioned choice of accents for the listening materials in all the textbooks examined" (Chan 2020: 254-255).

A similar pattern can also be found in the German ELT context. Syrbe and Rose (2018) analysed three state-approved textbooks used at secondary schools in NRW. All three textbooks displayed a strong focus on Received Pronunciation within the audio materials. Only a few instances of other Inner Circle varieties (American English, Australian English, and South African English) and only individual instances of Outer Circle varieties (Jamaican English and Nigerian English) or Expanding Circle varieties (Finnish English and Spanish English) were present in the audio materials. A more recent study of German ELT textbooks (Schildhauer, Schulte, and Zehne 2020) with one textbook used at the advanced secondary school level and one textbook used at the intermediate secondary school level revealed again a strong focus on Inner Circle accents of English – particularly Received Pronunciation and General American – in the audio materials of the textbook used at the advanced secondary school level. The textbook used at the intermediate secondary school level displayed a similar focus on Received Pronunciation, but not so much on General American. Instead, other Inner Circle accents as well as some Outer Circle accents of English were present in the audio recordings.

In general, these previous studies on the representation of World Englishes in ELT teaching materials reveal a strong focus on Inner Circle varieties of English. These studies mostly focused on which varieties were included in either the textbooks themselves or the audio materials. However, a more in-depth linguistic analysis of the representation of the varieties of English present in these textbooks has not yet been conducted. This thesis aims

to fill that gap by providing a new approach to analysing the representation of different varieties of English in ELT textbooks.

This thesis takes a corpus-phonological approach to textbook analysis with a focus on the representation of different varieties of English. A textbook corpus consisting of eight secondary school textbooks used in NRW and encompassing speech samples from Australian English, British English, and American English has been compiled for this thesis. On the basis of acoustic- as well as auditory analyses, the representation of three select varieties of English in EFL textbook audio materials will be investigated.

While previous research of textbook audio materials either only focused on the occurrence of different accents of English or has solely focused on auditory analyses of different accents of English (cf. Kopperoinen 2011; Schildhauer, Schulte and Zehne 2020; Syrbe and Rose 2018), an acoustic analysis of vowels and an auditory analysis for rhoticity has been selected for this approach. Thus, this thesis offers a more detailed analysis of the representation of different accents of English by exploring how salient accent features of the varieties under discussion are portrayed.

This research therefore contributes to the discourse around the global orientation of textbooks and the representation of the linguistic diversity of the English language by providing detailed insights into the representation of three select accents of English based on acoustic and auditory analyses.

1.4 This Present Research

Situated at the intersection of World Englishes research, corpus linguistics, applied linguistics, and ELT research, this thesis provides a linguistic description of three select varieties of English as portrayed in three textbook series used in EFL classrooms in NRW in Germany to answer the following research question: How are different varieties of English represented in the audio materials of German textbooks for English? Taking a corpus-phonological approach, the representation of British English in the form of Received Pronunciation, American English in the form of General American, and Australian English will be analysed acoustically with a focus on select salient features of each accent.

The analyses of Received Pronunciation focus on the representation of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START in the textbooks *English G Access 2* (Rademacher 2014a), *Green Line 2* (Weisshaar 2015a), and *Camden Town 2* (Hanus et al. 2013a). General American is evaluated based on the monophthongs BATH, LOT, and CLOTH

and rhoticity in the lexical sets NEAR, SQUARE, and CURE in the textbooks *English G Access 4* (Rademacher 2016a), *Green Line 4* (Weisshaar 2017a), and *Camden Town 4* (Claussen et al. 2015a). The analyses of Australian English focus on the representation of the diphthongs FACE, PRICE, and MOUTH in the textbooks *English G Access 5* (Rademacher 2017a) and *Green Line 5* (Weisshaar 2018a).

With this context, the research question of this study can be broken down for each variety of English:

Research Question 1: How is Received Pronunciation represented in the audio materials of the textbooks *English G Access 2*, *Green Line 2*, and *Camden Town 2*, exemplified by the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START?

Research Question 2: How is General American represented in the audio materials of the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4*, exemplified by the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE?

Research Question 3: How is Australian English represented in the audio materials of the textbooks *English G Access 5* and *Green Line 5*, exemplified by the diphthongs FACE, PRICE, and MOUTH?

For each research question the following hypotheses are formulated based on the research findings presented in Sections 1.1 to 1.3.

Hypothesis 1:

The textbooks *English G Access 2*, *Green Line 2*, and *Camden Town 2* consistently present pronunciation variants consistent with Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START.

Rationale: Due to the strong focus on British and American English as instructional models and teaching targets in ELT (Matsuda, 2013: 1; Matsuda and Matsuda, 2018: 66), it can be assumed that the reference accent for British English, Received Pronunciation, is presented consistently in German teaching materials.

Hypothesis 2: The textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4* consistently present a General American accent in the pronunciation of the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE.

Rationale: Due to the strong focus on British and American English as instructional models and teaching targets in ELT (Matsuda, 2013: 1; Matsuda and Matsuda, 2018: 66), it

can be assumed that the reference accent for American English, General American, is presented consistently in German teaching materials.

Hypothesis 3: The textbooks *English G Access 5* and *Green Line 5* display features of Australian English, but not consistently and with a focus on General Australian English.

Rationale: Apart from a strong focus on British and American English in ELT curricula and materials (see Sections 1.2 and 1.3), the only variation presented to learners of English is often in the form of one or several other Inner Circle varieties of English. Thus, it can be assumed that accent features of Australian English are presented to learners, but possibly not consistently. A focus on General Australian English can be expected, as this is the accent spoken by most Australians (Cox 2006: 148).

Outline of this study

Chapter 2 provides insights into the theoretical framework of this study. Section 2.2 introduces important terminology. Sections 2.3 and 2.4 establish the theoretical framework of how accent features can be analysed. Section 2.3 focuses on analysing vowels. Vowel quality is defined from an articulatory, auditory, and acoustic perspective and this section presents acoustic methods of analysing vowel quality. Section 2.4 defines the concept of rhoticity based on Harris' (2013) rhoticity systems and presents common approaches to analysing rhoticity. Sections 2.5 – 2.7 provide linguistic descriptions of Receive Pronunciation (2.5), General American (2.6), and Australian English (2.7). First, a general definition of what constitutes each accent is provided before describing phonological features of each accent. Section 2.8 summarises the characteristics of the features analysed in this study. The vowels analysed for each accent are presented in individual vowel quadrilaterals to illustrate their pronunciation in the respective accent. Section 2.9 concludes with a summary of the theoretical foundations of this study.

Chapter 3 introduces the textbook corpus and provides an overview of the corpus design and the methodological approaches toward the analyses. Section 3.2 illustrates the corpus design of the phonological textbook corpus. After considering practical aspects in the design process, the general corpus design, annotation practices, and the corpus design of the individual sub-corpora are presented. Section 3.3 discusses the processes and methods used in the analysis of the corpus data. Section 3.4 addresses methodological and data limitations of this study.

Chapter 4 presents the results of the acoustic and auditory analyses of the textbook corpus. The three Sections 4.2, 4.3, and 4.4 present the results of the analyses of the corpus data and the comparison with the linguistic descriptions of the accents under discussion presented in Chapter 2. Section 4.2 focuses on Received Pronunciation, Section 4.3 focuses on General American, and Section 4.4 focuses on Australian English.

Chapter 5 discusses findings of this study and highlights the for the teaching and learning realities in German EFL classrooms. The benefits of a corpus-phonological approach are also discussed.

Chapter 6 offers a summary of this study while also looking ahead outlining further research possibilities in the field.

2 Theory

2.1 Introduction

This chapter presents the study's theoretical framework. Section 2.2 defines key terminology used in this study. Section 2.3 defines the concept of vowel quality from three different phonetic perspectives: articulatory, auditory, and acoustic phonetics. Section 2.4 conceptualises rhoticity in different phonological concepts based on Harris' (2013) rhoticity systems. Sections 2.5 – 2.7 introduce the accents analysed in this study: Received Pronunciation, General American, and Australian English. Each section begins with a definition of the respective accent before describing the phonological features of each accent. These descriptions are structured consistently throughout these sections: first, the monophthongs are described from front to central to back. Within each of these categories the monophthongs are described from close to open. Afterwards, first the rising and then the centring diphthongs are presented, followed by a description of interesting consonant features. Each section concludes with a summary of the features chosen for analysis in this study. Section 2.8 summarises the phonological features chosen for the analysis and visualises the relevant vowels for each variety in vowel quadrilaterals.

2.2 Terminology

The terms *variety*, *dialect*, and *accent* are central concepts in this study and need to be defined. The term *variety* can be defined as "[any] form of a language seen as systematically distinct from others" (Matthews 2014c). Mair (2008: 141) states that varieties can, for example, differ based on regional, social, or situational contexts and further clarifies that "[the] identifying features of such a sub-type [of a language] can be orthographic […], phonetic, grammatical, lexical-semantic or pragmatic" (Mair 2008: 141). In this study, the term variety is used to refer to regional varieties of English such as Australian English.

The term *dialect* refers to "any distinct variety of a language, especially one spoken in a specific part of a country or other geographical area" (Matthews 2014b). Mair (2015: 144) specifies the concept of a dialect as "a variety of a language which differs […] from other varieties of the same language not only in its pronunciation, but also in its vocabulary, and in its grammar." Thus, in this study, the terms dialect and variety are understood as interchangeable concepts. The terms dialect and variety are used in this study to refer to a

regionally distinct variety of English, such as British English, American English, or Australian English.

The term *accent* is defined as "[a] variety of speech differing phonetically from other varieties" (Matthews 2014a). Thus, the term accent only refers to the pronunciation of different varieties of English. In this study, British English is represented by the accent Received Pronunciation, American English by the accent General American, and Australian English by the accent Standard (Mainstream) Australian English.

2.3 Analysing Vowel Quality

The English vowel system can be differentiated into monophthongs, which are vowels with a steady vowel quality (Ashby 2011: 108), and diphthongs, which are vowels that change "quality in the space of a single syllable" (Ashby 2011: 108). In order to refer to the individual vowels, this thesis uses Wells' (1982a) standard lexical sets written in SMALL CAPITALS. The lexical sets are a set of keywords "used to refer to (i) any or all of the words belonging to the standard lexical set in question; and (ii) the vowel sound used for the standard lexical set in question in the accent under discussion" (Wells 1982a: 124). Lexical sets are commonly used in linguistic research to talk about the phonological features of vowels in different accents of English, such as Received Pronunciation (cf. Fabricius 2002; 2007; Hinton 2015), General American (cf. Kretzschmar 2008), and Australian English (cf. Bradley 2008).

The following section describes vowel quality from different phonetic perspectives: an articulatory perspective, an auditory and more specifically, a speech perception perspective, and an acoustic perspective. Then the mechanics and procedure of analysing vowels acoustically—the methodology adopted in this study—are described in greater detail.

Traditionally, vowel quality is described from an articulatory perspective: vowels are categorised according to the respective tongue height and tongue advancement during the vowel articulation (Liebermann and Blumstein 1988: 164). Tongue height—sometimes referred to as vowel height—describes how far the tongue body is raised during vowel production. Tongue advancement describes which part of the tongue is raised during vowel articulation: the tip of the tongue, the mid part, or the back (IPA 1999: 10–12).

The vowel quadrilateral is used as a schematic representation of "the position of the tongue in vowel production" (IPA 1999: 10) using the cardinal vowels as reference points. The vowel quadrilateral distinguishes between front, central, and back vowels with respect to tongue advancement, and between close, close-mid, open-mid, and open vowels with

respect to tongue height (Cruttenden 2014: 39). The quadrilateral indicates tongue height using four equidistant horizontal lines which represent auditorily equal steps (IPA 1999: 11). To indicate tongue advancement, the quadrilateral has a right angle in the top right and bottom right corners, and the bottom horizontal line is half as long as the top horizontal line. The central area in the quadrilateral is indicated by connecting the midpoint of the top horizontal line with the midpoint of the bottom horizontal line.

An auditory perspective on vowel quality is interested in how listeners perceive and distinguish vowels. In this sense, speech perception can be seen as "a phonetic mode of listening, [focusing] (…) on the sounds of speech rather than the words" (Johnson 2012: 101). As listeners, we are able to distinguish between and categorize different vowels based on what we are familiar with as speakers (Johnson 2012: 107).

From an acoustic point of view, vowel quality is related to the vowel formant frequencies of the first (F1) and second (F2) formants (Johnson 2012: 142; 144). The articulatory feature of tongue height is negatively correlated to F1, while the articulatory feature of tongue frontness is negatively correlated to F2. Thus, open vowels have a higher F1 value than close vowels, and back vowels have a lower F2 value than front vowels (Johnson 2012: 144; Kent and Read 1992: 92; Roach 2001: 42; Rosner and Pickering 1994: 13). Consequently, a close front vowel such as FLEECE [iː] in Received Pronunciation has a relatively low F1 value, but a relatively high F2 value, whereas an open back vowel such as START in Received Pronunciation has a relatively high F1, but a relatively low F2 value. According to Johnson (2012: 144), "the distinctive features of vowels are tied to these acoustic properties, rather than to articulatory properties," meaning vowel quality is best analysed acoustically.

Acoustic measurements of speech recordings can be conducted with computer programs such as Praat (Boersma and Weenink 2020). Among other features, Praat displays the spectrogram of a sound file and shows the formant frequencies with the help of linear predictive coding (LPC).

Spectrograms visualise frequency information in hertz (Hz) with the frequencies ascending in order on the vertical axis and time on the horizontal axis (Roach 2001: 41; Rosner and Pickering 1994: 6). Amplitude is visualised using colour "with the peaks of the spectrum black, and the valleys white" (Johnson 2012: 78). The individual formants appear as "broad, dark zones" in the spectrogram and are easily identifiable (Thomas 2011: 48).

LPC "estimates the peaks in the acoustic spectrum – the vowel formants" (Johnson 2012: 73) and is thus an important tool in analysing vowel quality acoustically. This method "provides better estimates of formant parameters than does the spectrogram" (Rosner and

Pickering 1994: 8). Nonetheless, LPC is not infallible, and a few limitations have to be considered when working with LPC analyses. Due to the mechanics of LPC (see Johnson 2012 for further explanations), nasalized vowels, nasals, laterals, and some fricatives may not be adequately represented. Furthermore, the number of formants to measure in the spectrogram has to be set in advance. As a result, LPC analysis may show formant readings where none exist if the maximum number of formants to be shown is set too high or may fail to show formants in the spectrogram if the number of anticipated formants is set too low (Johnson 2012: 77). Thus, when working with LPC analyses it is important to check whether the LPC readings coincide with the formants seen in the spectrogram (Thomas 2011: 48).

Measurements for F1 and F2 are taken either from spectrograms or from the LPC readings to describe vowels acoustically. Monophthongs are usually measured at the temporal mid-point of the vowel, which reduces co-articulation effects from neighbouring consonants (Rosner and Pickering 1994: 79). Other methods of measuring vowel targets in both monophthongs and diphthongs are presented in Section 3.3.1. The monophthongs can then be represented in the F1xF2 plane with F1 plotted on the vertical axis, F2 plotted on the horizontal axis, and the origin of the plane in the upper right-hand corner (Rosner and Pickering 1994: 11). Diphthongs are represented in the F1xF2 plane by trajectories starting at the F1 and F2 values of the onglide and ending at the F1 and F2 values of the offglide (Kent and Read 1992: 103). The formant frequencies of diphthongs are frequently taken at 25 per cent through the vowel for the nucleus and at 75 per cent through the vowel for the offglide (Thomas 2011: 151–152). Visualising vowel quality in this fashion provides "a convenient way of representing a given vowel system" (Rosner and Pickering 1994: 13), and the F1xF2 plane also resembles the traditional IPA vowel chart (Johnson 2012: 144). Thus, acoustic vowel quality can ultimately be described using the traditional articulatory features of vowel quality: tongue height and tongue advancement.

Analysing vowels acoustically and plotting the results in an F1xF2 plot is a common method in linguistic research on different accents of English, for instance, Received Pronunciation (cf. Harrington et al. 2000), and Australian English (cf. Butcher 2006; Cox 2006; Harrington, Cox, and Evans 1997).

2.4 Analysing Rhoticity

A rhotic variety of English is a variety in which the consonant /r/ is articulated in all phonological contexts. This includes prevocalic contexts, such as in word-initial positions (e.g. *ready*), in consonant clusters (e.g. *free*), or in word-medial position (e.g. *merry*); preconsonantal contexts (e.g. *party*); and word-final contexts (e.g. *car*). In a non-rhotic variety of English, on the other hand, /r/ is only articulated in prevocalic contexts, for example, in the words *ready*, *free*, or *merry*. These prevocalic contexts, however, also include prevocalic contexts across word boundaries. Therefore, in non-rhotic varieties of English, /r/ is also articulated in word-final contexts if the following word begins with a vowel (e.g. *far away*). In these instances, the articulation of the consonant /r/ is referred to as 'linking-/r/' (Marsden 2017: 275-276).

Harris (2013) describes this difference between rhotic and non-rhotic varieties of English in greater detail in his theory on rhoticity systems of English. By going into a more fine-grained analysis of the distributional differences of /r/ articulation in different phonological contexts in varieties of English, he introduces three rhoticity systems R1, R2, and R3, which "by no means exhaust the set of attested r-systems in English (although their geographical coverage is pretty extensive)" (Harris 2013: 332). These rhoticity systems encompass varieties of English commonly referred to as 'rhotic' varieties (R1), 'non-rhotic' varieties (R2), and varieties that Harris terms 'broad non-rhotic' (R3) (2013: 332). General American is an example of an R1 system, Received Pronunciation and Australian English are examples of R2 systems, and African American Vernacular English (AAVE) is an example of an R3 system. Non-rhotic varieties of English display linking-/r/ as explained above and also 'intrusive-/r/' (Marsden 2017: 276). Intrusive-/r/ has no orthographic representation of /r/ and can occur when a word ends in a vowel and the next word starts with a vowel (e.g. *law and order*) resulting in a pronunciation such as [ˌlɔːɹ ənd ˈɔːdə]. Linking- and intrusive-/r/ "are only used in reference to 'non-rhotic' varieties (…) and are treated as separate phenomena from 'rhoticity proper'" (Marsden 2017: 276).

The three rhoticity systems R1, R2, and R3 can be differentiated by the distribution of /r/ articulation across different phonological contexts (see Table 1 for an overview). Speakers of varieties of English that belong to the system R1, such as speakers of General American, articulate /r/ in all possible phonological contexts (i) – (ix). Speakers of varieties of English belonging to the system R2, such as Received Pronunciation or Australian English, only articulate /r/ in the phonological contexts (i) – (vi). Thus, speakers of varieties

belonging to the R2 system only articulate /r/ in prevocalic contexts but not in preconsonantal or utterance final positions. Speakers of varieties that belong to the R3 system of rhoticity, such as AAVE, only articulate /r/ in prevocalic positions (i) – (iii). So, these speakers only articulate /r/ in word-initial positions, in consonant clusters, and in intervocalic positions within a word before a stressed syllable.

	Phonological contexts of /r/ articulation	**Example word**	**System R1**	**System R2**	**System R3**
(i)	Word initial position in stressed or unstressed syllables	*really, remission*	✓	✓	✓
(ii)	In consonant clusters	*three, photograph*	✓	✓	✓
(iii)	Intervocalic position within a word with succeeding stressed syllable	*arise*	✓	✓	✓
(iv)	Intervocalic position within a word with succeeding unstressed syllable	*merry*	✓	✓	x
(v)	Intervocalic position across a word boundary with succeeding stressed syllable	*for instance*	✓	✓	x
(vi)	Intervocalic position across a word boundary with succeeding unstressed syllable	*for example*	✓	✓	x
(vii)	Preconsonantal position within a word	*third*	✓	x	x
(viii)	Preconsonantal position across a word boundary	*car park*	✓	x	x
(ix)	Word/phrase/utterance final	*How far?*	✓	x	x

Table 1. Overview of Harris' (2013) rhoticity systems (adapted from Marsden 2017: 278).

Linguists conducting studies on varieties of English frequently use auditory and acoustic approaches to analyse rhoticity. An auditory approach to analysing rhoticity usually consists of the researcher, or researchers, listening to the sound samples and determining whether /r/ is articulated or not. This common approach to analysing rhoticity has been followed in linguistic studies on, for example, Scottish English (cf. Meer et al. 2021), English in the Black Country (cf. Asprey 2007), New Zealand English (cf. Bartlett 2002; Gibson 2005; Marsden 2017), New York City English (cf. Becker 2014; Guy 2018; Labov 1997; Mather 2012), and Chinese English (cf. Li and Kibak 2017).

Rhoticity can also be analysed acoustically. An articulation of /r/ is usually marked by a very low frequency in the third formant (F3) (Ladefoged and Johnson 2011: 203; Ladefoged and Maddieson 1996: 244). Such a low frequency in F3 can also be observed in rhotacized vowels such as NEAR or SQUARE in General American, with a considerable fall in frequency in F3 toward the end of the vowel (Ladefoged and Johnson 2011: 231). An acoustic analysis of rhoticity is frequently used in linguistic studies on various accents of English, as well, for example on New Zealand English (cf. Schilk and Pickert 2022), Scottish English (cf. Jauriberry, Sock and Hamm 2015), or American English (cf. Boyce and Espy-Wilson 1997; Kuecker, Lockenvitz and Müller 2015). However, Heselwood and Plug (2011) have shown, that a low F3 is not always a reliable indicator of rhoticity. Consequently, relying solely on acoustic measurements does not always render reliable results regarding rhoticity.

Therefore, a mixed auditory and acoustic approach is highly beneficial in analysing rhoticity and is an established approach in linguistic studies on accents of English. Several studies take an auditory approach to analysing rhoticity, with added acoustic analyses and inspections of the spectrogram for either all or a portion of the tokens (cf. Dickson and Hall-Lew 2017; Lonergan and Cox 2010; Redzwan 2016; Tan 2012) or for difficult-to-determine tokens (Stuart-Smith, Lawson and Scobbie 2014). Overall, analysing rhoticity mainly auditorily with acoustic analyses supporting these auditory judgements adheres to common practice in the field and, therefore, this study also makes use of this practice.

2.5 Received Pronunciation

Received Pronunciation is often defined as "a standardised accent of English" (Trudgill 1999: 118) and is seen as the prestige accent of British English (Zsiga 2013: 433). Associated with the upper and upper middle class of Britain (Wells 1982a: 117), Received Pronunciation is often seen as social rather than a regional accent (Trudgill 1999: 118). Although "it

originated in the south-east of England" (Trudgill and Hannah 2008: 15), it is now seen as a supra-regional accent (Kortmann and Upton 2004: 25; Mugglestone 2017: 165; Murphy 2016: 4; Upton 2008: 239).

Received Pronunciation is often the teaching target in the EFL classroom (Cruttenden 2014: 82). It carries overt prestige and is highly codified (Fabricius and Mortensen 2013: 377): Received Pronunciation, along with General American (see Section 2.6), is, for instance, the accent used in pronunciation dictionaries (cf. Jones 2011; Wells 2008).

The concept of Received Pronunciation as well as the term itself have been strongly debated (Cruttenden 2014: 80; Mugglestone 2017: 151; 162; Schmitt 2016: 30) and several alternatives have been proposed: for instance, *General British* (Cruttenden 2014), *Oxford English*, or *BBC English* (cf. Zsiga 2013: 433). However, the "obvious lack of consensus among different writers" (Mugglestone 2017: 163) poses certain difficulties with these new labels, as well. Received Pronunciation remains the term used to describe the "standard prestige dialect of Contemporary English spoken in England" (Zsiga 2013: 433).

The focus of this section lies on the phonological features of Received Pronunciation and particularly on vowel realisations and select consonantal features.

Received Pronunciation has five front monophthongs: FLEECE, KIT, DRESS, SQUARE, and TRAP. FLEECE is a long close front vowel [iː] which can also be realized with an onglide from KIT resulting in [ɪi] (Upton 2008: 245). KIT is a short close-mid front vowel [ɪ] which "is the norm in unstressed position in the morphemes *-ed, es,* as in *hunted, faces*, and in such words as *minutes, David*" (Upton 2008: 242). DRESS can be described as a short open-mid front vowel which is realized as [ɛ] in Received Pronunciation, although traditionally, this vowel has a slightly closer variant which can be transcribed as [e] (Upton 2008: 241-242). Traditionally, SQUARE used to be a centring diphthong [ɛə][1]; however, it has become monophthongal and is now a long open-mid front monophthong which is realized as [ɛː] (Cruttenden 2014: 80; 84; Upton 2008: 246). TRAP is an open front monophthong that, traditionally, has been transcribed with [æ] (Roach 2004: 243). In recent years, however, this vowel has significantly lowered in RP and is now better transcribed with [a] (Cruttenden 2014: 84; Upton 2008: 242).

Received Pronunciation has two stressed central monophthongs: NURSE, and STRUT. NURSE is a mid central vowel that is traditionally transcribed as [ɜː] (Roach 2004: 242). However, as Upton (2008: 244) observes, there is some variability in vowel height in this

[1] Cruttenden (2014: 80) transcribes the centring diphthong SQUARE as [eə] instead of [ɛə].

vowel, which can range from open-mid to close-mid. He suggests that NURSE is better transcribed as [ɜː]. STRUT is an open-mid central to open central vowel that is traditionally transcribed as [ʌ] (Roach 2004: 242; Upton 2008: 243), although [ɐ] would be a more accurate transcription for this vowel. Due to the lowering of TRAP, the TRAP-STRUT distinction is less pronounced than it used to be for many speakers (Upton 2008: 243).

The vowels GOOSE, FOOT, THOUGHT, NORTH, CLOTH, LOT, BATH, PALM, and START are usually back vowels in Received Pronunciation (Roach 2004: 242). GOOSE is a close back vowel in Received Pronunciation realized as [uː], and FOOT is a close-mid back vowel realized as [ʊ] (Roach 2004: 242; Upton 2008: 241). However, both vowels display considerable fronting to a more central position, especially among younger speakers, resulting in pronunciations such as [ʉː] and [ʊ̟] respectively (Cruttenden 2014: 84). GOOSE may even have fronted further to a close front position which would be indicated by [yː] (Harrington, Kleber, and Reubold 2011: 151; Jansen and Mompean 2023: 55). The symbol [ʊ̟], however, is not a phoneme symbol existent in the international phonetic alphabet. While this symbol is likely used to match the fronted realisation in GOOSE indicated by [ʉː], the use of diacritics would be more appropriate here. A fronted FOOT vowel can be represented by [ʊ̟] instead. THOUGHT and NORTH are both mid back vowels realized with the same phoneme [ɔː] (Roach 2004: 242; Upton 2008: 241-242). LOT and CLOTH are both fully back and "somewhat less open than secondary cardinal 5" (Wells 1982a: 130; see also Wells 2008: xxiii). They are considered open back vowels realized with the same sound [ɒ] (Roach 2004: 242; Upton 2008: 241). The vowels BATH, PALM, and START are often merged in Received Pronunciation and are traditionally realized as the open back long monophthong [ɑː] (Roach 2004: 242). This description is still true for many speakers. However, BATH is "becoming both increasingly centralized and more shortened" (Upton 2008: 244), especially in words with a succeeding nasal. Speakers with a shortened variant in BATH often still have a long [ɑː] in START (Upton 2008: 246), and by extension also PALM.

FACE, PRICE, MOUTH, GOAT, and CHOICE are closing diphthongs—sometimes also referred to as rising diphthongs—in Received Pronunciation. The diphthongs FACE, PRICE, and CHOICE are front rising diphthongs moving towards KIT, while MOUTH and GOAT are back-rising diphthongs moving towards FOOT (Roach 2004: 242; Upton 2008: 245-246). FACE has a front starting point just below close-mid (Upton 2008: 245) and is transcribed as [eɪ] (Roach 2004: 242). PRICE has an open central starting point (Upton 2008: 245) and is transcribed as [aɪ] (Roach 2004: 242). CHOICE starts from a mid back to an open-mid back position (Jones 2011: vii; Upton 2008: 246; Wells 2008: xxiii) and is transcribed as [ɔɪ] (Roach

2004: 242). MOUTH starts from an open central position (Jones 2011: vii; Wells 2008: xxiii) and is transcribed as [aʊ] (Roach 2004: 242). GOAT has a mid central starting point and can be transcribed as [əʊ] (Roach 2004: 242).

Received Pronunciation has three centring diphthongs: NEAR, SQUARE, and CURE. All three diphthongs move toward the mid central area in the vowel chart and are phonetically transcribed as [ɪə], [eə], and [ʊə] respectively (Roach 2004: 242). However, SQUARE and CURE are now increasingly not realized as centring diphthongs for many speakers of Received Pronunciation. SQUARE has increasingly become monophthongal and is realized as an open-mid front long [ɛː] (Cruttenden 2014: 84; Upton 2008: 246). The monophthongal variant of CURE, [ɔː], only seems to affect some monosyllabic words such as *sure* [ʃɔː], but not others, such as *poor* [pʊə] (Cruttenden 2014: 80; Trudgill and Hannah 2008: 16-17).

Three select consonant features are introduced here: the state of rhoticity, /t/, and /l/. Received Pronunciation is a non-rhotic variety of English that exhibits linking- and intrusive-/r/ (Upton 2008: 249; see also Section 2.4). /t/ is often realised as a glottal stop [ʔ] in pre-consonantal position except before /l/ (Cruttenden 2014: 84). Received Pronunciation has what is termed *clear l* and *dark l*. Clear l is realised as [l] before vowels, while dark l is realised as [ɫ] before consonants and at the end of words (Schmitt 2016: 39).

This thesis analyses the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START to investigate the representation of Received Pronunciation in German textbook audio materials. Section 2.8 provides a summary of the features under discussion in this thesis. The next section defines General American and presents the phonological features of this accent.

2.6 General American

The accent of American English that is often associated with "a certain norm or 'standard' of pronunciation" (Kövecses 2000: 81) is General American. This accent has its geographic origins in the area that stretches from "North of the Great Lakes region to the Far West" (Schneider 2011: 81). Despite this broad geographic origin, General American can be described as a "geographically neutral [accent]" (Kövecses 2000: 82), with the majority of Americans speaking General American "or at least a variety or accent very close to it" (Kövecses 2000: 222). General American is often described negatively as an accent without any "marked regional north-eastern or southern characteristics" (Trudgill and Hannah 2008: 49).

This is usually the American English accent taught to EFL learners (Wells 1982a: 118; 2008: xx), especially in parts of Asia and Latin America (Cruttenden 2014: 87). General American, like Received Pronunciation (see Section 2.5), carries overt prestige and is codified: General American is the accent used in pronunciation dictionaries to represent American English (cf. Jones 2011; Wells 2008).

The term *General American* has been subject to debate and the alternative term 'Network English' has been introduced but has not been widely accepted (Schmitt 2016: 30; Schneider 2011: 81; Wells 1982c: 470).

The focus of this section lies on the phonological features of General American and particularly on vowel realisations and select consonantal features. The vowels of General American lack the difference between phonemic long and short vowels (Collins and Mees 2006: 137) that we find, for example, in Received Pronunciation. They are instead differentiated into tense and lax vowels, with tense vowels corresponding to long vowels, and lax vowels corresponding to short vowels. (Jones 2011: viii). Thus, General American vowel features are presented without length markings in this section.

General American has five front monophthongs: FLEECE, KIT, DRESS, TRAP, and BATH (Jones 2011: vii). FLEECE is a close front vowel that is realised as [i] (Kretzschmar 2008: 44; Wells 2008: xxiii; Zsiga 2013: 430). KIT, realised as [ɪ], is a vowel situated between fully close and close-mid in terms of vowel height, and between fully front and central in terms of tongue advancement (Jones 2011: vii; Wells 2008: xxiii). DRESS is transcribed as either [e], (Jones 2011: vii; Wells 2008: xxiii), or [ɛ] (Kretzschmar 2008: 44; Schmitt 2016: 38; Zsiga 2013: 430). While Jones (2011: vii) and Wells (2008: xxiii) classify DRESS as a mid front vowel, the transcription choice of [ɛ] suggests an open-mid front vowel instead. Thus, DRESS ranges in terms of vowel height from mid to open-mid in General American. TRAP is a front vowel that is situated between fully open and open-mid and is realised as [æ] (Jones 2011: vii; Wells 1982a: 129; 2008: xxiii). BATH is realised with the TRAP vowel in General American and is thus also realised as [æ] (Tottie 2002: 17; Wells 1982a: 133-134). BATH can be categorised as an open front vowel in General American.

The vowels NURSE and STRUT are central monophthongs in General American. NURSE is a mid central vowel (Wells 2008: xxiv) that is rhotacized (Ladefoged and Johnson 2011: 94). NURSE is usually transcribed as [ɝ] (Jones 2011: ix; Wells 2008: xxiv; Zsiga 2013: 430). STRUT is an open-mid central vowel (Wells 2008: xxiii) that is traditionally transcribed as [ʌ] (Kretzschmar 2008: 44; Schmitt 2016: 38; Wells 2008: xxii; Zsiga 2013: 430). This transcription, however, does not completely fit an open-mid central vowel. Adhering to the IPA

(IPA 1999) principle of choosing the closest cardinal vowel for representation, [ɐ] might be a more appropriate symbol. However, [ʌ] is usually used to represent this vowel.

General American has nine back monophthongs: GOOSE, FOOT, FORCE, NORTH, THOUGHT, CLOTH, LOT, PALM, and START. Before these vowels are described, the lexical sets FORCE, NORTH, CLOTH, and LOT require some explanation. Some speakers of General American retain an opposition between the lexical sets FORCE and NORTH. For these speakers, FORCE is realised with [or], whereas NORTH is realised with [ɔr]. This leads to minimal pairs such as *for* (NORTH) and *four* (FORCE) or *born* (NORTH) and *borne* (FORCE). For many speakers, these two lexical sets have merged and are realised as [ɔr] (Wells 1982a: 159-162). In General American, CLOTH and LOT are two distinct lexical sets, whereas in Received Pronunciation these two vowels have merged. CLOTH includes those words that have LOT in Received Pronunciation with a succeeding fricative, /r/, or / ŋ/ (Collins and Mees 2006: 138; Wells 1982a: 130; 136-137).

GOOSE is a close back vowel in General American commonly realised as [u] (Wells 1982a: 147; 2008: xxiv). FOOT is a fairly close back vowel [ʊ] that is situated between fully back and central, and fully close and close-mid (Wells 1982a: 133; 2008: xxiii). If the split between FORCE and NORTH is retained, FORCE is a close-mid back vowel realised as [or] (Wells 1982a: 146; 161; Zsiga 2013: 430) NORTH is an open-mid back to open back vowel realised as [ɔr] (Wells 1982a: 145; 159; Zsiga 2013: 430) THOUGHT is also an open-mid back to open back vowel realised as [ɔ] (Wells 1982a: 145; 2008: xxiv; Zsiga 2013: 430). CLOTH is an open-mid back to open back vowel and is phonetically identical with THOUGHT. CLOTH is realised as [ɔ] (Wells 1982a: 136; Zsiga 2013: 430). LOT is "a central fully open unrounded vocoid, ranging from (retracted) [a] to advanced [ɑ]" (Wells 1982a: 130). Wells (2008: xxiv) and Jones (2011: vii; ix), however describe LOT as an open back vowel situated between fully back and central. They transcribe LOT as [ɑ][2]. Both articulatory descriptions of LOT are used in this study. PALM and LOT are realised with the same vowel in General American (Wells 1982a: 143). Thus, PALM is usually transcribed as [ɑ] (Jones 2011: ix). START is also phonetically identical to LOT and occurs in words with a succeeding /r/ (Wells 1982a: 158).

FACE, PRICE, MOUTH, GOAT, and CHOICE are closing diphthongs—sometimes also referred to as rising diphthongs—in General American. The diphthongs FACE, PRICE, and CHOICE move towards KIT, while MOUTH and GOAT move towards FOOT (Wells 2008: xxiii-xxiv). FACE has a front starting point between close-mid and mid (Wells 2008: xxii) and is

[2] Wells (2008) and Jones (2011) use [ɑː] in their transcription. The length marks were omitted here to be consistent with the transcription convention used in this chapter.

transcribed as [eɪ] (Tottie 2002: 18). PRICE has an open central or slightly fronted starting point and is transcribed as [aɪ] (Wells 1982a: 149). CHOICE has a mid back to open-mid back starting position (Wells 2008: xxiii) and is transcribed as [ɔɪ] (Jones 2011: vii; Upton 2008: 246; Wells 2008: xxiii). MOUTH has an open central starting point (Wells 1982a: 151) and is transcribed as [aʊ] (Wells 2008: xxiv). GOAT starts between mid central and mid back and can be transcribed as [oʊ] (Wells 2008: xxiv).

General American lacks phonemic centring diphthongs NEAR, SQUARE, and CURE. Instead, these lexical sets are pronounced with a monophthong and /r/ (Tottie 2002: 18). NEAR is realised [ɪr], SQUARE is realised as [er], and CURE is realised as [ʊr] (Jones 2011: ix).

Two select consonant features are introduced here: the state of rhoticity, and /t/. General American is a rhotic variety of English, so post-vocalic /r/ is articulated in all phonological contexts (Tottie 2002: 16; see also Section 2.4). General American exhibits 'medial-t-voicing' in intervocalic position before an unstressed vowel (Schmitt 2016: 39). [ɾ] is used in these instances (Wells 2008: xxv).

This thesis analyses the three monophthongs BATH, LOT, and CLOTH and rhoticity in the lexical sets NEAR, SQUARE, and CURE to investigate the representation of General American in German textbook audio materials. Section 2.8 provides a summary of the features under discussion in this thesis.

2.7 Australian English

This section defines Australian English and describes the accent used by the majority of Australian English speakers. First, different varieties within Australian English and their respective speech community are described. Standard (Mainstream) Australian English is given particular focus, as it is the variety used as this study's theoretical framework. A phonological description of Standard (Mainstream) Australian English is provided.

Australian English can be subdivided into three main dialect types: Australian Aboriginal Englishes, spoken by many Indigenous Australians (Butcher 2008), Ethnocultural Australian English, used to express non-mainstream or ethnic Australian identities, and Standard (Mainstream) Australian English (henceforth Australian English), spoken by non-indigenous people who were born in Australia or immigrated at an early age (Cox and Palethorpe 2007: 341). Australian English is the variety spoken by most Australians and is "a salient marker of national identity" (Cox and Palethorpe 2007: 341). It is the variety codified in dictionaries and used in education, government institutions, the courts, broadcasting, and

trade. Australian English is "characterised by specific vowel pronunciations, intonation patterns, lexical items, and various paralinguistic features which distinguish it from other types of English" (Harrington, Cox, and Evans 1997: 155).

The Australian English accent shows variation along a three-point socio-stylistic continuum from Broad Australian English at one end to Cultivated Australian English at the other end. General Australian English falls between Broad and Cultivated Australian English (Horvath 2008: 89). These accents are, however, not "discrete entities as they display considerable phonetic overlap" (Harrington, Cox, and Evans 1997: 156). Broad Australian English is "the most overtly local [and marked] form of Australian English" (Cox and Palethorpe 2007: 341), spoken by roughly a third of Australian English speakers (Horvath 2008: 89). Even though this accent of Australian English has historically been highly stigmatized (Harrington, Cox, and Evans 1997: 156), acceptance is growing (Bradley and Bradley 2001: 275). At the other end of the spectrum, Cultivated Australian English has the least local features and "bears some resemblance to Received Pronunciation" (Cox and Palethorpe 2007: 341). This accent is "estimated to be spoken by only about 10% of Australians" (Horvath 2008: 89). However, over the past decades, most speakers of Australian English can be located somewhere in the middle of this continuum, favouring General Australian English (Cox and Palethorpe 2007: 341). Most speakers of Australian English use General Australian English (Cox 2006: 148; Horvath 2008: 89-90), which therefore forms the theoretical basis for the following research. Phonetic differences to Broad or Cultivated Australian English, are included to provide a more extensive account of the phonological features of Australian English.

Traditionally, Australian English has been described as a relatively homogeneous variety with little regional variation (Cox and Fletcher 2017: 19). Even though first regional studies on the Australian English accent in the areas of New South Wales (cf. Cox 2006) and South Australia (cf. Butcher 2006) found some regional differences, they do not seem to be systematic and cannot be used to identify a speaker's regional origin. Thus, "[b]y global standards, [Australian English] displays relative homogeneity" (Cox and Palethorpe 2007: 342).

The focus in this section lies on the phonological features of Australian English and particularly on vowel realisations, as they carry "much of the responsibility for differentiating [Australian English] from other world Englishes" (Cox and Palethorpe 2012: 297). This section follows the transcription system used by Harrington, Cox, and Evans (1997), which follows "the IPA principle that symbols corresponding to the closest cardinal vowel should

be selected to represent phonemes" (Cox 2008: 329). The robustness of this transcription system to describe Australian English pronunciation has been confirmed by Butcher (2006) and Cox (2006) in two separate acoustic studies. They have shown that this phoneme set accurately reflects Australian English pronunciation. A description of the phonological features of Australian English with a strong focus on vowels follows.

The vowels FLEECE, KIT, DRESS, and TRAP are front vowels in Australian English (Cox 2006: 150). FLEECE and KIT are both front close vowels. FLEECE is generally realised as [iː], but often contains an onglide, giving it a slight diphthongal quality of [ᵊiː] (Cox and Fletcher 2017: 65), which, however, is often more pronounced among speakers of Broad Australian English (Cox and Palethorpe 2007: 345). KIT is a close front vowel in Australian English that is phonetically realised as [ɪ] (Cox and Fletcher 2017: 65). DRESS is traditionally raised in Australian English to a close-mid front position and is phonetically realised as [e] (Harrington, Cox, and Evans 1997: 164; 178; Watson, Harrington, and Evans 1998: 192-193). Cox and Palethorpe (2008: 345) have, however, shown that DRESS is starting to lower again in Australian English. Nevertheless, DRESS is still considered a close-mid front vowel in Australian English (Cox and Fletcher 2017: 65). TRAP, although traditionally also raised in Australian English, has considerably lowered in the past decades, resulting in an open front position for this vowel. TRAP is realised as [æ] (Cox and Fletcher 2017: 17; Cox and Palethorpe 2007: 345).

The vowels GOOSE, NURSE, STRUT, BATH, PALM, and START are considered central vowels in Australian English. GOOSE is considerably fronted in Australian English and can phonetically be described as [ʉː] (Cox 2006: 150). NURSE is quite fronted in Australian English and has a rounded quality for many speakers. NURSE is transcribed as [ɜː] (Cox 2006: 157), although this phoneme does not quite reflect the fronted and rounded quality of Australian English NURSE. The lexical sets BATH, PALM, and START are all realised with the vowel [ɐː]—an open central vowel (Cox 2006: 157; Cox and Fletcher 2017: 17). Since these three lexical sets all represent the same vowel, they are henceforth treated as one cluster BATH/PALM/START to refer to the vowel [ɐː]. If the individual lexical sets or words belonging to either lexical set are under discussion, however, each individual lexical set is referred to. STRUT can be transcribed as [ɐ], as it only contrasts to BATH/PALM/START in length.

FOOT, THOUGHT, NORTH, and LOT are considered back vowels in Australian English (Cox 2006: 150). FOOT is a close back vowel that can be transcribed as [ʊ]. THOUGHT and NORTH, both close back vowels—although slightly more open than FOOT—do not differ

phonetically and are both realised with the vowel [oː] (Cox 2006: 157). The combined cluster of these vowels will be referred to as THOUGHT/NORTH in this study.

The diphthongs, in particular, differentiate Australian English from other varieties of English (Cox 2008: 329). Australian English has five closing diphthongs: FACE, PRICE, MOUTH, CHOICE, and GOAT. FACE has an open front first target in the vicinity of TRAP and a closing glide that moves toward KIT. The transcription convention for this diphthong is [æɪ] (Cox 2006: 185; Cox 2008: 331; Harrington, Cox, and Evans: 1997: 171). The first target in this diphthong is slightly retracted for male speakers, with a starting point between TRAP and START (Cox 2006: 158). FACE also displays accent differences regarding the socio-stylistic continuum: speakers of Cultivated Australian English display a more fronted variant compared to General and Broad Australian English (Harrington, Cox, and Evans 1997: 174).

PRICE is a front rising diphthong that has an open back first target starting between START and LOT and a closing glide that moves toward DRESS. This diphthong is phonetically represented by [ɑe] (Cox 2008: 331). Speakers of Broad Australian English have a more retracted and raised first target in this diphthong compared to speakers of General or Cultivated Australian English (Harrington, Cox, and Evans 1997: 171-173). The first target in Broad Australian English can be categorised as open-mid back (Cox and Palethorpe 2010: 176) that is situated close to LOT (Cox 1998: 40). The first target in PRICE is more open and fronted in Cultivated compared to General Australian English. Thus, the first target in PRICE has an open central position in Cultivated Australian English (Cox 1998: 40-41). Overall, the main accent difference "between broad/general/cultivated accents is in the diphthongs [PRICE and MOUTH]" (Harrington, Cox, and Evans 1997: 179).

MOUTH has an open front first target between TRAP and START and a back-rising glide moving toward LOT. This vowel is transcribed as [æɔ] (Cox 2006: 159). Speakers of Broad Australian English show a raised first target in this diphthong compared to speakers of General or Cultivated Australian English (Harrington, Cox, and Evans 1997: 171). Broad talkers have a rather open-mid front first target in MOUTH between TRAP and DRESS (Harrington, Cox, and Evans 1997: 179). Speakers of Cultivated Australian English have an open front to open central first target in MOUTH that is slightly fronted compared to START (Cox 1998: 40-41).

Traditionally, Australian English, like other non-rhotic varieties of English, has three centring diphthongs: NEAR, SQUARE, and CURE. The vowels NEAR and SQUARE display a range from fully diphthongal realisations to more monophthongal realisations in Australian English (Harrington, Cox and Evans 1997: 175), which seem to become increasingly

monophthongal (Cox and Palethorpe 2012: 297). In these cases, NEAR and SQUARE only differ from KIT and DRESS in terms of length and are realised as long monophthongs (Cox and Fletcher 2017: 65-66; Cox and Palethorpe 2012: 297). While NEAR is still transcribed as a centring diphthong [ɪə], SQUARE is increasingly transcribed as [eː] (Cox and Fletcher 2017: 65; 67). CURE is only found as centring diphthong in the speech of older speakers of Australian English (Cox and Palethorpe 2007: 345). Among younger speakers of Australian English, CURE is either pronounced monophthongally with [oː] or disyllabically as [ʉːə] (Cox and Fletcher 2017: 66; Cox and Palethorpe 2007: 345).

Consonant features of Australian English have not been a major research focus in studies on Australian English as they "display many of the same variations present in other major dialects of English" (Cox and Palethorpe 2007: 342). For the sake of completeness, three select consonant features of Australian English are briefly introduced: the state of rhoticity, /l/, and the flapping of /t/. Australian English is a non-rhotic variety of English that has both linking and intrusive /r/ (Trudgill and Hannah 2008: 24). Dark *l* [ɫ] occurs in Australian English "in pre-pausal and pre-consonantal positions and (…) often (…) before a morpheme boundary preceding a vowel" (Cox and Palethorpe 2007: 343) and is often vocalised in Australian English depending on the phonetic environment (cf. Borowsky 2001; Horvath and Horvath 2002). In intervocalic position, /t/ can be realised as a voiced flap [ɾ] both in intervocalic final positions (*lot of*) and in word-medial positions (*bitter*) (Tollfree 2001: 57).

This thesis analyses the three rising diphthongs, FACE, PRICE, and MOUTH to investigate the representation of Australian English in German textbook audio materials. As discussed above, these vowels are salient features of Australian English differentiating this accent from other accents of English, and account for many of the differences between Broad, General, and Cultivated Australian English. The analysis focuses on the first targets in the diphthongs FACE, PRICE, and MOUTH, as "the second target of diphthongs is much more variable than the first and often not attained [...] even in citation-form speech" (Harrington, Cox, and Evans 1997: 174). Section 2.8 provides a summary of the features under discussion in this thesis.

2.8 Phonological Features Chosen for Analysis

This section provides a summary of the features chosen for the analysis for Received Pronunciation, General American, and Australian English. The vowels chosen for the analysis

are described according to their articulatory features of vowel height and tongue advancement and are visualised in stylised vowel quadrilaterals for each accent. Vowels with various realisations in an accentare visualised with all possible variants. The section on General American also includes a summary regarding rhoticity in the lexical sets NEAR, SQUARE, and CURE.

Received Pronunciation

The monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START are analysed for Received Pronunciation in this study. Table 2 summarises the articulatory features of the monophthongs under discussion as presented in Section 2.5. The vowels are described according to vowel height (close, close-mid, mid, open-mid, and open) and tongue advancement (front, central, back). This categorisation obscures some variety present within these categories: LOT and START, for instance, are both classified as open back vowels, but LOT is slightly further back and not quite as open as START, for example. Such a simplification in categorisation is, however, necessary for the analyses presented in this study. This categorisation of the monophthongs regarding vowel height and tongue advancement forms the basis of the comparison (see Section 3.3.3) between the data analysed in this study and the linguistic description of Received Pronunciation.

Lexical Set	Articulatory description
GOOSE	close back or close central to close front
LOT	open back
CLOTH	open back
TRAP	open front
BATH	open back or open central
START	open back

Table 2. Articulatory categorisation of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START in Received Pronunciation.

The vowel quality of the monophthongs under discussion is illustrated in the stylised vowel quadrilaterals in Figure 1 and Figure 2 below. These representations are based on the linguistic descriptions provided in Section 2.5 above. The variability present in the articulatory categories of vowel height and tongue advancement is included in these vowel quadrilaterals. As BATH and GOOSE have more than one realisation in Received Pronunciation, these

realisations are represented in different vowel plots. There is no correlation between the two pronunciation variants. So, speakers with a fronted GOOSE vowel don't necessarily have a fronted BATH vowel as well. This is done purely for a better overview.

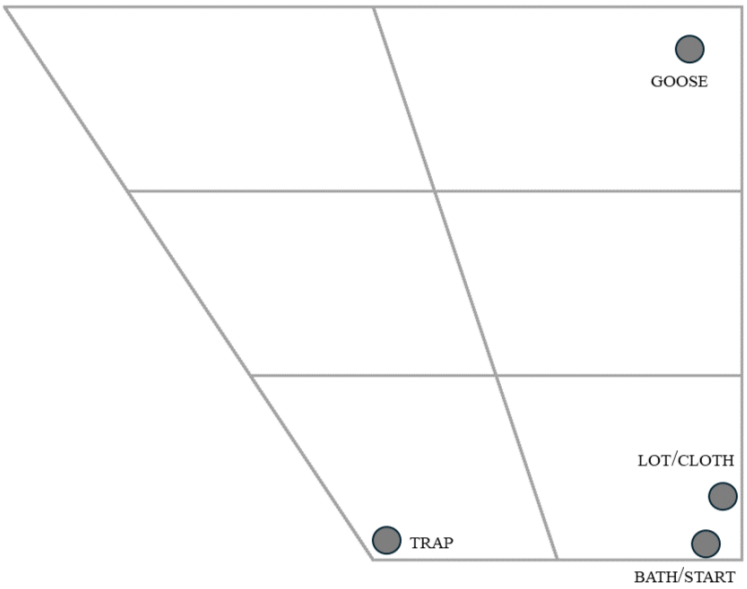


Figure 1. Received Pronunciation GOOSE, LOT, CLOTH, TRAP, BATH, START with back GOOSE and BATH.

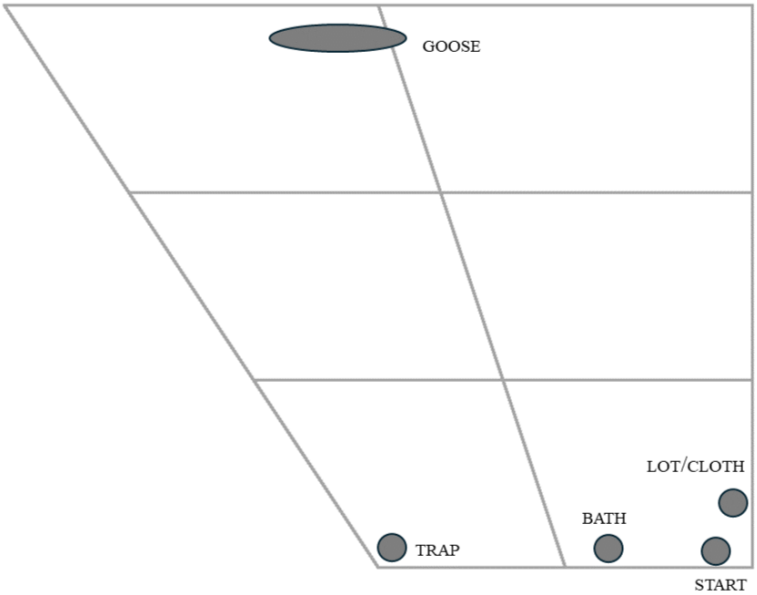


Figure 2. Received Pronunciation GOOSE, LOT, CLOTH, TRAP, BATH, START with fronted GOOSE and BATH.

General American

The monophthongs BATH, LOT, and CLOTH, and rhoticity in the lexical sets NEAR, SQUARE, and CURE are analysed for General American in this study. Table 3 summarises the

articulatory features of the monophthongs under discussion as presented in Section 2.6. The vowels are described according to vowel height—close, close-mid, mid, open-mid, and open—and tongue advancement—front, central, back. This categorisation obscures some variety present within these categories: LOT, and CLOTH both have variants that are categorised as open back vowels, but LOT is more open and less back than CLOTH. Such a simplification in categorisation is, however, necessary for the analyses presented in this study. This categorisation of the monophthongs regarding vowel height and tongue advancement forms the basis of the comparison between the data analysed in this study and the linguistic description of General American.

Lexical Set	**Articulatory description**
BATH	open front
LOT	open central to open back
CLOTH	open-mid back to open back

Table 3. Articulatory categorisation of the monophthongs BATH, LOT, and CLOTH in General American.

The vowel quality of the monophthongs under discussion is illustrated in the stylised vowel quadrilateral in Figure 3 below. These representations are based on the linguistic descriptions provided in Section 2.6 above. The variability present in the articulatory categories of vowel height and tongue advancement is included in these vowel quadrilaterals.

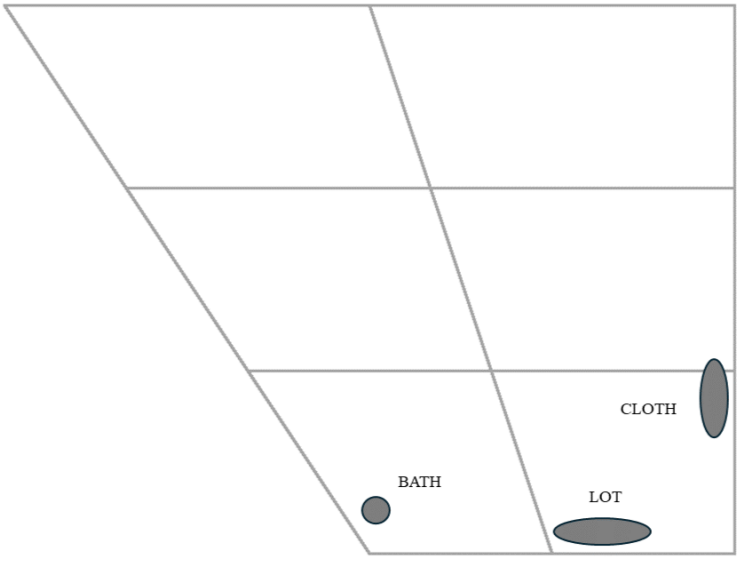


Figure 3. General American BATH, LOT, and CLOTH.

General American is a rhotic accent of English, meaning that post-vocalic /r/ is articulated in all phonological contexts. Thus, /r/ is always articulated in NEAR, SQUARE, and CURE.

Australian English

The first targets in the diphthongs FACE, PRICE, and MOUTH are analysed for Australian English in this study. Table 4 summarises the articulatory features of the first targets in the diphthongs FACE, PRICE, and MOUTH as presented in Section 2.7. The first targets in these diphthongs are described according to vowel height—close, close-mid, mid, open-mid, and open—and tongue advancement—front, central, back. This categorisation obscures some variety present within these categories: FACE, for example, has an open front first target in Cultivated and General Australian English, but the first target in Cultivated is fronted compared to the first target of FACE in General Australian English. Such a simplification in categorisation is, however, necessary for the analyses presented in this study. This categorisation of the first targets in the diphthongs regarding vowel height and tongue advancement forms the basis of the comparison between the data analysed in this study and the linguistic description of Australian English.

Lexical Set	**Articulatory description**		
	Cultivated	**General**	**Broad**
FACE	open front	open front	open front
PRICE	open central	open back	open-mid back
MOUTH	open front to open central	open front	open-mid front

Table 4. Articulatory categorisation of the first targets in FACE, PRICE, and MOUTH in Australian English.

The vowel quality of the diphthongs under discussion is illustrated in the stylised vowel quadrilaterals in Figure 4 – Figure 6 below. Cultivated Australian English is represented in Figure 4, General Australian English is represented in Figure 5, and Broad Australian English is represented in Figure 6. These representations are based on the linguistic descriptions provided in Section 2.7 above. The variability present in the articulatory categories of vowel height and tongue advancement is included in these vowel quadrilaterals. These vowel quadrilaterals also include select Australian English monophthongs to illustrate the position of the diphthongs further. As most of the accent differences between Cultivated,

General, and Broad Australian is found in the diphthongs, the monophthongs included do not display any accent differences. The diphthongs FACE, PRICE, and MOUTH are illustrated as stylised trajectories from the first to the second vowel target.

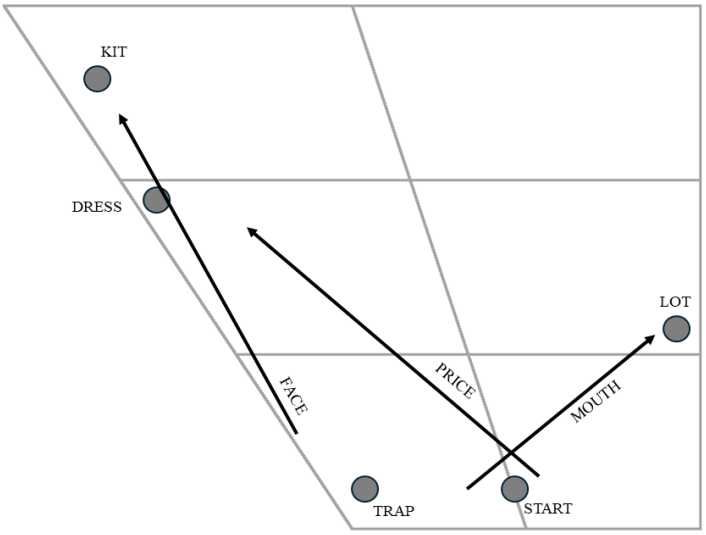


Figure 4. FACE, PRICE, and MOUTH in Cultivated Australian English.

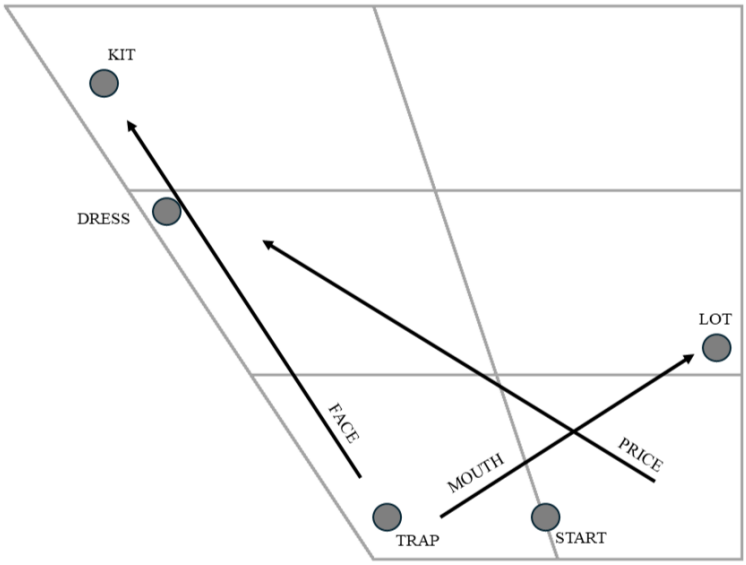


Figure 5. FACE, PRICE, and MOUTH in General Australian English.

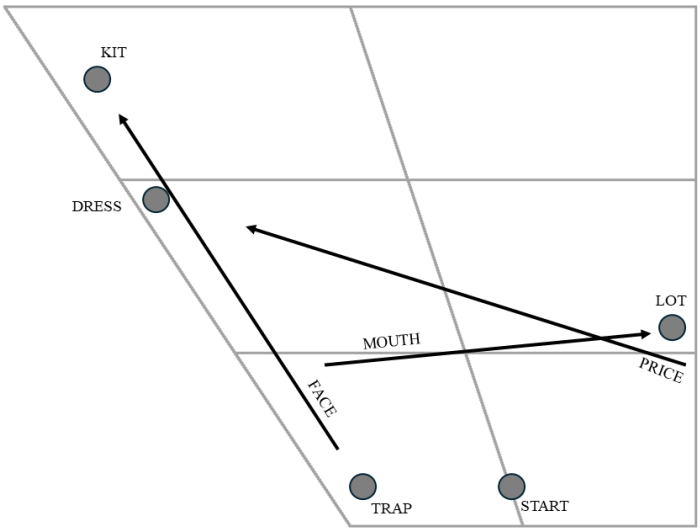


Figure 6. FACE, PRICE, and MOUTH in Broad Australian English.

This section summarised the features chosen for the analysis of Received Pronunciation, General American, and Australian English and provided illustrations of the articulatory features of the vowels through stylised vowel quadrilaterals.

2.9 Summary

This chapter presented the study's theoretical framework. The concepts of vowel quality and rhoticity were presented. The concept of vowel quality was defined from three different phonetic perspectives: articulatory, auditory, and acoustic phonetics and an introduction into analysing vowel quality was provided. Rhoticity was presented within the framework of Harris' (2013) rhoticity systems and common procedures of analysing rhoticity were presented. The accents included in this study—Received Pronunciation, General American, and Australian English—were each defined, and the vowels and salient consonantal features of each accent were described. Section 2.8 concluded with a summary of the features chosen for analysis. Chapter 3 presents the Methodology of this study.

3 Methodology

3.1 Introduction

Expanding on the theoretical foundations presented in Chapter 2, this chapter outlines the specifics of the data and methodology employed in this study. Section 3.2 discusses the corpus design of the specialized phonological audio textbook corpus. This is followed by an examination of practical considerations that influenced the corpus design, leading to a presentation of the general corpus design and annotation practices. Following this, the corpus design of the individual varietal sub-corpora is presented, providing an overview of the textbooks included, as well as the speakers and audio tracks that were analysed.

Section 3.3 presents the methodology of data analysis, including the methodology of the vowel analysis and the rhoticity analysis. Additionally, the section includes a detailed methodological description of the comparison between the findings of the vowel and rhoticity analyses conducted in the individual sub-corpora, and accent descriptions and findings from other acoustic studies on the respective accents under discussion. The primary objective of this chapter is to establish the framework for analysing the data presented in Chapter 4 and to outline the limitations related to the study's data and methodology.

3.2 Corpus Design: The Specialized Phonological Audio Textbook Corpus

The Specialized Phonological Audio Textbook Corpus (henceforth: the phonological textbook corpus or just the corpus) is—as the name suggests—a specialized phonological corpus comprised of audio textbook materials. Gut and Voorman (2014: 16) are the first researchers to provide a definition of a phonological corpus:

> A phonological corpus is thus defined here as a representative sample of language that contains
>
> - primary data in the form of audio or video data;
> - phonological annotations that refer to the raw data by time information (time-alignment); and
> - metadata about the recordings, speakers and corpus as a whole.

The corpus presented in this study is comprised of audio data from three textbook series' used in ELT in secondary schools in NRW, Germany: *English G Access, Green Line,*

and *Camden Town.* The textbooks used in grades 6, 8, and 9 are included in this corpus. The focus of this corpus is to display the representation of Received Pronunciation, General American, and Australian English in these textbooks. The following sections provide an overview of the corpus design and annotation practices employed in the compilation of this corpus.

3.2.1 Practical Considerations

Several key considerations had to be addressed before compiling the phonological textbook corpus. First, access to and availability of the ELT textbooks, along with their accompanying audio material and orthographic transcripts, had to be ensured. The audio material was available in digital format on CDs, while the transcripts of the audio material were available in electronic PDF files or in print teacher manuals or in the textbooks themselves. The audio files were saved in the Waveform Audio format, and the transcripts were digitized—if necessary—and saved as PDF files.

Given the constraints of time and resources inherent in a PhD project, the specialized phonological textbook corpus (referred to hereafter as 'the corpus') had to be designed accordingly. Consequently, this corpus focuses specifically on textbooks used in ELT at secondary schools (*Gymnasium*) in North Rhine-Westphalia. Moreover, the corpus was limited to three accents of English to ensure a manageable scope and allow for analytical depth. The accents included are Australian English, General American, and Received Pronunciation.

The following section provides a detailed account of the specific corpus design for the corpus as a whole and the individual sub-corpora dedicated to each accent.

3.2.2 Corpus Design and Annotation

This phonological textbook corpus is compiled of audio material from the newest editions (at the time of data collection in 2017/2018) of the volumes two, four, and five of the three textbook series *English G Access, Green Line,* and *Camden Town.* These textbooks are used in years six (volume 2), eight (volume 4), and nine (volume 5) in North Rhine-Westphalia. The accents included in this corpus are Received Pronunciation, General American, and Australian English.

Table 5 below provides an overview of the size of the corpus and the size of the individual sub-corpora. The entire corpus consists of three hours, 36 minutes, and 40 seconds

of recorded speech from the textbook audio materials. The Australian English sub-corpus is the smallest sub-corpus with 36 minutes of recorded speech included, while the General American sub-corpus consists of one hour, 45 minutes, and 28 seconds of recorded speech. The difference in size for the three sub-corpora is due to the fact that the Australian English sub-corpus is only comprised of two textbooks, while the General American sub-corpus, as well as the Received Pronunciation sub-corpus consist of audio materials from three textbooks each. Detailed descriptions of the corpus design of each sub-corpus are provided in Section 3.2.3 for the Received Pronunciation sub-corpus, 3.2.4 for the General American sub-corpus, and 3.2.5 for the Australian English sub-corpus.

Sub-corpora	Length of Audio Materials
Received Pronunciation sub-corpus	01:15:12
General American sub-corpus	01:45:28
Australian English sub-corpus	00:36:00
Total	**03:36:40**

Table 5. Size of the corpus and the individual sub-corpora.

In general, the corpus includes seven speakers per textbook and an analysis of one to two audio tracks per speaker, in most instances. These speakers per textbook include all available speakers for the Australian English sub-corpus and a selection of speakers for the Received Pronunciation and General American sub-corpora, as more speakers were available in the audio materials for the Received Pronunciation and General American sub-corpora. However, to keep the corpus sizes comparable, only a selection of seven speakers[3] and one to two audio tracks per speaker were selected per textbook. An overview of the chosen speakers and audio tracks can be found in sections 3.2.3 for the Received Pronunciation sub-corpus, 3.2.4 for the General American sub-corpus, and 3.2.5 for the Australian English sub-corpus.

The speakers were chosen on the grounds that they were either explicitly described as a speaker of the respective accent under discussion, or their status as a speaker of that accent could be inferred from the context of the audio text.[4] The rationale for this approach

[3] For some textbooks from the General American and Australian English sub-corpora, only six speakers were selected. The choice of this selection is explained further in the respective sub-chapters below.

[4] The textbook publishers were contacted about information on the voice actors and their native accents as well as which characters each voice actor voiced. However, the publishers did not offer any information on the voice actors used for the recordings.

is that the aim of this thesis is to analyse how different accents are represented in the audio materials Therefore, (supposed) speakers of these accents are analysed to evaluate how well they portray the respective accents. In all analysed audio tracks all occurrences of the features under discussion and usually up to five instances of the remaining monophthongs were analysed for each speaker.

Sub-corpora	Speakers	Tokens
Received Pronunciation sub-corpus	21	1.633
General American sub-corpus	20	1.732
Australian English sub-corpus	13	1.365
Total	**54**	**4.730**

Table 6. Number of speakers and tokens per sub-corpus.

Table 6 provides an overview of the number of speakers included in the individual sub-corpora as well as the entire corpus, as well as the number of tokens analysed per sub-corpus and in the entire corpus. Overall, the corpus includes 54 speakers and 4.730 analysed tokens. These token numbers include all features that were part of the analysis, and the additional monophthongs analysed to create each speaker's individual vowel spaces (see Section 3.3.1).

Any phonological corpus requires at least a time-aligned annotation on an orthographic level, and level of phonological annotations (Gut and Voormann 2014: 17). This corpus was annotated in Praat on three different tiers: the speaker tier, the word tier, and the vowel tier. All relevant words for the analysis were annotated orthographically in a time-aligned manner on the word tier. For each word, an annotation was added on the speaker tier to indicate which speaker uttered these words. Within the word, the relevant vowels were annotated in a time-aligned manner on the vowel tier using the lexical sets as labels for the vowels. While many phonological corpora use phonetic symbols at a phonological level of annotation (cf. Delais-Roussarie and Post 2014: 58), Well's LEXICAL SETS were chosen for this level of annotation in this corpus. As manual annotations, especially phonological annotations, are very time-consuming (Gut and Voormann 2014: 24), a "query-driven annotation approach" (Brinckmann 2014: 364) was chosen for this corpus: Thus, only the words and vowels of interest for this research project were annotated.

The speakers were coded according to the textbook, track number, and occurrence of the speaker on the track. For example, speaker SpEGA4-2-1 is a speaker from the textbook

English G Access 4, track number two, and the first speaker on this audio track. If a textbook is accompanied by several CDs, the CD number is added between the textbook code and the track number. In this case, SpGL5-1-3-1 is a speaker from the textbook Green Line 5, CD1, track number three, and the first speaker on this audio track. Using this notation, each speaker can be identified easily in the annotation.

As two or more audio tracks per speaker are usually included in this corpus, each speaker received a new speaker code to show which characters were voiced by the same speaker. The final speaker codes were coded for the speaker's gender—M for male speakers and F for female speakers—and the textbook and volume of the textbook—E for *English G Access*, G for *Green Line*, and C for *Camden Town*. The speakers were numbered consecutively by gender throughout the three sub-corpora starting with the Australian English sub-corpus, followed by the General American sub-corpus and lastly the Received Pronunciation sub-corpus. Therefore, speaker M1E5 is the first male speaker from the Australian English sub-corpus, as E5 stands for the textbook *English G Access 5*, which is part of the Australian English sub-corpus. Speaker F8G4 is a female speaker from the General American sub-corpus, as G4 stands for *Green Line 4*, which is part of the General American sub-corpus. Male and female speakers were each coded separately but consecutively, meaning that the male speakers are counted consecutively from 1 through to 30 starting with the textbook *English G Access* in each sup-corpus, followed by the textbook *Green Line*, and then, if applicable, *Camden Town*.

Sub-corpus	Speakers
Australian English	F1E5, F2G5, F3G5, F4G5, M1E5, M2E5, M3E5, M4E5, M5E5, M6G5, M7G5, M8G5, M9G5
General American	F5E4, F6E4, F7E4, F8G4, F9G4, F10G4, F11C4, F12C4, F13C4, F14C4, M10E4, M11E4, M12E4, M13E4, M14G4, M15G4, M16G4, M17G4, M18C4, M19C4
Received Pronunciation	F15E2, F16E2, F17G2, F18G2, F19G2, F20C2, F21C2, F22C2, F23C2, F24C2, M20E2, M21E2, M22E2, M23E2, M24E2, M25G2, M26G2, M27G2, M28G2, M29C2, M30C2

Table 7. Overview of the speakers per sub-corpus.

Table 7 displays an overview of the speaker codes per sub-corpus. A list of which final speaker code encompasses which original speaker codes (i.e. the speaker codes

showing the track number and occurrence on the track) are included in Sections 3.2.3-3.2.5, where each sub-corpus is discussed separately.

When designing a new corpus, the question of representativeness is always at the forefront. According to Gut and Voormann (2014: 20):

> The term 'representativeness' is usually used to refer to the objective that the raw data of a corpus should constitute a sample of a (…) language variety that includes its full range of variability. It should thus provide the researchers with as accurate as possible a picture of the occurrence and variation of linguistic phenomena, and the potential to generalize the corpus-based findings to a language or language variety as a whole.

This corpus provides a relatively high level of representativeness with regards to the representation of Received Pronunciation, General American, and Australian English in the textbook series English G Access, Green Line and Camden Town, as not only a wide range of speakers are included in this corpus, but also all instances of the variables under discussion for each speaker are included and analysed. Thus, this corpus provides the opportunity of generalizing the findings concerning the three accents of English—Received Pronunciation, General American, and Australian English—to the variety of 'textbook English' as a whole for these textbooks used in North Rhine-Westphalia.

3.2.3 The Received Pronunciation Sub-Corpus

This section describes the corpus design of the Received Pronunciation sub-corpus with regards to the textbooks used, the speakers and audio tracks included, and the linguistic features included in this sub-corpus.

The Received Pronunciation sub-corpus is comprised of the three textbooks *English G Access 2* (Rademacher 2014a), *Green Line 2* (Weisshaar 2015a), and *Camden Town 2* (Hanus et al. 2013a), which are used in the 6th grade at secondary schools (*Gymnasium*) in North Rhine-Westphalia. Twenty-one speakers in total—seven speakers from each textbook—are included in this sub-corpus. A total of one hour, 15 minutes, and 12 seconds of audio data from 36 different audio tracks—ten from *English G Access 2* (Rademacher 2014b), 14 from *Green Line 2* (Weisshaar 2015b), and 12 from *Camden Town 2* (Hanus et al. 2013b)—are included. The sub-corpus includes 1.633[5] tokens in total.

[5] Originally, the lexical set SQUARE was included in this sub-corpus. These tokens are still included in the total number of tokens for this sub-corpus, even though they were later no longer considered for the analysis.

The 21 speakers of this sub-corpus are comprised of ten female and 11 male speakers. Seven speakers—two female and five male speakers—are included from the textbook *English G Access 2*. Seven speakers—three female and four male speakers—are included from the textbook *Green Line 2*. Seven speakers—five female and two male speakers—are included from *Camden Town 2*. For all three textbooks, more than seven speakers were available. The seven speakers per textbook were chosen on the grounds that each speaker had a relatively high speaking portion within the textbook to ensure the availability of a substantial number of tokens for the analysis.

Table 8 provides an overview of the speakers and indicates the original, track-individual speaker codes, as well as the audio tracks analysed for each speaker. Two audio tracks per speaker are included in this sub-corpus, except for two speakers: M23E2 and F22E2. For speaker M23E2 one track was available. Three audio tracks were included for speaker F22C2. As these textbook audio materials are on several CDs, the track numbers include the CD number. For instance, audio track 1-39 from the textbook *English G Access 2* is track number 39 on CD 1.

The main features of interest in this sub-corpus are the vowels GOOSE, LOT, CLOTH, TRAP, BATH, and START. Table 9 displays the raw frequency of the features under discussion per speaker, as well as the total frequency per vowel.

Of the six vowels under discussion, TRAP occurs most frequently with 230 tokens in total, followed by LOT with 194 tokens. CLOTH appears the least with only 58 tokens. The distribution of these six vowels across the individual speakers varies. The lowest frequency of these vowels across speakers ranges from zero to two. The highest frequency per vowel and speaker ranges from seven tokens for CLOTH to 31 tokens for TRAP.

Textbook	Speaker	Original Speaker Code	Audio Tracks
English G Access 2	F15E2	SpEGA2-1-39-2, SpEGA2-1-44-2	1-39, 1-44
	F16E2	SpEGA2-2-24-2, SpEGA2-3-1-1	2-24, 3-1
	M20E2	SpEGA2-1-9-1, SpEGA2-1-44-3	1-9, 1-44
	M21E2	SpEGA2-2-23-3, SpEGA2-2-24-1	2-23, 2-24
	M22E2	SpEGA2-3-1-2, SpEGA2-3-10-1	3-1, 3-10
	M23E2	SpEGA2-3-26-1	3-26
	M24E2	SpEGA2-3-2-1, SpEGA2-3-28-3	3-2, 3-28
Green Line 2	F17G2	SpGL2-2-12-2, SpGL2-2-13-2	2-12, 2-13
	F18G2	SpGL2-1-14-2, SpGL2-2-17-2	1-14, 2-17
	F19G2	SpGL2-2-33-2, SpGL2-2-34-2	2-33, 2-34
	M25G2	SpGL2-2-18-1, SpGL2-2-19-1	2-18, 2-19
	M26G2	SpGL2-1-9-2, SpGL2-1-10-2	1-9, 1-10
	M27G2	SpGL2-3-2-1, SpGL2-3-3-1	3-2, 3-3
	M28G2	SpGL2-1-6-1, SpGL2-1-20-2	1-6, 1-20
Camden Town 2	F20C2	SpCT2-2-20-1, SpCT2-2-28-2	2-20, 2-28
	F21C2	SpCT2-2-7-2, SpCT2-2-34-2	2-7, 2-34
	F22C2	SpCT2-2-20-2, SpCT2-2-28-1, SpCT2-1-36-1	2-20, 2-28, 1-36
	F23C2	SpCT2-1-36-2, SpCT2-1-38-2	1-36, 1-38
	F24C2	SpCT2-2-36-1, SpCT2-2-30-1	2-36, 2-30
	M29C2	SpCT2-1-35-3, SpCT2-2-23-3	1-35, 2-23
	M30C2	SpCT2-1-16-2, SpCT2-2-14-2	1-16, 2-14

Table 8. Overview of the Received Pronunciation speakers and audio tracks per textbook.

Speaker	GOOSE	LOT	CLOTH	TRAP	BATH	START
F15E2	4	8	3	14	1	2
F16E2	4	2	0	4	2	0
M20E2	9	11	3	5	4	3
M21E2	4	4	7	9	2	14
M22E2	11	10	4	26	8	5
M23E2	12	11	3	10	15	6
M24E2	16	11	2	31	19	22
F17G2	5	13	5	9	1	1
F18G2	2	6	1	3	2	2
F19G2	5	8	0	5	3	0
M25G2	5	4	3	11	2	2
M26G2	10	10	1	8	3	3
M27G2	13	5	3	13	9	2
M28G2	18	22	5	18	12	10
F20C2	2	1	3	2	0	2
F21C2	3	10	1	2	5	3
F22C2	6	13	3	5	4	7
F23C2	4	9	2	5	3	1
F24C2	18	18	6	20	8	4
M29C2	6	5	1	10	1	0
M30C2	9	13	2	20	3	2
Total	**166**	**194**	**58**	**230**	**107**	**91**

Table 9. Token numbers per speaker for GOOSE, LOT, CLOTH, TRAP, BATH, and START.

Moreover, the corpus was tagged for the remaining stressed monophthongs to show the characteristics of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. in relation to their position in each speaker's vowel space (see Section 3.3.1). The monophthongs FLEECE, KIT, DRESS, NURSE, STRUT, PALM, FOOT, THOUGHT and NORTH were added to the corpus. In general, up to five tokens per monophthong were analysed per speaker, though occasionally more than five tokens were annotated. In these cases, all annotated tokens were analysed. Table 1 (Appendix 6) provides an overview of the token numbers analysed for each monophthong per speaker.

3.2.4 The General American Sub-Corpus

This section describes the corpus design of the General American sub-corpus with regards to the textbooks used, the speakers and audio tracks included, and the relevant linguistic features.

The General American sub-corpus is comprised of the three textbooks *English G Access 4* (Rademacher 2016a), *Green Line 4* (Weisshaar 2017a), and *Camden Town 4* (Claussen et al. 2015a), which are used in the 8th grade at secondary schools (*Gymnasium*) in North Rhine-Westphalia. 20 speakers in total—seven speakers from *English G Access 4*, seven speakers from *Green Line 4*, and six speakers from *Camden Town 4*—are included. A total of one hour, 45 minutes, and 28 seconds of audio data from 35 different audio tracks—12 from *English G Access 4* (Rademacher 2016b), 12 from *Green Line 4* (Weisshaar 2017b), and 11 from *Camden Town 4* (Claussen et al. 2015b)—are included. The sub-corpus includes 1.732 tokens for analysis.

The 20 speakers are comprised of ten female and ten male speakers. Seven speakers—three female and four male speakers—are included from the textbook *English G Access 4*. Seven speakers—three female and four male speakers—are included from the textbook *Green Line 4*. Six speakers—four female and two male speakers—are included from *Camden Town 4*. For all three textbooks, more than seven speakers were available. The number of speakers per textbook were chosen on the grounds that each speaker had a relatively high speaking portion within the textbook to ensure the availability of a substantial number of tokens for the analysis. Only six speakers were included for *Camden Town 4*, as no other speaker showed a substantial speaking portion throughout the audio tracks. Thus, only six speakers were included for this textbook.

Table 10 provides an overview of the speakers and indicates the original, track-individual speaker codes, as well as the audio tracks analysed for each speaker. Two audio tracks per speaker are included in this sub-corpus, except for two speakers—M10E4 and M12E4—for whom only one track was available. Three audio tracks were included for speaker F11C4, as the third track—track number 1-14—was also included for speaker F12C4. As these textbook audio materials are on several CDs, the track numbers include the CD number. For instance, audio track 1-27 from the textbook *English G Access 4* is track number 27 on CD 1.

Textbook	Speaker	Original Speaker Code	Audio Tracks
English G Access 4	F5E4	SpEGA4-1-27-1, SpEGA4-1-28-1	1-27, 1-28
	F6E4	SpEGA4-2-15-1, SpEGA4-2-18-1	2-15, 2-18
	F7E4	SpEGA4-1-51-1, SpEGA4-1-53-2	1-51, 1-53
	M10E4	SpEGA4-2-7-1	2-7
	M11E4	SpEGA4-1-7-1, SpEGA4-2-11-1	1-7, 2-11
	M12E4	SpEGA4-2-13-1	2-13
	M13E4	SpEGA4-2-16-1, SpEGA4-2-17-1	2-16, 2-17
Green Line 4	F8G4	SpGL4-1-17-2, SpGL4-2-6-1	1-17, 2-6
	F9G4	SpGL4-1-18-2, SpGL4-3-9-2	1-18, 3-9
	F10G4	SpGL4-2-5-1, SpGL4-3-2-1	2-5, 3-2
	M14G4	SpGL4-1-3-1, SpGL4-1-20-2	1-3, 1-20
	M15G4	SpGL4-2-15-1, SpGL4-3-4-1	2-15, 3-4
	M16G4	SpGL4-2-7-2, SpGL4-2-8-2	2-7, 2-8
	M17G4	SpGL4-1-20-1, SpGL4-2-5-2	1-20, 2-5
Camden Town 4	F11C4	SpCT4-1-5-2, SpCT4-1-9-1, SpCT4-14-3	1-5, 1-9, 1-14
	F12C4	SpCT4-1-14-2, SpCT4-2-18-1	1-14, 2-18
	F13C4	SpCT4-1-5-4, SpCT4-2-27-2	1-5, 2-27
	F14C4	SpCT4-1-4-1, SpCT4-1-24-1	1-5, 1-24
	M18C4	SpCT4-1-3-1, SpCT4-2-27-1	1-3, 2-27
	M19C4	SpCT4-2-6-1, SpCT4-2-29-2	2-6, 2-29

Table 10. Overview of the General American speakers and audio tracks per textbook.

The main features of interest in this sub-corpus are the vowels BATH, LOT, and CLOTH, as well as rhoticity in the lexical sets NEAR, SQUARE, and CURE. Table 11 displays the raw frequency of the features under discussion per speaker, as well as the total frequency per vowel and for rhoticity in the lexical sets NEAR, SQUARE, and CURE.

Speaker	BATH	LOT	CLOTH	Rhoticity
F5E4	7	11	4	11
F6E4	11	24	4	11
F7E4	5	20	6	8
M10E4	5	21	1	9
M11E4	16	21	3	12
M12E4	4	20	5	7
M13E4	7	10	5	6
F8G4	2	10	2	22
F9G4	1	10	1	8
F10G4	2	9	1	7
M14G4	13	48	6	30
M15G4	4	12	7	6
M16G4	0	6	1	2
M17G4	4	15	0	9
F11C4	4	15	2	9
F12C4	2	11	12	20
F13C4	3	9	4	3
F14C4	0	8	0	1
M18C4	6	24	7	17
M19C4	6	20	5	14
Total	**102**	**324**	**76**	**212**

Table 11. Token numbers per speaker for BATH, LOT, CLOTH, and rhoticity in the features NEAR, SQUARE, and CURE.

Of the three vowels under discussion, LOT occurs most frequently with 324 tokens in total. BATH occurs 102 times in the sub-corpus, and CLOTH only 76 times. The distribution of these vowels across the individual speakers varies. The raw frequency for BATH ranges from zero tokens for speaker M16G4 to 16 tokens for speaker M11E4. LOT occurs between six times for speaker M16G4 and 48 times for speaker M14G4. CLOTH is not present for speaker M17G4 but occurs 12 times for speaker F12C4. Two hundred and twelve tokens were analysed for rhoticity. The raw frequency for rhoticity tokens per speaker—represented by the lexical sets NEAR, SQUARE, and CURE— vary greatly. For speaker F14C4 only one token of rhoticity could be analysed, while 30 tokens were analysed for speaker M14G4.

Moreover, the corpus was tagged for the remaining stressed monophthongs to show the characteristics of the monophthongs BATH, LOT, and CLOTH in relation to their position in each speaker's vowel space (see Section 3.3.1). The monophthongs FLEECE, KIT, DRESS, TRAP, NURSE, STRUT, PALM, START, GOOSE, FOOT, THOUGHT, NORTH, and FORCE were added to the sub-corpus. In general, up to five tokens per monophthong were analysed per speaker, though occasionally more than five tokens were annotated. In these cases, all annotated tokens were analysed. Table 2 (Appendix 6) provides an overview of the token numbers analysed for each monophthong per speaker.

3.2.5 The Australian English Sub-Corpus

This section describes the corpus design of the Australian English sub-corpus including the textbooks used, the speakers and audio tracks included, and the relevant linguistic features. This sub-corpus design has been previously described in Scheiwe (2022).

The Australian English sub-corpus of the phonological textbook corpus is comprised of the two textbooks *English G Access 5* (Rademacher 2017a) and *Green Line 5* (Weisshaar 2018a), which are used in the 9th grade at secondary schools (*Gymnasium*) in North Rhine-Westphalia. Audio samples from a unit on Australia from both textbooks are part of this sub-corpus, including speech samples from 13 different speakers. This sub-corpus consists of 36 minutes of audio data from 13 different audio tracks—seven from *English G Access 5* (Rademacher 2017b) and six from *Green Line 5* (Weisshaar 2018b). The sub-corpus provided 1.365 tokens in total for analysis.

The 13 speakers are comprised of four female and nine male speakers. Six speakers—one female and five male speakers—are included from the textbook *English G Access 5*, and seven speakers—three female and four male speakers—are included from *Green Line 5*. All available speakers were analysed for the Australian English sub-corpus. For each speaker, one to two audio tracks were analysed. The remaining audio tracks within the Australian English textbook units that were not analysed for this study either included speakers for whom two audio tracks were already included in the corpus, or these audio tracks did not display an Australian setting.

Table 12 below provides an overview of the speakers in this sub-corpus and indicates the original, track-individual speaker codes, as well as the tracks analysed for each speaker.

Textbook	Speaker	Original Speaker Code	Audio Tracks
English G Access 5	F1E5	SpEGA5-3-2, SpEGA5-5-4	3, 5
	M1E5	SpEGA5-2-1	2
	M2E5	SpEGA5-3-1, SpEGA5-5-2	3, 5
	M3E5	SpEGA5-3-3, SpEGA5-4-2, SpEGA5-5-5	3, 4, 5
	M4E5	SpEGA5-6-1	6
	M5E5	SpEGA5-11-1, SpEGA5-12-1	11, 12
Green Line 5	F2G5	SpGL5-1-19-1, SpGL5-1-10-2	10, 19
	F3G5	SpGL5-1-10-4, SpGL5-1-13-2	10, 13
	F4G5	SpGL5-1-14-1	14
	M6G5	SpGL5-1-3-1	3
	M7G5	SpGL5-1-4-1	4
	M8G5	SpGL5-1-10-1, SpGL5-1-13-1	10, 13
	M9G5	SpGL5-1-10-3, SpGL5-1-19-2	10, 19

Table 12. Overview of the Australian English speakers and audio tracks per textbook.

Although, one or two audio tracks were analysed per speaker most often, three audio tracks were analysed in one instance. Audio tracks three and four were analysed first for speaker M3E5. Track number five was added to the analysis for speaker M3E5, as it was part of the analysis for speakers F1E5 and M2E5, and therefore already part of the corpus. The audio tracks from the textbook *English G Access 5* are taken from the sole CD, while the audio tracks from the textbook *Green Line 5* are all taken from the first of three CDs.

The main features of interest in this sub-corpus are the diphthongs FACE, PRICE, and MOUTH. These diphthongs are salient markers that differentiate Australian English from other accents and are also markers of social variation within Australian English. PRICE and MOUTH particularly, display phonetic differences between Broad and General Australian English, while FACE shows accent differences in Cultivated Australian English compared to Broad and General Australian English. The analysis focuses on the first targets in the diphthongs FACE, PRICE, and MOUTH, as "the second target of diphthongs is much more variable than the first and often not attained [...] even in citation-form speech" (Harrington, Cox, and Evans 1997: 174).

Speaker	FACE	PRICE	MOUTH
F1E5	8	7	4
M1E5	50	36	27
M2E5	35	24	20
M3E5	8	10	1
M4E5	30	20	19
M5E5	24	47	15
F2G5	9	27	4
F3G5	30	20	9
F4G5	18	24	11
M6G5	15	27	7
M7G5	18	37	10
M8G5	18	18	9
M9G5	17	10	6
Total	**280**	**307**	**142**

Table 13. Token numbers per speaker for FACE, PRICE, and MOUTH.

Table 13 displays the raw frequency of the features under discussion per speaker and the total frequency per vowel in this sub-corpus. As is immediately apparent, the distribution of the features across the corpus is quite uneven. PRICE occurs the most with 307 tokens in total and ranging from only seven tokens for speaker F1E5 to 47 tokens for speaker M5E5. FACE occurs a total of 280 times in the corpus. Speaker M1E5 has by far the most FACE tokens with 50. Overall, MOUTH occurs less in this corpus with only 142 tokens in total and a range of one token for speaker M3E5 to 27 tokens for speaker M1E5. This distribution, however, is not that surprising considering the nature of the data set. The individual audio tracks differ considerably in length, ranging from just over a minute to almost six and half minutes. Therefore, the available data set per speaker varies greatly, resulting in a wide range of token numbers per speaker. Moreover, the vowels FACE and PRICE occurred more often in the data set compared to MOUTH, for two possible reasons: (1) FACE and PRICE have a higher frequency of occurrence[6] in English compared to MOUTH (cf. Cruttenden 2014: 159), and (2) the highly frequent pronoun *I*, which already accounts for almost 30% of the PRICE tokens,

[6] While Cruttenden shows these frequencies of occurrence for British English vowels, he also states that the frequency in General American is similar. Thus, it can be inferred that at least the ratio of frequency between these vowels is likely to be similar in Australian English.

can be expected to occur frequently in textbooks consisting largely of conversation-style texts. Moreover, in textbook units about Australia, the word *Australia* or its derivatives, containing the vowel FACE, occur quite frequently as well. Thus, the overall distribution of these vowels in the corpus is not surprising.

Furthermore, the sub-corpus was tagged for the stressed monophthongs to show the characteristics of the three diphthongs in relation to the individual speakers' monophthongal vowel space (see Section 3.3.1). The monophthongs FLEECE, KIT, DRESS, TRAP, NURSE, STRUT, BATH/PALM/START, GOOSE, FOOT, THOUGHT/NORTH, and LOT were added to the corpus. In general, up to five tokens per monophthong were analysed per speaker, though occasionally more than five tokens were annotated. In these cases, all annotated tokens were analysed. Table 14 and Table 15 below provide an overview of the token numbers analysed for each monophthong per speaker.

Speaker	**FLEECE**	**KIT**	**DRESS**	**TRAP**	**NURSE**	**STRUT**
F1E5	5	4	4	5	0	2
M1E5	5	5	5	5	5	5
M2E5	5	5	5	7	5	6
M3E5	4	6	6	3	0	2
M4E5	5	5	5	5	4	5
M5E5	5	5	5	5	5	5
F2G5	5	5	5	6	3	5
F3G5	5	5	5	5	2	4
F4G5	5	4	5	5	5	5
M6G5	5	5	6	4	5	6
M7G5	6	10	5	6	0	6
M8G5	5	5	5	6	1	4
M9G5	5	5	5	2	2	3
Total	**65**	**69**	**66**	**64**	**37**	**58**

Table 14. Token numbers per speaker for the monophthongs FLEECE, KIT, DRESS, TRAP, NURSE, and STRUT.

Speaker	BATH/PALM/ START	GOOSE	FOOT	THOUGHT/ NORTH	LOT
F1E5	2	5	2	2	5
M1E5	2	6	4	5	5
M2E5	5	5	4	5	9
M3E5	6	3	2	3	1
M4E5	4	5	3	5	5
M5E5	5	5	5	5	5
F2G5	5	4	5	3	5
F3G5	3	2	3	4	6
F4G5	5	5	4	5	5
M6G5	5	3	3	4	5
M7G5	9	4	3	7	5
M8G5	4	5	2	7	4
M9G5	3	4	1	0	5
Total	**58**	**56**	**41**	**55**	**65**

Table 15. Token numbers per speaker for the monophthongs BATH/PALM/START, GOOSE, FOOT, THOUGHT/NORTH, and LOT.

3.3 Data Analysis

The data analysis of this study focuses on conducting acoustic and auditory analyses to provide insights into how Received Pronunciation, General American, and Australian English are represented in the phonological textbook corpus. This section outlines the methodologies and tools used to examine and interpret the speech data systematically. Section 3.3.1 outlines the process of acoustically analysing vowels and visualising the results. This process has been partially described before in Scheiwe (2022). Section 3.3.2 focuses on the auditory analysis of rhoticity. Section 3.3.3. presents how the results of the acoustic and auditory analyses are evaluated in relation to the accent descriptions presented in Chapter 2.

3.3.1 Acoustic Analysis of Vowels and Visualisation

The corpus data was analysed acoustically using Praat (Boersma and Weenink 2020). The first three formants (F1, F2, and F3) were analysed for each vowel using Praat's automated LPC formant analysis (see Ladefoged and Johnson, 2011 for a detailed explanation).

In preparation for analysis, the audio data was annotated in a time-aligned manner on three different tiers: the speaker tier, the word tier, and the vowel tier (see Section 3.2.2. for a detailed description of the annotation practices). The relevant sections for the analysis were annotated on the speaker tier indicating each speaker with an individual code. This annotation was performed for each word included in the analysis. The relevant words for the analysis were annotated on the word tier and each relevant vowel within these words was annotated on the vowel tier. The vowel segmentation was based on spectrographic cues. During this annotation process, a visual inspection of the spectrogram corroborated Praat's automated LPC analysis (cf. Thomas 2011: 48). Figure 7 below shows a screenshot of the annotated Praat file for the audio track EGA5-5 from the Australian English sub-corpus.

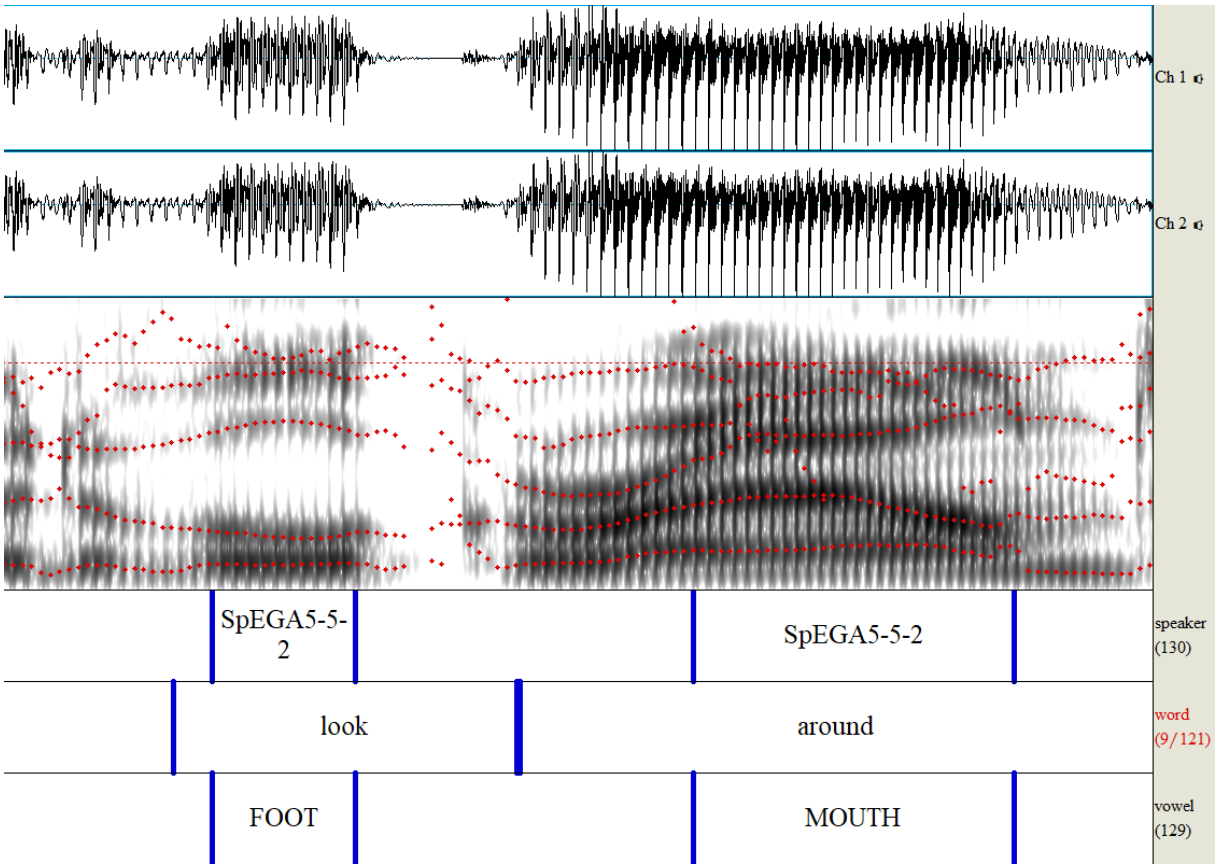


Figure 7. Screenshot of the Praat window for the annotated track EGA5-5 from the Australian English sub-corpus.

Figure 7 shows the annotation practice at the speaker, word, and vowel level for the audio track EGA5-5. The speakers are labelled with the original speaker code that indicates the audio track number as well as the speakers' occurrence on the audio track. The words are annotated orthographically, while the vowels are annotated using Wells' (1982a) lexical sets. The words and vowels are annotated in a time-aligned manner. The word- and vowel-boundaries are indicated by the blue vertical boundary marks in the annotation tiers.

The annotated Praat file of the audio track EGA5-5 is included in the Digital Appendix to further showcase the annotation and segmentation practices. The exact time stamps of all segmented words and vowels are included in the Excel files in the Digital Appendix (An overview of files included in the Digital Appendix can be found in Appendix 7).

The formant values F1, F2, and F3 were extracted with a formant ceiling of 5.000 Hz for male speakers, and 5.500 Hz for female speakers (Boersma 2020). These formant ceilings were used for all speakers from the Australian English and General American sub-corpora. As the speakers from the Received Pronunciation sub-corpus portrayed mainly children, these values had to be adjusted manually, as children often require a ceiling value of up to 8.000 Hz (Boersma 2020). Some speakers in the Received Pronunciation sub-corpus did not yield accurate readings at the standard ceilings of 5.000 Hz for male speakers and 5.500 Hz for female speakers. In those cases, the formant ceiling was determined individually by adjusting the formant ceilings to match the automatic readings from the LPC analysis with the formant bands visible on the spectrogram.

Table 16 below provides an overview of the individual formant ceiling values for the children in the Received Pronunciation sub-corpus, for whom the traditional formant ceilings of 5.000 Hz for male speakers and 5.500 Hz for female speakers did not provide an accurate match between the automatic LPC readings and the spectrogram. These adjusted formant ceilings yielded more accurate results.

Speaker	Original speaker code	Formant ceiling (Hz)
F15E2	SpEGA2-1-39-2	7000
	SpEGA2-1-44-2	6500
F19G2	SpGL2-2-33-2	6000
	SpGL2-2-34-2	6000
F20C2	SpCT2-2-20-1	7000
	SpCT2-2-28-2	7000
F21C2	SpCT2-2-7-2	6000
	SpCT2-2-34-2	6000
F22C2	SpCT2-2-20-2	7000
	SpCT2-2-28-1	7000
	SpCT2-1-36-1	6000
F23C2	SpCT2-1-36-2	6000
	SpCT2-1-38-2	6000
F24C2	SpCT2-2-36-1	7000
	SpCT2-2-30-1	6500
M20E2	SpEGA2-1-44-3	6500
M26G2	SpGL2-1-9-2	5300
	SpGL2-1-10-2	5300
M28G2	SpGL2-1-20-2	5500
	SpGL2-1-6-1	5000
M29C2	SpCT2-1-35-3	6000
	SpCT2-2-23-3	6000
M30C2	SpCT2-1-16-2	5500
	SpCT2-2-14-2	5500

Table 16. Overview of individual formant ceiling settings for the children in the Received Pronunciation sub-corpus.

One vowel target was identified for monophthongs to analyse the pronunciation of the individual monophthongs. The monophthongal vowel target was selected at the midpoint 50% into the vowel (cf. Harrington and Cassidy 1999:59-60; Rosner and Pickering 1994: 79). While the monophthongal vowel target is also often measured at a section in the vowel with relatively steady-state formant readings (Harrington and Cassidy 1999: 59), measurements at the temporal midpoint were selected for this study, as "many monophthongal vowels have no clearly identifiable steady-state section or else because the steady-state interval is different for each formant" (Harrington and Cassidy 1999: 59). Van Son and Pols (1990) investigated the validity of different methods of obtaining the vowel target in monophthongs and found that "the differences between the various methods used are, in most respects, marginal and all methods used essentially give the same outcome. When studying vowel targets, the method that is most convenient can be used" (van Son and Pols 1990: 1692).

Two vowel targets were selected in diphthongs to analyse the onset and glide in each vowel. The first vowel target was measured at 25% and the second target at 75% of the vowel's duration (cf. Thomas 2011: 151-152). Another option is to measure the two vowel targets at 30% and 70% into the vowel. Both are frequently used options for determining vowel targets in diphthongs (Thomas 2011: 151-152). Measurements at all three targets—25%, 50%, and 75% of the vowel's duration—were taken for all vowels, regardless of whether they are monophthongs or diphthongs to simplify the data collection process. Only the measurements at the relevant vowel targets were considered in the analysis: the measurement at 50% for monophthongs, 25% for the first target in diphthongs, and 75% for the second target in diphthongs. The formant values (in hertz) for F1, F2, and F3 at the three vowel targets are reported as whole numbers without decimals.

The formant readings were extracted using two different Praat scripts (Appendix 3). The first script was used to extract the data for the Australian English sub-corpus and was then slightly modified for the subsequent analyses of the Received Pronunciation and the General American sub-corpora. These practical modifications did not change the technical mechanics of the script. In the first script, the formant ceilings for male and female speakers had to be adjusted manually. Thus, the script had to be run twice for tracks that included both male and female speakers. The resulting .txt files then had to be manually filtered for the appropriate formant readings of the individual speakers. The second Praat script runs two different loops for tracks with both male and female speakers and therefore provides the appropriate formant readings in only one iteration. Instead of saving the results in three individual .txt files, this script saves the results in one Excel file. All automatic formant

readings were manually re-checked with an inspection of the spectrogram and adapted if necessary.

After extracting the formant values as described above, each speaker's monophthongs were visualised in an F1xF2 scatter plot based on the F1 and F2 values extracted at the vowel target. These vowel plots were created using R Statistical Software (R Core Team 2021). The scripts used for creating the vowel plots and an overview of the packages used in the scripts are included in Appendix 4. Each individual speaker's monophthongs were plotted on an F1xF2 scatter plot to evaluate the vowel analysis. Average F1 and F2 formant values for each vowel were plotted to create each individual speaker's vowel space. The resulting vowel plots approximate the articulatory vowel quadrilateral, with F1 inversely corresponding to vowel height, and F2 inversely corresponding to vowel frontness (Johnson 2012: 144; Rosner and Pickering 1994: 11).

While Received Pronunciation and General American are evaluated solely on the pronunciation of monophthongs—and rhoticity in the case of General American, the representation of Australian English in the textbook audio materials is analysed regarding the diphthongs FACE, PRICE, and MOUTH. Diphthong trajectories from the first to the second vowel target were superimposed onto the vowel plots for the speakers of the Australian English sub-corpus. For each speaker, the mean F1 and F2 values from the two vowel targets in the diphthongs FACE, PRICE, and MOUTH were used to plot these diphthong trajectories.

In order to interpret a speaker's realisation of the respective vowels under discussion, the vowel plots were segmented manually into the articulatory categories of front, central, and back on the horizontal (F2) axis, and open, open-mid, mid, close-mid, and close on the vertical (F1) axis. To achieve this articulatory categorisation of the individual vowel spaces, each plot was vertically segmented into three equidistant areas to categorise each vowel in terms of vowel height. The segmentation into front, central, and back vowels was achieved by dividing the vowel spaces in the middle connecting the midpoints of the two parallel sides in the quadrilateral. These segmentations were constructed using GeoGebra® (2024). Figure 8 below illustrates this segmentation.

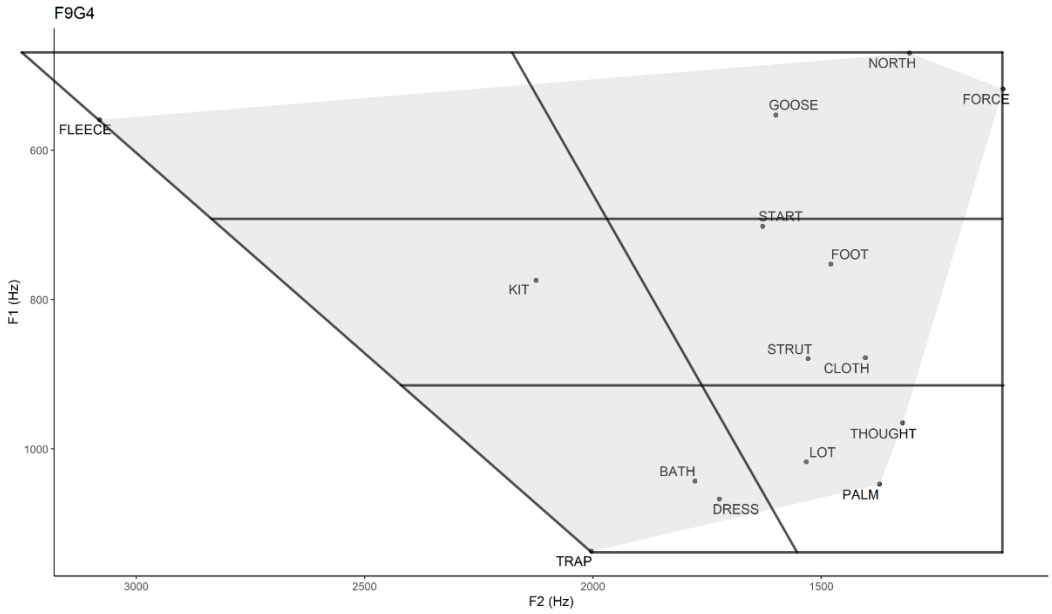


Figure 8. Averaged vowel plot from speaker F9G4 with added articulatory categorisation.

Each vowel plot is categorised in this manner. The top parallel side of the quadrilateral is constructed to run through the vowel with the lowest F1 value, in this case NORTH. In the upper right-hand corner, there is a right angle with the right-hand side of the quadrilateral running through the vowel with the lowest F2 value, in this case FORCE. The second parallel side of the quadrilateral at the bottom goes through the vowel with the highest F1 value, in this case TRAP. The left-hand side of the quadrilateral is constructed in such a way, that it goes through the vowel with the highest F2 value, in this case, FLEECE. The central line is constructed by connecting the two midpoints of the two parallel sides at the top and the bottom of the quadrilateral. The two parallel lines in the middle and the two parallel sides at the top and bottom of the quadrilateral are equidistant.

Four of the Australian English vowel spaces have a rather triangular or diamond shape. For these plots, the categorisation is slightly amended to resemble this triangular/diamond shape. The categorisation in these plots does not have two parallel sides at the top and bottom, instead, the categorisation matches the triangular point in the open area (see Appendix 2 for these amended categorisations). The segmentation into three equidistant levels is maintained. The segmentation into front, central, and back is achieved by connecting the midpoint of the top line to the triangular point at the bottom. This method of segmentation was used for the vowel plots of speakers F1E5, F2E5, M3E5, and F4G5.

This categorisation of each plot does not fully match the dimensions of the cardinal vowel chart, as the bottom parallel side of the quadrilateral is not necessarily half the length of the top side of the quadrilateral. Instead, these sides were matched to fit each individual person's vowel space. Nonetheless, categorising each individual vowel plot in this way makes it possible to describe the vowels according to the articulatory features of vowel height and tongue advancement. The individual vowel plots with these categorisations can be found in Appendix 2.

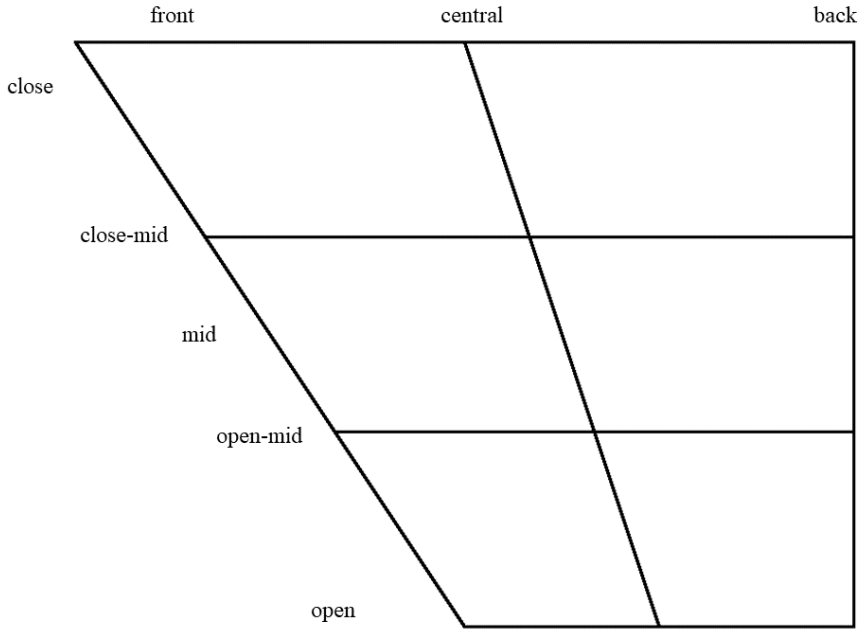


Figure 9. Vowel quadrilateral with articulatory labels.

Figure 9 shows a conventional vowel quadrilateral with the articulatory labels attached. Thus, a vowel was classified as open front when it was located in the lower third and front half of the quadrilateral. Vowels in close vicinity to or directly on the central line were classified as central vowels. Vowels in the back half of the quadrilateral are classified as back vowels. Similarly, vowels in close vicinity to or on the lines for an open-mid or close-mid vowel height were categorised as such. Vowels between the open-mid and close-mid areas were categorised as mid vowels. Vowels above the close-mid line were categorised as close vowels.

A vowel plot was created for each individual speaker based on the average F1 and F2 values for each of that speaker's monophthongs. Three to four additional vowel plots were created per sub-corpus: a vowel plot based on the pooled averages for all speakers of

the sub-corpus and vowel plots based on the pooled averages for the speaker groups from the individual textbooks. As the Australian English sub-corpus only consists of two textbooks, three additional vowel plots were created: the vowel plot of the pooled averages for all speakers and the two vowel plots for the textbook speaker groups. These pooled averages were calculated based on each speaker's individual average formant values per vowel. For instance, the Australian sub-corpus consists of 13 speakers. The mean F1 and mean F2 values per monophthong, and for these speakers also the mean F1 and F2 values for the two vowel targets in the diphthongs, were calculated for each speaker. These mean formant values were used again to calculate the pooled average for the entire speaker group. All 13 mean F1 and mean F2 values per vowel were added and then divided by the number of speakers. These pooled average F1 and F2 values were then used to plot the vowel space of the entire speaker group. The same procedure was repeated for the speakers of the individual textbooks. All vowel plots from each sub-corpus can be found in Appendix 1. The categorised versions of these vowel plots and the categorised versions of the grouped vowel plots can be found in Appendix 2. The R scripts used to create the grouped and speaker-individual vowel plots are found in Appendix 4.

All vowels in this corpus were categorized according to their position in the individual vowel spaces. The monophthongs that are used as the reference points in these vowel plots are always the speaker-individual mean values. For the pooled averages, these monophthongs are the calculated pooled mean values across that respective speaker group.

This section provided an overview of the process of analysing vowels, reported the analysis settings in Praat, and explained the visualisation of the results. The next section focuses on the auditory and acoustic analysis of rhoticity.

3.3.2 Auditory and Acoustic Analysis of Rhoticity

In the General American sub-corpus, the lexical sets NEAR, SQUARE, and CURE were analysed with regard to rhoticity. A total of 212 tokens were analysed. These tokens were categorised according to Harris' (2013) rhoticity systems and the different phonological contexts in which /r/ can be articulated. As the rhoticity analysis was confined to the lexical sets NEAR, SQUARE, and CURE, not all nine phonological contexts that Harris identified (see Section 2.4) are applicable in this analysis. As /r/ cannot occur either word-initially or in consonant clusters within the lexical sets of NEAR, SQUARE, and CURE, the categories of the phonological context of /r/ in the General American sub-corpus used in this analysis are:

(i) Intervocalically within a word before a stressed syllable
(ii) Intervocalically within a word before an unstressed syllable
(iii) Intervocalically across a word boundary before a stressed syllable
(iv) Intervocalically across a word boundary before an unstressed syllable
(v) Preconsonantally within a word
(vi) Preconsonantally across a word boundary
(vii) Word/phrase/utterance final

Tokens in category (vii) were identified according to the methodology presented by Robb and Gillon (2007: 3), who defined "[an] utterance […] as a string of syllables that did not contain a silent interval in excess of 250 ms." Consequently, a token was marked as word- or phrase-final if it was followed by a pause exceeding 250 ms, which signals a definite pause or utterance break. Tokens were identified as utterance-final if either a change in speakers occurred or the token was the final word on the respective audio track.

Rhoticity in the lexical sets NEAR, SQUARE, and CURE in the General American sub-corpus was analysed auditorily, categorising /r/ as either present or absent in these tokens. In unclear cases, an additional acoustic analysis in the form of an inspection of the spectrogram was conducted, as acoustically, rhoticity in these lexical sets is marked with a distinct lowering of F3 towards the end of the vowel (Kent and Read 1992: 139; Ladefoged and Maddieson 1996: 244). Analysing rhoticity primarily auditorily and only using additional acoustic analyses in either a portion of the data or in unclear cases is commonly used in the field (cf. Dickson and Hall-Lew 2017; Lonergan and Cox 2010; Redzwan 2016; Stuart-Smith, Lawson and Scobbie 2014; Tan 2012).

The validity of the analysis was ensured by further auditory analysis of the data from a second judge, also trained in linguistics, who categorised /r/ as present or absent in the tokens. Both the researcher and the second judge are non-rhotic speakers of English: the second judge is a native speaker of Singapore English. The individual classifications of /r/ in the tokens as either present or absent by both researchers yielded a different classification in 28 out of 212 tokens, creating an inter-judge agreement rate of 86%. The results from the auditory and acoustic analyses by the researcher are the basis for discussing the state of rhoticity in the lexical sets NEAR, SQUARE, and CURE in the General American sub-corpus.

Whether /r/ in the lexical sets NEAR, SQUARE, and CURE is articulated or not, is the main focus of this analysis. First, an overview was created of how many tokens per lexical set and phonological context were analysed. Then the rate of /r/ articulation was calculated

per phonological context across all speakers, as well as the individual textbook groups. In these cases, the absolute frequencies of articulated and not-articulated /r/, as well as the percentage of articulated /r/, are calculated. The level of rhoticity per speaker is also calculated to determine how many speakers are fully rhotic in the lexical sets NEAR, SQUARE, and CURE, and how many speakers display variable rhoticity. These cases of variable rhoticity are then examined more closely focusing on the individual instances in which /r/ is not articulated for those speakers.

This section discussed the methodological steps in the rhoticity analysis. The next section turns to the evaluation of the findings from the acoustic and auditory analyses and explains the methodological process of comparing the findings to the theoretical framework presented in Chapter 2.

3.3.3 Synthesis of Corpus Findings and Linguistic Descriptions of Accents

The results of the acoustic and auditory analyses are compared to linguistic descriptions of the accents in question, in order to evaluate the representation of Received Pronunciation, General American, and Australian English in the textbook series *English G Access*, *Green Line*, and *Camden Town*. The categorisation of the vowels under investigation into the articulatory categories of vowel height and tongue advancement are compared to the linguistic descriptions presented in Chapter 2. The results are then evaluated as to whether the speakers' pronunciation matches common pronunciation variants of the accent under discussion. First, the overall pronunciation per sub-corpus is evaluated based on the grouped vowel plots for all speakers and the textbook groups. Then, each individual speaker's pronunciation is evaluated based on their averaged vowel space.

3.4 Identifying and Addressing Limitations

The limitations of this study have been defined regarding the data and the methodology employed. The following sections explain and address limitations related to the corpus size and the resulting representativeness of the data, as well as the method of identifying and selecting speakers for the corpus.

3.4.1 Data Limitations

The specialized phonological textbook corpus is fairly small, with only 4.730 tokens. The sub-corpora are consequently quite small as well, ranging from 1.365 tokens for the Australian English sub-corpus to 1.732 tokens for the General American sub-corpus. This results in low token numbers per speaker and feature in the individual sub-corpora. A good corpus generally should be both, representative and of sufficient size (Gut and Voorman 2014: 19). Representativeness, in this case, refers "to the objective that the raw data of a corpus should constitute a sample of language or a language variety that includes its full range of variability" (Gut and Voorman 2014: 20).

The focus of the specialized phonological textbook corpus lies in the representation of three select accents of English based on select phonological features. Certainly, an analysis of more accents of English represented by all phonological features of the respective accents may represent the full range of variability within the textbook data. Although this may suggest that the corpus may not be truly representative of the representation of varieties of English in textbooks used in NRW in general, the corpus still offers a high level of representativeness. The analysed features represent salient aspects of the accents under discussion. As all available tokens of these features within the audio tracks are included in the analyses, this phonological corpus still showcases the range of variability present in the textbook data.

Factors such as the project's time constraints influenced the corpus compilation, annotation, and analysis and, therefore, also the degree of representativeness that can be achieved. However, "true representativeness is not always possible in corpus design" (Gut and Voorman 2014: 21). Larger corpora that show a high degree of representativeness usually also involve sizeable research teams and numerous support staff. The corpus size and data selection for this project take into account that this research project is a solo-authored project with minimal outside assistance.

3.4.2 Speaker Limitations

As stated in Section 3.2.2, the speakers chosen for this sub-corpus were selected because they were assumed to be representative of the respective accent under discussion. While all available speakers were chosen for the Australian English sub-corpus, a selection of seven speakers per textbook were analysed for each textbook in the Received Pronunciation sub-corpus, as well as for the textbooks *English G Access 4* and *Green Line 4* in the General American sub-corpus. Six speakers were selected for analysis from the textbook *Camden Town 4* in the General American sub-corpus. As illustrated in Sections 3.2.3 and 3.2.4, these speakers were chosen based on the amount of data available per speaker to ensure as substantial a number of tokens for analysis as possible. However, as textbooks tend to contain rather short texts, the data pool for each speaker is small and varies significantly per speaker. While speakers with longer texts were selected to ensure a sufficient number of tokens for analysis, the final amount of data available still varies significantly per speaker (cf. Table 9, Table 11, and Table 13).

Multiple audio files were chosen for the analysis of each speaker. In some cases, multiple audio files covered the same text; in other cases, the speakers voiced different characters. Information about which voice actor voiced which character was not available. The 'identity' of the speakers had to be determined auditorily. A certain degree of error in determining whether a textbook character was voiced by the same voice actor is possible; however, the validity of the results was ensured by a second judge with linguistic training who independently identified which voice actor voiced which character or text.

3.5 Closing Remarks

This chapter outlined the specialized phonological textbook corpus design, and the acoustic and auditory analysis methodology used in this research. Following practical considerations of the corpus compilation process, the general corpus design and annotation methods were discussed before the corpus design for each varietal sub-corpus. Subsequently, the chapter presented the acoustic and auditory analysis methodology, as well as the process of comparing and contrasting the findings from these analyses with linguistic descriptions of the accents under discussion. Finally, the chapter addressed the study's limitations. The next chapter presents the results of the acoustic and auditory analyses of the corpus data.

4 Results

4.1 Introduction

This chapter presents the findings of the acoustic and auditory analyses of the textbook audio materials regarding the representation of Received Pronunciation, General American, and Australian English in English textbooks used at secondary schools in NRW. Section 4.2 presents the findings from the Received Pronunciation sub-corpus, Section 4.3 presents the findings from the General American sub-corpus, and Section 4.4 presents the findings from the Australian English sub-corpus.

4.2 Received Pronunciation

This section investigates how Received Pronunciation, and specifically the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START, are represented in the German textbooks of English: *English G Access 2*, *Green Line 2*, and *Camden Town 2*. Section 4.2.1 illustrates the results of the acoustic analyses of the corpus data to depict the Received Pronunciation accent of the textbooks. The section provides a general overview of the pronunciation of the entire speaker group and detailed analyses of the three textbook speaker groups and the individual speakers. This analysis provides a comprehensive overview of how Received Pronunciation is depicted in German textbooks of English.

Section 4.2.2 expands on these findings: a comparison with the linguistic description of Received Pronunciation as presented in Section 2.5 facilitates an interpretation of these findings. The results of the acoustic analyses are evaluated against the descriptions of Received Pronunciation.

4.2.1 The Vowels in the Received Pronunciation Sub-Corpus

Speech samples from 21 textbook speakers have been analysed to investigate the representation of Received Pronunciation as exemplified by the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. Seven speakers each were analysed from the textbooks *English G Access 2*, *Green Line 2*, and *Camden Town 2*. A combined amount of 846 tokens for GOOSE, LOT, CLOTH, TRAP, BATH, and START were analysed across these 21 speakers.

To get a first impression of the representation of Received Pronunciation in German textbooks, four speaker groups were created: a group of all speakers, all speakers from

English G Access 2 (EGA2), all speakers from *Green Line 2* (GL2), and all speakers from *Camden Town 2* (CT2). For each group, pooled averages for F1 and F2 of the vowel targets in GOOSE, LOT, CLOTH, TRAP, BATH, and START were calculated based on the groups' speakers mean raw F1 and F2 values in these vowels.

Speaker groups	GOOSE				LOT				CLOTH			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
EGA2	F1	411.7	67.3	7	F1	676.4	63.3	7	F1	685.5	132.9	6
	F2	1717.2	399.8		F2	1168.7	135.9		F2	1234.9	184.2	
GL2	F1	371.5	45.6	7	F1	593.4	147.1	7	F1	625.3	67.6	6
	F2	1800.6	361.9		F2	1120.3	90.6		F2	1186.6	137.1	
CT2	F1	414.9	65.2	7	F1	720.7	114.4	7	F1	814.1	84.2	7
	F2	1892.7	342.0		F2	1250.0	160.9		F2	1274.8	170.1	
All	**F1**	**399.4**	**59.1**	**21**	**F1**	**663.5**	**117.7**	**21**	**F1**	**713.9**	**120.5**	**19**
	F2	**1803.5**	**348.8**		**F2**	**1179.7**	**133.7**		**F2**	**1234.3**	**156.0**	

Table 17. Pooled mean formant values and standard deviations for GOOSE, LOT, and CLOTH for the speaker groups from English G Access 2, Green Line 2, Camden Town 2, and all speakers.

Speaker groups	TRAP				BATH				START			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
EGA2	F1	976.4	78.7	7	F1	919.5	116.1	7	F1	829.2	75.1	6
	F2	1604.1	108.1		F2	1340.0	130.5		F2	1206.4	88.2	
GL2	F1	923.2	88.1	7	F1	760.6	152.1	7	F1	681.3	183.0	6
	F2	1612.7	120.4		F2	1231.0	129.3		F2	1202.2	91.3	
CT2	F1	1062.6	121.4	7	F1	972.0	134.3	6	F1	926.8	141.4	6
	F2	1695.0	101.1		F2	1524.5	233.7		F2	1475.6	294.6	
All	**F1**	**987.4**	**107.2**	**21**	**F1**	**879.6**	**153.5**	**20**	**F1**	**812.4**	**163.1**	**18**
	F2	**1637.3**	**109.9**		**F2**	**1357.2**	**194.6**		**F2**	**1294.7**	**211.9**	

Table 18. Pooled mean formant values and standard deviations for TRAP, BATH, and START for the speaker groups from English G Access 2, Green Line 2, Camden Town 2, and all speakers.

Table 17 and Table 18 display the pooled averages in hertz of the F1 and F2 values for the vowel targets in the vowels GOOSE, LOT, CLOTH, TRAP, BATH, and START and the standard deviations (SD) per speaker group per vowel. Both the mean F1 and F2 values and the SD are rounded to one decimal place. The token number (*n*) indicates the number of speakers per group, as these mean values were calculated based on the speaker-individual mean F1 and F2 values per vowel. The pooled averages of F1 and F2 for the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START are visualised in Figure 10–Figure 13 below. Figure 10 displays the mean vowel plot for the entire speaker group, Figure 11 displays the mean vowel plot for the speakers from *English G Access 2*, Figure 12 displays the mean vowel plot for the speakers from *Green Line 2*, and Figure 13 displays the mean vowel plot for the speakers from *Camden Town 2*.

The pooled mean vowel realisations of all monophthongs for each speaker group are used as the reference points in each plot. Thus, the vowel plot for the entire speaker group also displays the pooled mean vowel realisations of the remaining monophthongs for this speaker group. The vowel spaces for the speaker groups from *English G Access 2, Green Line 2,* and *Camden Town 2* are each created from the pooled mean vowel realisations of the monophthongs for these respective speaker groups. This procedure provides a general first impression of Received Pronunciation in German textbooks of English. These results are interpreted regarding their representation of Received Pronunciation, i.e. whether these features represent pronunciation features consistent with Received Pronunciation or not, in Section 4.2.2.

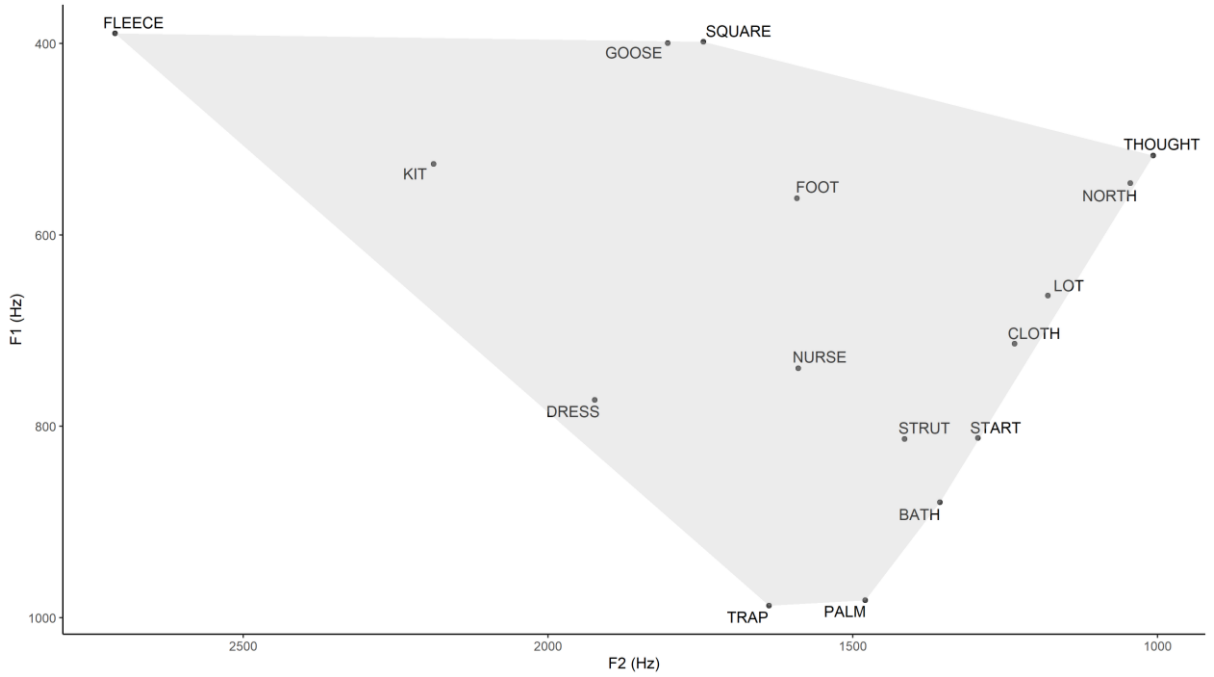


Figure 10. Averaged realisations of the monophthongs for all 21 speakers from the Received Pronunciation sub-corpus.

Figure 10 displays an averaged vowel plot of all speakers' individual averaged F1 and F2 values in hertz. These mean values were calculated from each speaker's averaged F1 and F2 values in hertz for each vowel.

The vowels of interest here are the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. The categorisation of these vowels according to the articulatory features of vowel height and tongue advancement are based on the 'segmented' vowel plots in Appendix 2. These vowel plots include a segmentation into the articulatory categories of front, central, and back on the horizontal (F2) axis, and close, close-mid, mid, open-mid, and open on the vertical (F1) axis.

If all speakers in the Received Pronunciation sub-corpus are considered as one group, the following realisations of the features under discussion can be observed. GOOSE has a fronted position for this speaker group and can be described as a close central vowel. TRAP has an open front position and has the most open position in the vowel space. BATH and START are both open back vowels for this speaker group, with BATH more open than START. The vowels LOT and CLOTH are both mid back vowels for this speaker group.

While considering all speakers of the sub-corpus as one cohort provides a comprehensive overview of the depiction of Received Pronunciation in English textbooks used in NRW, examining the speaker groups from each textbook individually offers unique insights

into pronunciation patterns. Therefore, the subsequent paragraphs delve into the results for each textbook speaker group, starting with the textbook *English G Access 2*.

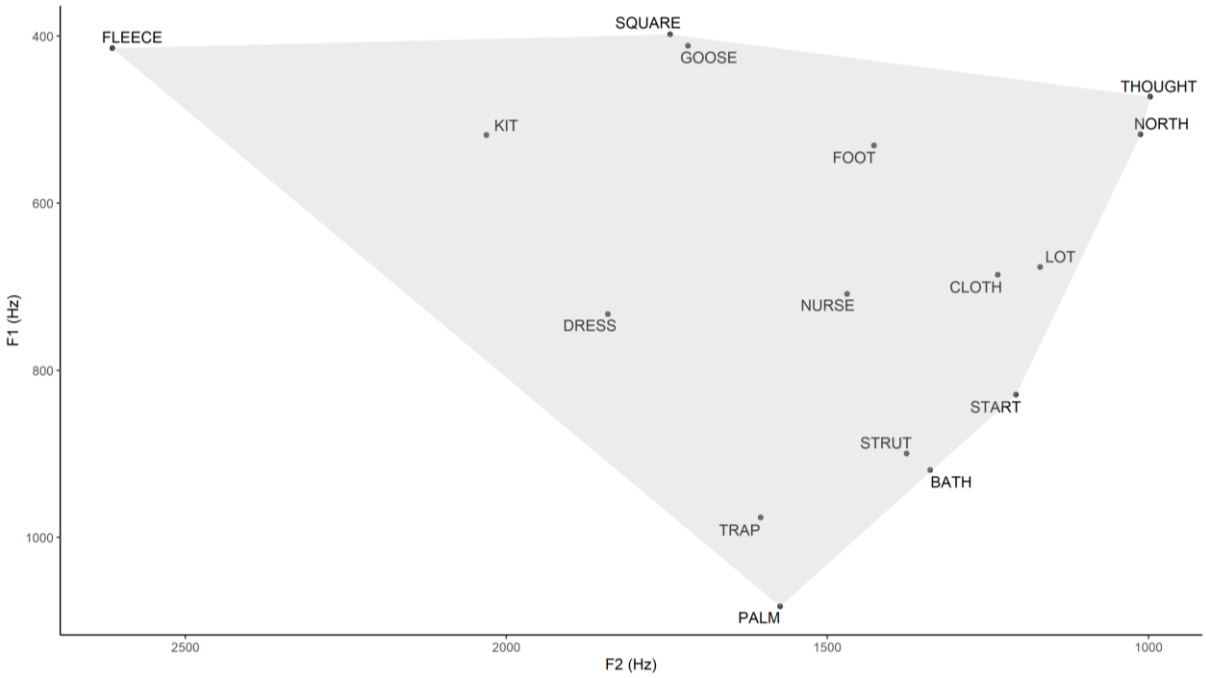


Figure 11. Averaged realisations of the monophthongs for the seven English G Access 2 speakers.

Figure 11 shows the averaged vowel plot for all seven speakers from the textbook *English G Access 2*. The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook. This vowel plot strongly resembles the one for the entire speaker group in form and in the realisation of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. This speaker group shows a fronted realisation of the GOOSE vowel with a close central pronunciation. TRAP is a fairly open front vowel with a quality that is less open than PALM but more open than BATH and STRUT. BATH is an open back vowel for this speaker group, which is slightly more open and retracted than STRUT. This vowel is positioned between fully open and open-mid. START has an open-mid back position for this speaker group. LOT and CLOTH are both mid back vowels for this speaker group. CLOTH is slightly fronted compared to LOT.

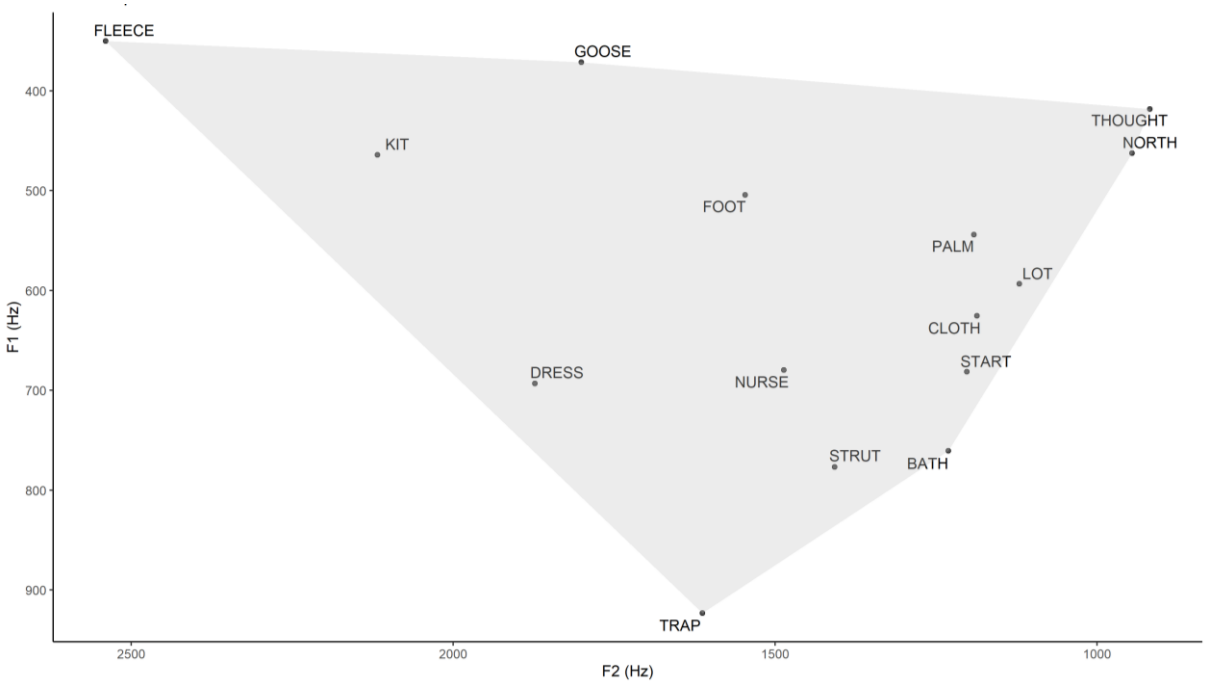


Figure 12. Averaged realisations of the monophthongs for the seven Green Line 2 speakers.

Figure 12 shows the averaged vowel plot for all seven speakers from the textbook *Green Line 2*. The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook. GOOSE is a close central vowel for this speaker group. TRAP is an open front vowel and the vowel with the highest F1 value for this speaker group. Thus, TRAP is the most open vowel for this speaker group. BATH has an open-mid back position. START, LOT, and CLOTH are all mid back vowels for this speaker group. LOT is the closest vowel of the three, followed by CLOTH and then START.

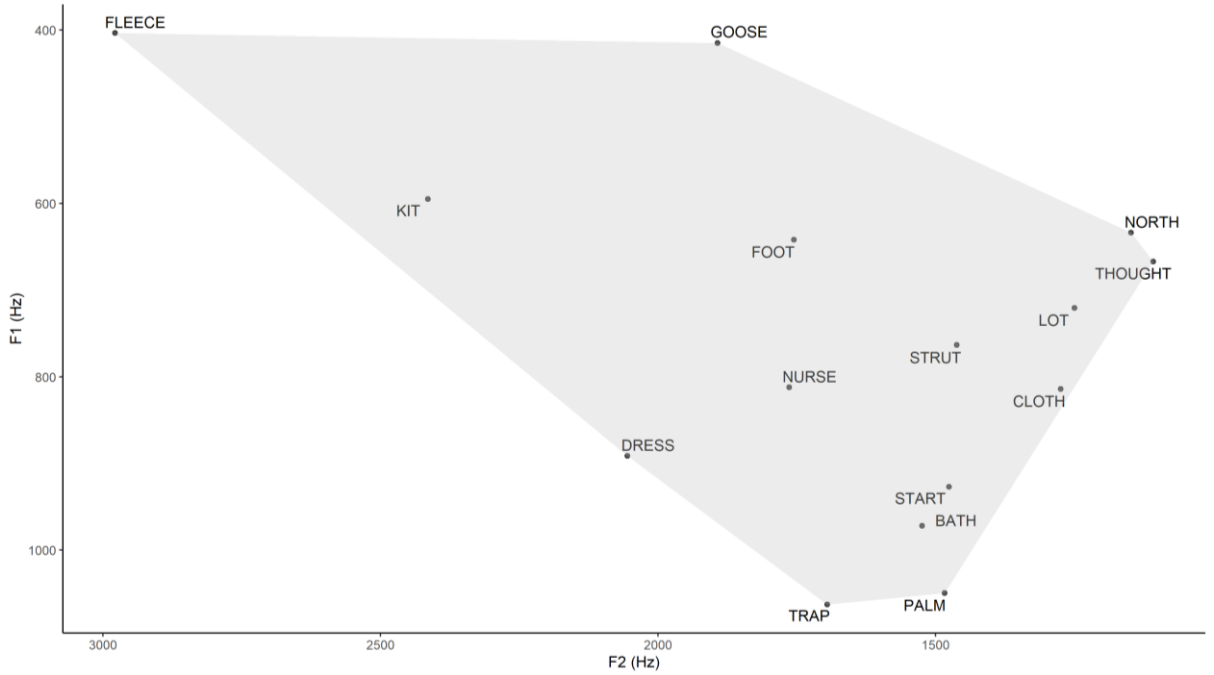


Figure 13. Averaged realisations of the monophthongs for the seven Camden Town 2 speakers.

Figure 13 displays the averaged vowel plot for all seven speakers from the textbook *Camden Town 2*. The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook. GOOSE is fronted from a fully back position but is still a close back vowel for this speaker group. TRAP is an open front vowel and again the most open vowel in the vowel space. BATH is an open central vowel and START is slightly further back and considered an open back vowel. LOT has a mid back position and CLOTH is slightly more open and is considered an open-mid back vowel.

Speaker group	GOOSE	LOT	CLOTH
EGA2	close central	mid back	mid back
GL2	close central	mid back	mid back
CT2	close back	mid back	open-mid back
All speakers	**close central**	**mid back**	**mid back**

Table 19. Overview of articulatory categorisation of GOOSE, LOT, and CLOTH for the speaker groups from English G Access 2, Green Line 2, Camden Town 2, and all speakers.

Speaker group	TRAP	BATH	START
EGA2	open front	open back	open-mid back
GL2	open front	open-mid back	mid back
CT2	open front	open central	open back
All speakers	**open front**	**open back**	**open back**

Table 20. Overview of articulatory categorisation of TRAP, BATH, and START for the speaker groups from English G Access 2, Green Line 2, Camden Town 2, and all speakers.

Table 19 and Table 20 summarise the results from the acoustic analysis and categorisation into the articulatory features of vowel height and tongue advancement for the three textbook speaker groups, and all speakers as a group. The realisation of the monophthongs LOT, and TRAP is rather homogenous across the three textbook speaker groups, and all speakers as a group: LOT is a mid back vowel, and TRAP is an open front vowel for all four speaker groups. GOOSE, CLOTH, BATH, and START display some variability across the different speaker groups. GOOSE is a close central vowel the speaker groups from *English G Access 2*, *Green Line 2*, and for all speakers as a group. The speaker group from *Camden Town 2* has a close back realisation of GOOSE. CLOTH is a mid back vowel for the speakers from the textbooks *English G Access 2*, and *Green Line 2*, and also for all speakers as a group. The speakers from *Camden Town 2*, however, have an open-mid back realisation of this vowel. BATH varies in vowel height and tongue advancement across the speaker groups. For all speakers as a group and for the speakers from the textbook *English G Access 2*, BATH is an open back vowel. For the speakers from *Green Line 2*, BATH is an open-mid back vowel, and for the speakers from *Camden Town 2*, BATH is an open central vowel. START is an open back vowel for all speakers as a group, and also for the speakers from *Camden Town 2*. This vowel has a mid back realisation for the speakers from *Green Line 2*, and an open-mid back realisation for the speakers from *English G Access 2*.

These analyses of the individual speaker groups highlight overall trends within the dataset. A detailed examination of the pronunciation patterns at the speaker level offers more nuanced insights into the depiction of Received Pronunciation in the English textbooks used at German secondary schools. The analyses at the speaker level are based on each speaker's individual vowel space (see Appendix 1), which are built from each speaker's mean F1 and F2 values of their monophthongs. Tables 1-3 (Appendix 5) provides an overview of the speakers' mean F1 and F2 values in hertz and the standard deviation of the vowels under discussion: GOOSE, LOT, CLOTH, TRAP, BATH, and START. The pronunciation of the vowels

under discussion is evaluated based on their position in the vowel space. GOOSE, LOT, CLOTH, TRAP, BATH, and START are categorised according to the articulatory features of vowel height and tongue advancement. The 'segmented' vowel plots for each speaker used for this categorisation can be found in Appendix 2. Table 21 – Table 26 below indicate the different pronunciation patterns of GOOSE, LOT, CLOTH, TRAP, BATH, and START and indicate how many speakers produce these variants.

	Position of GOOSE		
	Close front	Close central	Close back
Number of speakers	9	1	11

Table 21. Realisation of GOOSE based on each speaker's mean formant values.

Table 21 provides an overview of the averaged pronunciation variants of GOOSE in the Received Pronunciation sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong GOOSE has three different pronunciation variants in the Received Pronunciation sub-corpus: close front, close central, and close back. All 21 speakers produce tokens of GOOSE. For 9 of these speakers, GOOSE is a close front vowel. One speaker has a close central realisation in this vowel, and 11 speakers have a close back pronunciation of GOOSE. Thus, just under half of the speakers in this sub-corpus have a fronted realisation in GOOSE ranging from central to front.

	Position of LOT			
	Open-mid back	Mid back	Close-mid back	Close back
Number of speakers	4	10	5	2

Table 22. Realisation of LOT based on each speaker's mean formant values.

Table 22 provides an overview of the averaged pronunciation variants of LOT in the Received pronunciation sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong LOT has four different pronunciation variants in the Received Pronunciation sub-corpus:

open-mid back, mid back, close-mid back, and close back. All 21 speakers produce tokens of LOT. LOT is a back vowel for all 21 speakers that varies in vowel height from open-mid to close. For ten of these speakers, LOT is a mid back vowel. Five speakers have a close-mid back pronunciation of LOT, and four speakers have an open-mid back realisation in this vowel. Two speakers have a close back vowel in LOT.

	Position of CLOTH				
	Open front	Open back	Open-mid back	Mid back	Close-mid back
Number of speakers	1	1	6	5	6

Table 23. Realisation of CLOTH based on each speaker's mean formant values.

Table 23 provides an overview of the averaged pronunciation variants of CLOTH in the Received pronunciation sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong CLOTH has five different pronunciation variants in the Received Pronunciation sub-corpus: open front, open back, open-mid back, mid back, and close-mid back. Only 19 of the 21 speakers produce tokens of CLOTH. CLOTH is a back vowel for almost all of these speakers: CLOTH is a back vowel for 18 speakers, and a front vowel for one speaker. The 18 speakers with a back realisation of CLOTH differ considerably regarding vowel height: one speaker has an open back vowel, six speakers each have an open-mid back, close-mid back realisation in this vowel, and five speakers have a mid back realisation in CLOTH. The only speaker with front vowel has an open front realisation of CLOTH.

	Position of TRAP			
	Open-mid front	Open front	Open central	Open back
Number of speakers	1	17	2	1

Table 24. Realisation of TRAP based on each speaker's mean formant values.

Table 24 provides an overview of the averaged pronunciation variants of TRAP in the Received Pronunciation sub-corpus and how many of the speakers produce these variants

based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong TRAP has four different pronunciation variants in the Received Pronunciation sub-corpus: open-mid front, open front, open central, and open back. All 21 speakers produce tokens of TRAP. For 18 of these speakers, TRAP is an open front vowel: 17 of these speakers have an open front realisation, while one speaker has an open-mid front realisation for TRAP. Two speakers have an open central realisation in this vowel, and one speaker has an open back pronunciation of TRAP.

	Position of BATH					
	Open front	Open central	Open-mid central	Open back	Open-mid back	Mid back
Number of speakers	1	3	1	9	3	2

Table 25. Realisation of BATH based on each speaker's mean formant values.

Table 25 provides an overview of the averaged pronunciation variants of BATH in the Received Pronunciation sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong BATH has six different pronunciation variants in the Received Pronunciation sub-corpus: open front, open central, open-mid central, open back, open-mid back, and mid back. Nineteen of the 21 speakers produce tokens of BATH. BATH is a back vowel for 14 of these speakers: nine of these 14 have an open back vowel for BATH, three have an open-mid back vowel, and two have a mid back vowel. BATH is a central vowel for four speakers: three of these have an open central vowel and one speaker has an open-mid central vowel. One speaker has an open front realisation of BATH. Overall, BATH is an open vowel for 13 speakers, an open-mid vowel for four speakers, and a mid vowel for two speakers.

	Position of START					
	Mid front	Open central	Open back	Open-mid back	Mid back	Close back
Number of speakers	1	2	8	5	1	1

Table 26. Realisation of START based on each speaker's mean formant values.

Table 26 provides an overview of the averaged pronunciation variants of START in the Received Pronunciation sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong START has six different pronunciation variants in the Received Pronunciation sub-corpus: mid front, open central, open back, open-mid back, mid back, and close back. Only 18 of the 21 speakers produce tokens of START. For 15 of these 18 speakers, START is a back vowel: eight speakers have an open back realisation, five an open-mid back realisation, and one speaker each has a mid back and even a close back realisation in START. Two speakers have an open central pronunciation in this vowel and one speaker has a mid front realisation of START.

Overall, GOOSE, LOT, CLOTH, TRAP, BATH, and START display a certain variability in vowel quality in this corpus data. A closer look at how the individual textbook speakers realise these six vowels allows for a more detailed analysis of how Received Pronunciation is presented in the textbook materials. Table 27 and Table 28 below provide an overview of the pronunciation variants in GOOSE, LOT, CLOTH, TRAP, BATH, and START across the individual speakers. The sequence of speakers in these tables is adapted from the sequence introduced in Section 3.2.3 to better showcase interesting pronunciation patterns across several speakers. The sequence of speakers is the same for both tables.

Speaker	GOOSE	LOT	CLOTH
M22E2	close back	mid back	mid back
M23E2	close back	mid back	close-mid back
M24E2	close back	close-mid back	close-mid back
M26G2	close front	mid back	open-mid back
M30C2	close front	mid back	open front
F19G2	close front	mid back	-
M28G2	close front	mid back	mid back
F21C2	close back	mid back	open-mid back
F16E2	close front	mid back	-
F17G2	close front	mid back	close-mid back
M29C2	close back	close-mid back	mid back
F15E2	close front	close-mid back	close-mid back
F24C2	close back	open-mid back	open-mid back
F18G2	close back	open-mid back	open-mid back
M21E2	close back	open-mid back	mid back
F22C2	close back	open-mid back	open back
M25G2	close back	close back	open-mid back
M27G2	close front	close back	close-mid back
F20C2	close front	close-mid back	close-mid back
M20E2	close central	close-mid back	open-mid back
F23C2	close back	mid back	mid back

Table 27. Overview of the position of GOOSE, LOT, and CLOTH for the Received Pronunciation speakers.

Speaker	TRAP	BATH	START
M22E2	open front	open back	open back
M23E2	open front	open back	open back
M24E2	open front	open back	open back
M26G2	open front	open back	open back
M30C2	open front	open back	open back
F19G2	open back	open back	-
M28G2	open front	open back	open-mid back
F21C2	open front	open back	open central
F16E2	open-mid front	open-mid back	-
F17G2	open front	open-mid back	open-mid back
M29C2	open front	open front	-
F15E2	open front	open central	open-mid back
F24C2	open front	open central	open back
F18G2	open front	open central	open central
M21E2	open front	-	open-mid back
F22C2	open central	open back	open back
M25G2	open front	mid back	close back
M27G2	open front	mid back	open-mid back
F20C2	open front	-	open back
M20E2	open central	open-mid central	mid back
F23C2	open front	open-mid back	mid front

Table 28. Overview of the position of TRAP, BATH, and START for the Received Pronunciation speakers.

With six different vowel features, each with four–six different pronunciation variants, under investigation, finding consistent pronunciation patterns across all six vowels is not to be expected. Thus, interesting patterns across two to three vowels are presented here. The first five speakers presented in Table 27 and Table 28 above—M22E2, M23E2, M24E2, M26G2, and M30C2—have an open back vowel in both BATH and START, and an open front vowel in TRAP. This group is split in their pronunciation of GOOSE: speakers M26G2 and M30C2 display GOOSE fronting with a close front position in this vowel, while the remaining three speakers have a close back vowel in GOOSE.

Eight speakers have the same articulatory features for LOT and CLOTH: M22E2, M24E2, M28G2, F15E2, F18G2, F24C2, F20C2, and F23C2. The pronunciation of LOT and CLOTH for these speakers includes open-mid back, mid back, and close-mid back. TRAP is an open front vowel for all these speakers, and GOOSE is split between a close front or close-mid front vowel and a close back vowel for this speaker group.

To illustrate these results and the analysis further, three speakers, speakers F18G2, M24E2, and M28G2 have been selected for more detailed analysis. These speakers display different realisations and combinations across the six vowels. The individual vowel plots of these speakers are represented in Figure 14 – Figure 16 below and their individual realisations of the mean formant values for GOOSE, LOT, CLOTH, TRAP, BATH, and START are discussed.

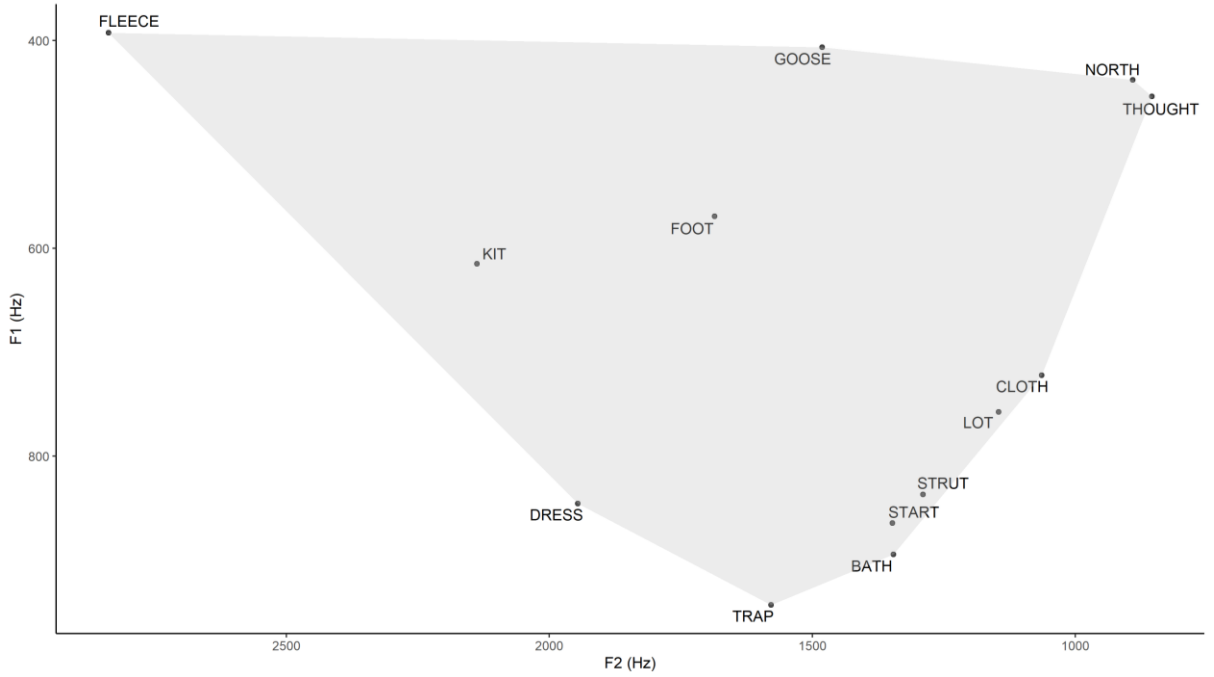


Figure 14. Vowel plot of mean formant values for speaker F18G2.

Figure 14 displays speaker F18G2's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels GOOSE, LOT, CLOTH, TRAP, BATH, and START are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker F18G2 has a close back vowel in GOOSE, that is situated between fully back and central. LOT and CLOTH are open-mid back vowels for this speaker. TRAP is an open front

vowel for this speaker and the vowel with the highest F1 value. TRAP is thus the most open vowel in this speaker's vowel inventory. BATH and START are situated close together. They are both open central vowels for this speaker.

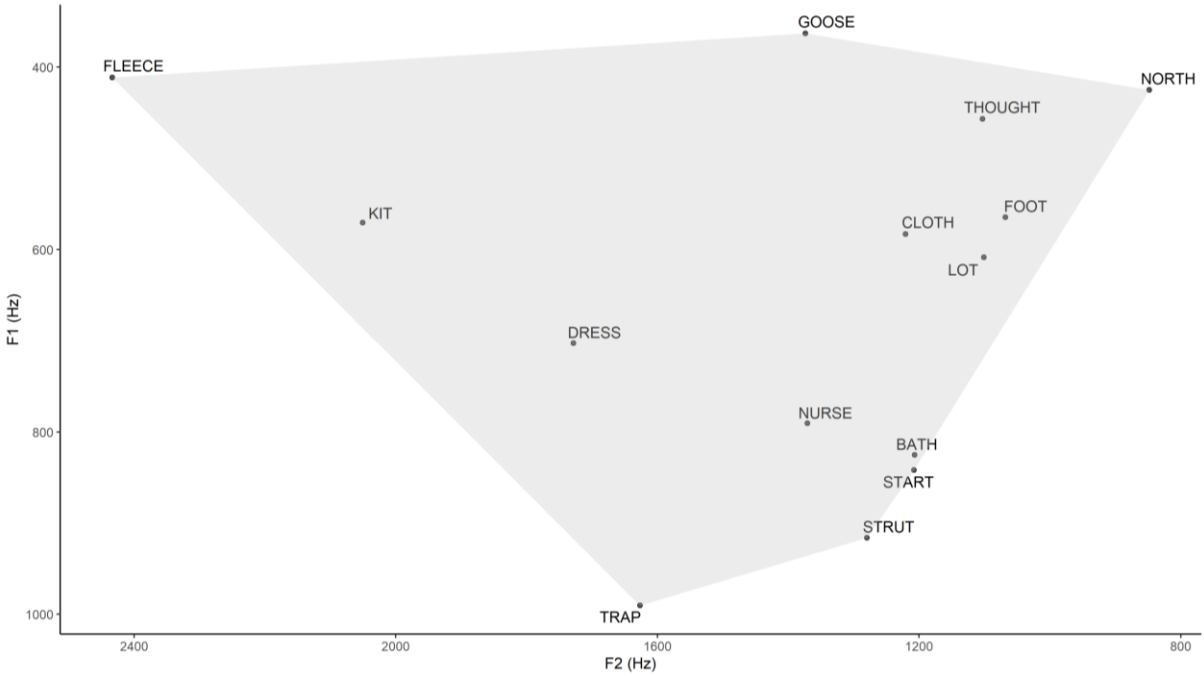


Figure 15. Vowel plot of mean formant values for speaker M24E2.

Figure 15 displays speaker M24E2's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels GOOSE, LOT, CLOTH, TRAP, BATH, and START are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker M24E2 has a close back vowel in GOOSE, that is situated between fully back and central. LOT and CLOTH are close-mid back vowels for this speaker, with CLOTH more fronted than LOT. TRAP is an open front vowel for this speaker and the vowel with the highest F1 value. TRAP is thus the most open vowel in this speaker's vowel inventory. BATH and START are situated close together. They are both open back vowels for this speaker.

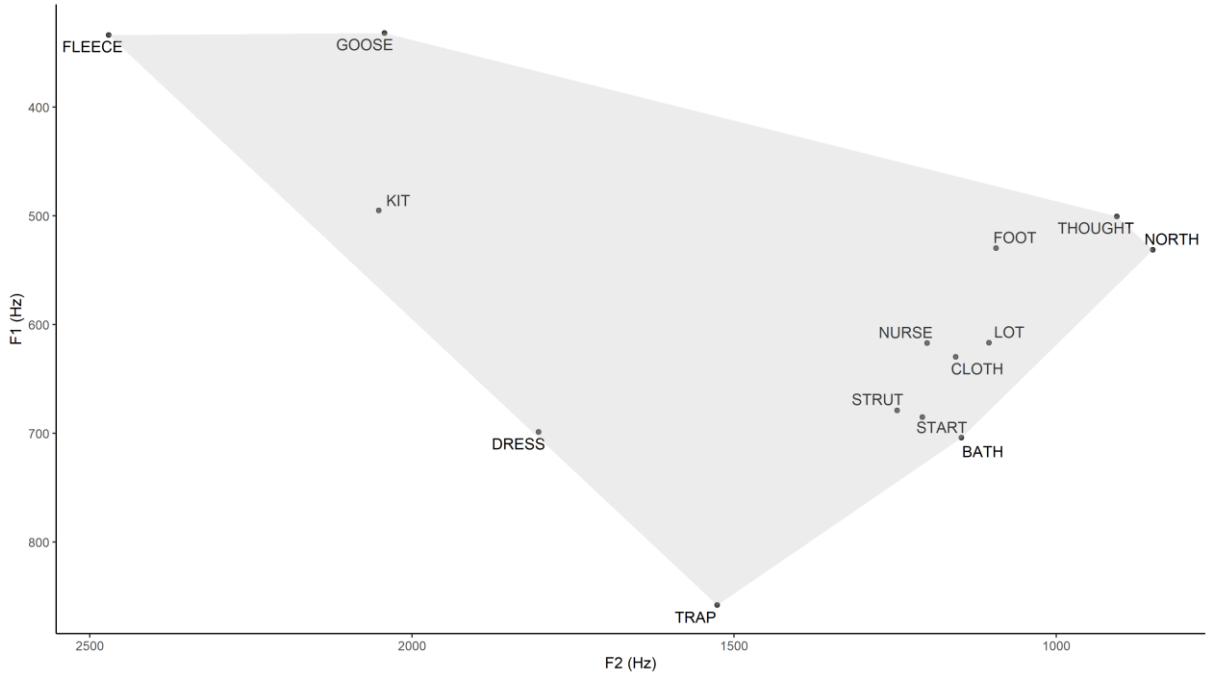


Figure 16. Vowel plot of mean formant values for speaker M28G2.

Figure 16 displays speaker M28G2's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels GOOSE, LOT, CLOTH, TRAP, BATH, and START are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker M28G2 has a close front vowel in GOOSE, that is closer than but as front as KIT. LOT and CLOTH are mid back vowels for this speaker, with CLOTH slightly more fronted than LOT. TRAP is an open front vowel for this speaker and the vowel with the highest F1 value. TRAP is by far the most open vowel in this speaker's vowel inventory. BATH and START are situated close together, differ however, slightly in openness. START is categorised as an open-mid back vowel, while bath is categorised as an open back vowel. While BATH is categorised as open back, this vowel is close to the open-mid back area.

The vowel plots presented in Figure 14 – Figure 16 above depict the individual vowel spaces of speakers F18G2, M24E2, and M28G2. They serve to illustrate the articulatory categorisation of the acoustic analyses presented for the individual speakers in Table 27 and Table 28 above exemplified by these three speakers. These results have demonstrated the variability in vowel realisations within this corpus data. While the grouped results for the textbook groups and all speakers as a cohort suggest a more homogeneous depiction of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START in the textbook audio materials,

the analyses of the individual speakers display a higher degree of variability in the realisation of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START.

This section presented the results of the acoustic analyses of the Received Pronunciation corpus data to illustrate how Received Pronunciation is depicted in English textbooks used at German secondary schools in NRW. These findings are expanded on and interpreted in the following section.

4.2.2 Synthesis—the Received Pronunciation Corpus Data and Received Pronunciation in Linguistic Literature

This section expands on the results presented in Section 4.2.1 and offers a detailed analysis of the representation of Received Pronunciation in English textbooks used at German secondary schools in NRW. The articulatory categorisations of the grouped and the speaker-individual results are compared to the linguistic description of Received Pronunciation presented in Section 2.5 and summarised in Section 2.8. This comparison facilitates a detailed interpretation of the representation of Received Pronunciation in these textbook materials.

The features of interest in this sub-corpus are the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. The following paragraph summarises the linguistic description of these features as presented in Section 2.5.

GOOSE is either a close back vowel in Received Pronunciation (Roach 2004: 242; Upton 2008: 241) or displays considerable fronting to a close central or even close front position (Cruttenden 2014: 84; Harrington, Kleber, and Reubold 2011: 151; Jansen and Mompean 2023: 55). LOT and CLOTH are both realised with the same open back vowel in Received Pronunciation. This vowel is located between fully open and open-mid (Roach 2004: 242; Upton 2008: 241; Wells 1982a: 130). TRAP is an open front vowel (Roach 2004: 243). BATH and START are either merged and realised with an open back vowel (Roach 2004: 242), or BATH is slightly fronted to an open central position (Upton 2008: 244).

These descriptions of the phonetic realisations of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START are used to contextualise and interpret the results presented in Section 4.2.1. First, the articulatory categorisation of the results regarding vowel height and tongue advancement is compared to the theory presented above. This facilitates an interpretation of the realisations of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and

START of the grouped and speaker-individual results regarding the representation of Received Pronunciation in the corpus data.

Table 29 and Table 30 below provide an overview of pronunciation patterns consistent with Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START for the speaker groups from the textbooks *English G Access 2*, *Green Line 2*, *and Camden Town 2*, as well as for all speakers as a group based on the articulatory classifications presented in Table 19 (page 69) and Table 20 (page 70). These articulatory classifications are matched to the descriptions of Received Pronunciation presented above. Articulatory classifications of the corpus data that match Received Pronunciation are indicated with RP (for Received Pronunciation) in Table 29 and Table 30 below. Three dashes (---) indicate that an articulatory classification does not match Received Pronunciation. Overall, the speaker groups display Received Pronunciation consistently for GOOSE, TRAP, and partially for BATH, but not for LOT, and CLOTH. Only the speaker group from *Camden Town 2* displays a pronunciation consistent with Received Pronunciation in START.

Speaker group	GOOSE	LOT	CLOTH
EGA2	RP	---	---
GL2	RP	---	---
CT2	RP	---	---
All speakers	**RP**	**---**	**---**

Table 29. Overview of pronunciation patterns consistent with Received Pronunciation in the monophthongs GOOSE, LOT, and CLOTH for the different speaker groups based on articulatory descriptions.

Speaker group	TRAP	BATH	START
EGA2	RP	RP	---
GL2	RP	---	---
CT2	RP	RP	RP
All speakers	**RP**	**RP**	RP

Table 30. Overview of pronunciation patterns consistent with Received Pronunciation in the monophthongs TRAP, BATH, and START for the different speaker groups based on articulatory descriptions.

For all speaker groups, GOOSE matches Received Pronunciation. For the speaker groups from *English G Access 2*, *Green Line 2*, and all speakers as a group, GOOSE is a close central vowel, and portrays GOOSE fronting commonly found in Received Pronunciation.

The speaker group from *Camden Town 2* has a close back realisation in GOOSE. TRAP is an open front vowel for all speaker groups and thus represents a feature of Received Pronunciation. BATH only displays Received Pronunciation characteristics for the textbook speaker groups from *English G Access 2*, and *Camden Town 2*, as well as for all speakers as a group. For all speakers and the speaker group from *English G Access 2*, BATH is an open back vowel. BATH is an open central vowel for the speaker group from *Camden Town 2*. The speaker group from *Green Line 2* had an open-mid back realisation in BATH, which is a pronunciation pattern inconsistent with Received Pronunciation. All speaker groups display pronunciation patterns inconsistent with Received Pronunciation in LOT, CLOTH. The speaker groups from *English G Access 2* and *Green Line 2* display pronunciation patterns inconsistent with Received Pronunciation in START. The speaker group from *Camden Town* 2 and all speakers as a group have an open back vowel in start, which matches Received Pronunciation.

A closer look at the individual speakers is necessary to analyse the depiction of Received Pronunciation in these textbooks and provide a complete picture of how these textbooks portray Received Pronunciation. Each individual speaker's averaged vowel pronunciations of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START were classified into the articulatory categories of vowel height and tongue advancement in Section 4.2.1. These articulatory classifications are matched to the descriptions of Received Pronunciation presented above. Articulatory classifications that match Received Pronunciation are indicated with RP (for Received Pronunciation) in Table 31 and Table 32 below. Three dashes (---) indicate that an articulatory classification does not match Received Pronunciation and the letter 'x' indicates that a vowel is not present in that speaker's data set.

Speaker	GOOSE	LOT	CLOTH
M22E2	RP	---	---
M23E2	RP	---	---
M24E2	RP	---	---
M26G2	RP	---	---
M30C2	RP	---	---
F19G2	RP	---	x
M28G2	RP	---	---
F21C2	RP	---	---
F16E2	RP	---	x
F17G2	RP	---	---
M29C2	RP	---	---
F15E2	RP	---	---
F24C2	RP	---	---
F18G2	RP	---	---
M21E2	RP	---	---
F22C2	RP	---	RP
M25G2	RP	---	---
M27G2	RP	---	---
F20C2	RP	---	---
M20E2	RP	---	---
F23C2	RP	---	---

Table 31. Overview of pronunciation patterns consistent with Received Pronunciation in GOOSE, LOT, and CLOTH across the individual speakers.

Speaker	TRAP	BATH	START
M22E2	RP	RP	RP
M23E2	RP	RP	RP
M24E2	RP	RP	RP
M26G2	RP	RP	RP
M30C2	RP	RP	RP
F19G2	---	RP	x
M28G2	RP	RP	---
F21C2	RP	RP	---
F16E2	---	---	x
F17G2	RP	---	---
M29C2	RP	---	x
F15E2	RP	RP	---
F24C2	RP	RP	RP
F18G2	RP	RP	---
M21E2	RP	x	---
F22C2	---	RP	RP
M25G2	RP	---	---
M27G2	RP	---	---
F20C2	RP	x	RP
M20E2	---	---	---
F23C2	RP	---	---

Table 32. Overview of pronunciation patterns consistent with Received Pronunciation in TRAP, BATH, and START across the individual speakers.

All speakers' pronunciation of the monophthong LOT and 18 of the 19 speakers' pronunciation of CLOTH is inconsistent with Received Pronunciation. While an open back vowel would be expected for these two vowels in Received Pronunciation, all speakers (except speaker F24C2 in CLOTH) display a less open vowel in both lexical sets. LOT is articulated as a close back, close-mid back, mid back, or open-mid back vowel by these speakers. These pronunciation patterns are all inconsistent with Received Pronunciation. CLOTH is a back vowel for all speakers except one—speaker M30C2, who has an open front variant in CLOTH. The remaining speakers range from a close-mid back pronunciation to an open-mid back pronunciation in this vowel. These pronunciation patterns are also inconsistent with

Received Pronunciation. Only speaker F24C2 has an open back pronunciation of CLOTH, which is consistent with Received Pronunciation.

Six speakers display pronunciation patterns consistent with Received Pronunciation in the remaining four vowels: M22E2, M23E2, M24E2, M26G2, M30C2, and F24C2. Five of these speakers (M22E2, M23E2, M24E2, M26G2, and M30C2) have an open back vowel in BATH and START and an open front vowel in TRAP. Speakers M22E2, M23E2, M24E2, and F24C2 have a close back vowel in GOOSE, while the other two speakers, M26G2 and M30C2 display GOOSE fronting with a close front pronunciation in this vowel.

This section offered a detailed analysis of the representation of Received Pronunciation in English textbooks used at German secondary schools in NRW, based on the analysis of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START. The analysis of the data at the superordinate level of the textbook speaker groups, and also the analysis at the speaker-individual level revealed interesting insights into the representation of Received Pronunciation in the textbook audio materials. The interpretation of the acoustic vowel analyses revealed that the corpus data matches the linguistic description of Received Pronunciation in some, but not all variables.

Section 4.3 presents and interprets the findings from the analyses of the General American sub-corpus following a similar structure.

4.3 General American

This section investigates how General American, and specifically the monophthongs BATH, LOT, and CLOTH, as well as rhoticity in the vowels NEAR, SQUARE, and CURE, are represented in the German textbooks of English: *English G Access 4*, *Green Line 4*, and *Camden Town 4*. The following sections present the results of the analysis. Section 4.3.1 illustrates the results of the acoustic analyses of the monophthongs BATH, LOT, and CLOTH to depict the General American accent of the textbooks. The section moves from a general overview of the pronunciation of the entire speaker group to more fine-grained analyses of the two textbook speaker groups and the individual speakers. Section 4.3.2 presents the state of rhoticity in the General American sub-corpus. The section provides a general overview of the pronunciation of the entire speaker group and detailed analyses of the three textbook speaker groups and individual speakers. This analysis provides a comprehensive overview of how General American is depicted in German textbooks of English.

Section 4.3.3 expands on these findings: a comparison with the linguistic description of General American as presented in Section 2.6 facilitates an interpretation of these findings. The results of the acoustic vowel analyses and the auditory rhoticity analyses are evaluated against the description of General American.

4.3.1 The Vowels in the General American Sub-Corpus

Speech samples from 20 textbook speakers have been analysed to investigate the representation of General American as exemplified by the monophthongs BATH, LOT, and CLOTH. Seven speakers were analysed from the textbook *English G Access 4*, seven speakers were analysed from the textbook *Green Line 4*, and six speakers were analysed from the textbook *Camden Town 4*. A combined total of 502 tokens for BATH, LOT, and CLOTH were analysed across these 20 speakers.

To get a first impression of the representation of General American in German textbooks, four speaker groups were created: a group of all speakers, all speakers from *English G Access 4* (EGA4), all speakers from *Green Line 4* (GL4), and all speakers from *Camden Town 4* (CT4). For each group, pooled averages for F1 and F2 of the vowel targets in BATH, LOT, and CLOTH were calculated based on the groups' speakers mean raw F1 and F2 values in these vowels.

Speaker groups		**BATH**				**LOT**				**CLOTH**		
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
EGA4	F1	891.8	156.4	7	F1	814.4	88.1	7	F1	711.3	162.2	7
	F2	1698.8	153.7		F2	1270.8	161.2		F2	1151.9	184.8	
GL4	F1	858.0	170.6	6	F1	821.1	137.6	7	F1	713.5	140.2	6
	F2	1818.6	189.0		F2	1333.8	168.7		F2	1237.7	210.4	
CT4	F1	931.1	111.9	5	F1	878.5	94.2	6	F1	786.2	42.1	5
	F2	1819.8	156.7		F2	1371.4	175.0		F2	1161.0	79.3	
All	**F1**	**891.4**	**144.9**	**18**	**F1**	**836.0**	**107.7**	**20**	**F1**	**732.8**	**129.0**	**18**
	F2	**1772.4**	**168.1**		**F2**	**1323.0**	**164.4**		**F2**	**1183.1**	**167.8**	

Table 33. Pooled mean formant values and standard deviations for BATH, LOT, and CLOTH for the speaker groups from English G Access 4, Green Line 4, Camden Town 4, and all speakers.

Table 33 displays the pooled averages in Hz of the F1 and F2 values for the vowel targets in the vowels BATH, LOT, and CLOTH and the standard deviations (SD) per speaker group per vowel. Both the mean F1 and F2 values, and the SD are rounded to one decimal place. The token number (n) indicates the number of speakers per group, as these mean values were calculated based on the speaker-individual mean F1 and F2 values per vowel. The pooled averages of F1 and F2 for the monophthongs BATH, LOT, and CLOTH are visualised in Figure 17–Figure 20 below. Figure 17 displays the mean vowel plot for the entire speaker group, Figure 18 displays the mean vowel plot for the speakers from *English G Access 4*, Figure 19 displays the mean vowel plot for the speakers from *Green Line 4*, and Figure 20 displays the mean vowel plot for the speakers from *Camden Town 4*.

The pooled mean vowel realisations of all monophthongs for each speaker group are used as the reference points in each plot. Thus, the vowel plot for the entire speaker group also displays the pooled mean vowel realisations of the remaining monophthongs for this speaker group. The vowel spaces for the speaker groups from *English G Access 4*, *Green Line 4*, and *Camden Town 4* are created from the pooled mean vowel realisations of the monophthongs for these respective speaker groups. This procedure provides a general first impression of the General American accent in German textbooks of English. These results are interpreted regarding their representation of General American, i.e. whether these features represent a General American pronunciation or not, in Section 4.3.3.

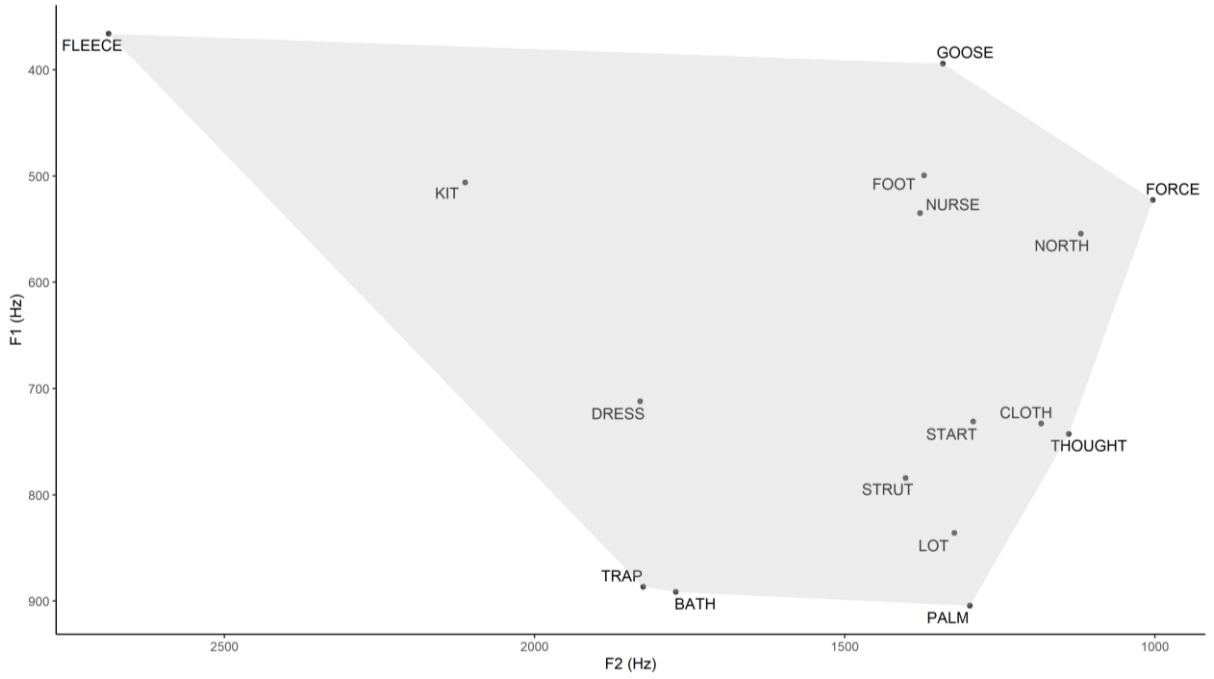


Figure 17. Averaged realisations of the monophthongs for all 20 speakers from the General American sub-corpus.

Figure 17 displays an averaged vowel plot of all speakers' individual averaged F1 and F2 values in hertz. These mean values were calculated from each speaker's averaged F1 and F2 values in hertz for each vowel.

The vowels of interest here are the monophthongs BATH, LOT, and CLOTH. The categorisation of these vowels according to the articulatory features of vowel height and tongue advancement is based on the 'segmented' vowel plots in Appendix 2. These vowel plots include a segmentation into the articulatory categories of front, central, and back on the horizontal (F2) axis and close, close-mid, mid, open-mid, and open on the vertical (F1) axis.

For this speaker group, BATH is an open front vowel with a quality close to TRAP. LOT and CLOTH display a visible split for this speaker group. LOT is an open back vowel which is slightly closer than PALM, but more open and slightly retracted compared to STRUT. CLOTH is an open-mid back vowel with a vowel quality close to THOUGHT.

While considering all speakers of the sub-corpus as one cohort provides a comprehensive overview of the depiction of General American in English textbooks used in NRW, examining the speaker groups from each textbook individually offers insights into unique pronunciation patterns. Therefore, the subsequent paragraphs delve into the results for each textbook speaker group, starting with *English G Access 4*.

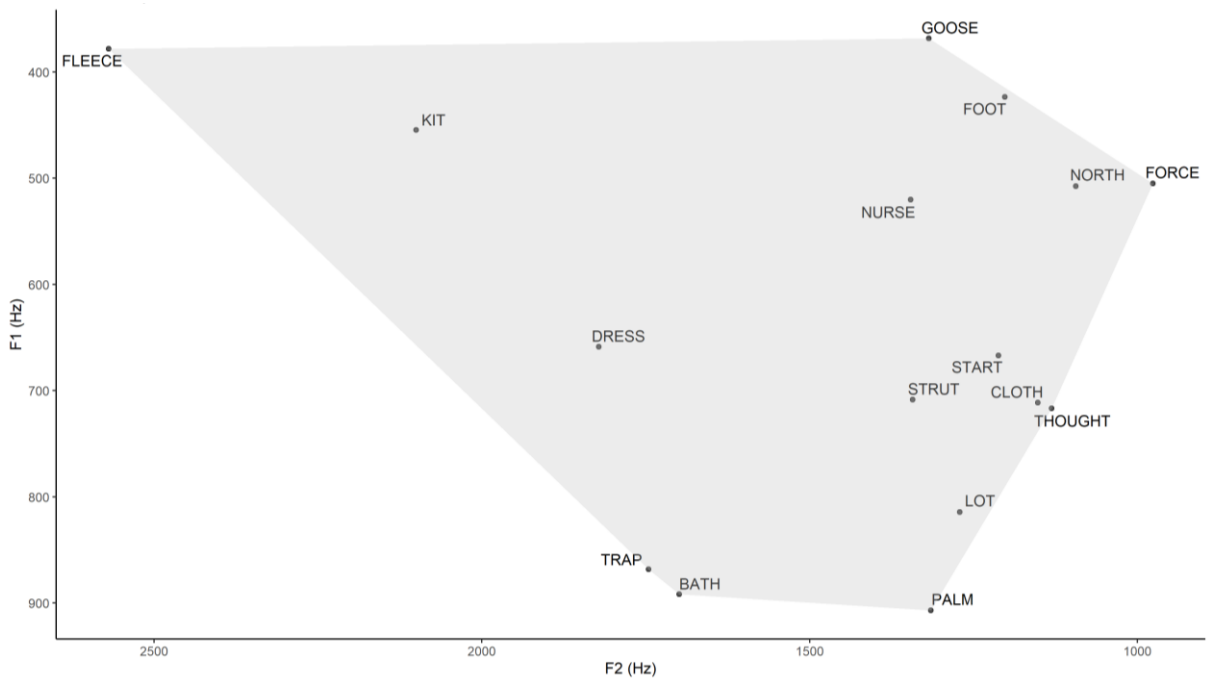


Figure 18. Averaged realisations of the monophthongs for the seven English G Access 4 speakers.

Figure 18 shows the averaged vowel plot for all speakers from the textbook *English G Access 4*. These mean values were calculated from each speaker's averaged F1 and F2 values in hertz for each vowel.

This vowel plot strongly resembles the one for the entire speaker group in form and in the realisation of the monophthongs BATH, LOT, and CLOTH. This speaker group shows an open front pronunciation in BATH with a quality close to but slightly more open than TRAP. LOT is an open back vowel for this speaker group, which is less open than PALM, but more open than the open-mid back vowel THOUGHT. CLOTH is an open-mid back vowel for this speaker group, with a quality close to THOUGHT.

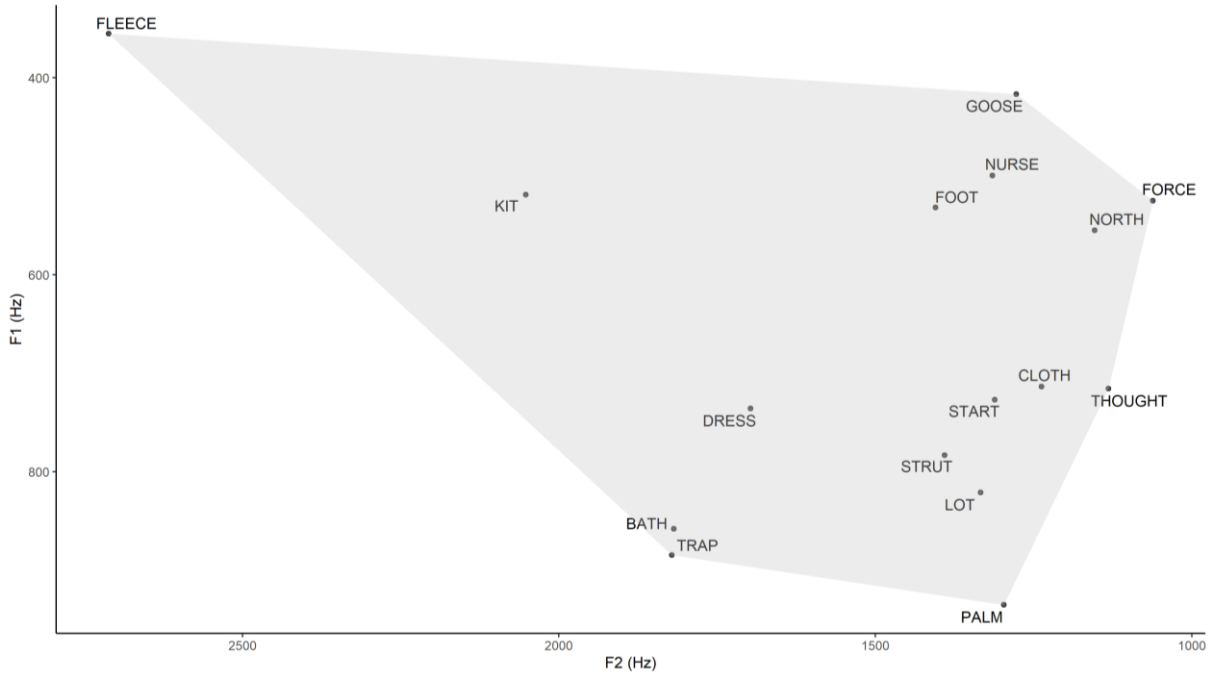


Figure 19. Averaged realisations of the monophthongs for the seven Green Line 4 speakers.

Figure 19 shows the averaged vowel plot for all seven speakers from the textbook *Green Line 4.* The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook.

BATH is an open front vowel for this speaker group. However, this speaker group shows BATH as slightly more closed than TRAP. LOT is an open back vowel for this speaker group. It is more open than STRUT, but closer than PALM. It is situated closer to STRUT than to PALM, but more retracted compared to STRUT. CLOTH is an open-mid back vowel for this speaker group. With regard to vowel height, CLOTH has a similar vowel height to THOUGHT. CLOTH is, however, slightly more fronted than THOUGHT.

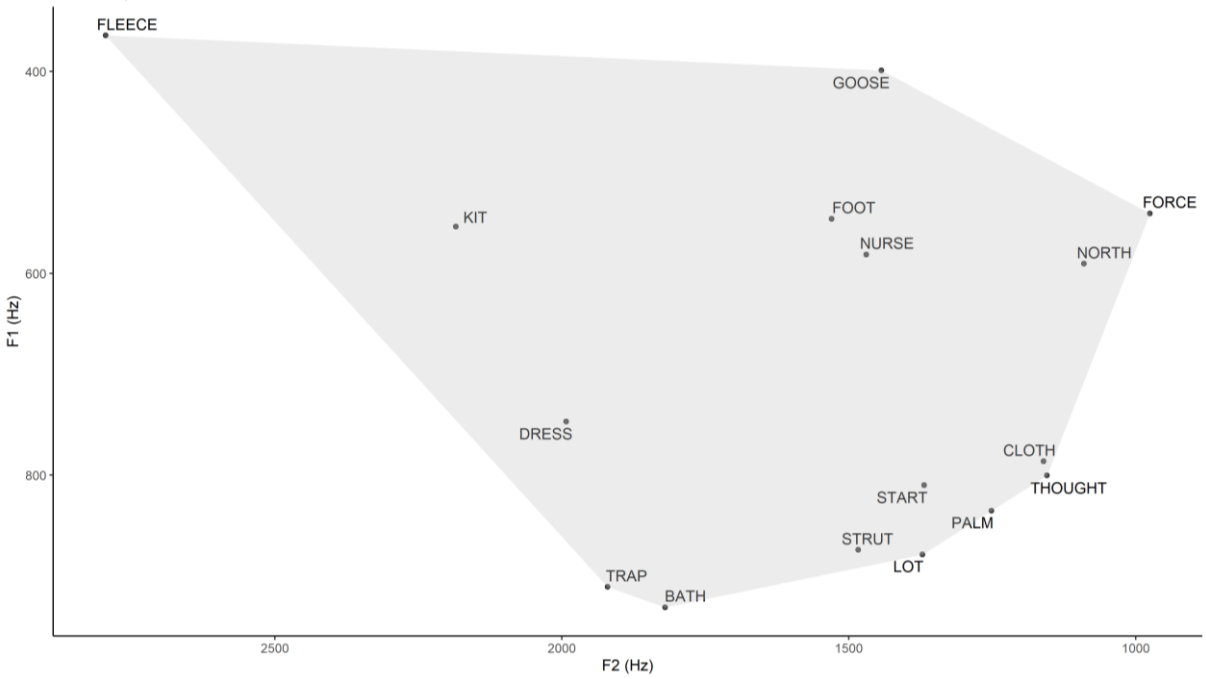


Figure 20. Averaged realisations of the monophthongs for the six Camden Town 4 speakers.

Figure 20 shows the averaged vowel plot for all six speakers from the textbook *Camden Town 4.* The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook.

For these speakers, all three vowels—BATH, LOT, and CLOTH—are open vowels. BATH is an open front vowel that is slightly more open and retracted compared to TRAP. LOT is an open back vowel with a more open quality compared to STRUT and PALM. CLOTH is also an open back vowel with a quality close to THOUGHT.

Speaker group	BATH	LOT	CLOTH
EGA4	open front	open back	open-mid back
GL4	open front	open back	open-mid back
CT4	open front	open back	open back
All speakers	**open front**	**open back**	**open back**

Table 34. Overview of articulatory categorisation of BATH, LOT, and CLOTH for the speaker groups from English G Access 4, Green Line 4, Camden Town 4, and all speakers.

Table 34 summarises the results from the acoustic analysis and categorisation into the articulatory features of vowel height and tongue advancement for the three textbook

speaker groups, and all speakers as a group. Overall, the grouped vowel plots suggest a rather homogeneous realisation of the monophthongs BATH, LOT, and CLOTH across the three textbooks included in this sub-corpus. For the three textbook groups and the group of all speakers, BATH has an open front vowel quality. LOT is an open back vowel across the three textbook groups, and for the entire speaker group. CLOTH has an open-mid back vowel quality for the entire speaker group, and for the textbook groups from *English G Access 4* and *Green Line 4*. For the speaker group from *Camden Town 4*, this vowel is slightly more open resulting in an open back quality.

These analyses of the individual speaker groups highlight overall trends within the dataset. A detailed examination of the pronunciation patterns at the speaker level offers more nuanced insights into the depiction of General American in the English textbooks used at German secondary schools. The analyses at the speaker level are based on each speaker's individual vowel space (see Appendix 1), which are built from each speaker's mean F1 and F2 values of their monophthongs. Tables 4-6 (Appendix 5) provides an overview of the speakers' mean F1 and F2 values in hertz and the standard deviation of the vowels under discussion: BATH, LOT, and CLOTH. The pronunciation of the vowels under discussion is evaluated based on their position in the vowel space. BATH, LOT, and CLOTH are categorised according to the articulatory features of vowel height and tongue advancement. The 'segmented' vowel plots for each speaker used for this categorisation can be found in Appendix 2. Table 35 – Table 37 below indicate the different pronunciation variants of BATH, LOT, and CLOTH, and indicate how many speakers produce these variants.

	Position of BATH	
	Open-mid front	Open front
Number of speakers	2	16

Table 35. Realisation of BATH based on each speaker's mean formant values.

Table 35 provides an overview of the different realisations of the BATH vowel and how many speakers in the General American sub-corpus produce such a realisation based on each speaker's mean F1 and F2 formant values for this vowel. BATH has two different realisations in the General American sub-corpus: it can be either an open-mid front vowel or an open front vowel. Eighteen of the 20 speakers produce tokens of BATH. Overall, the speakers are fairly consistent in their pronunciation of BATH, which varies slightly in vowel height

but not in vowel frontness. For all 18 speakers, BATH is a front vowel. Sixteen of the 18 speakers have an open front realisation in this monophthong. The remaining two speakers have an open-mid front realisation of BATH.

	Position of LOT			
	Open front	Open central	Open back	Open-mid back
Number of speakers	1	4	12	3

Table 36. Realisation of LOT based on each speaker's mean formant values.

Table 36 provides an overview of the different pronunciation variants in LOT and how many speakers in the General American sub-corpus produce such variants based on each speaker's mean F1 and F2 formant values for this vowel. The monophthong LOT has four different pronunciation variants in the General American sub-corpus: open front, open central, open back, and open-mid back. The vowel target in LOT varies in vowel backness and also in vowel height for the 20 speakers of the General American sub-corpus. LOT is produced as an open front, open central, open back, or open-mid back vowel in this data set. Just over half of the speakers—12 of the 20—have an open back realisation of LOT. Five speakers produce a fronter variant in this vowel, with four speakers producing an open central vowel and one speaker even producing an open front realisation of LOT. Three speakers differ in vowel height rather than frontness in this vowel: they produce an open-mid back vowel for LOT.

	Position of CLOTH				
	Open central	Open back	Open-mid back	Mid back	Close back
Number of speakers	1	6	5	5	1

Table 37. Realisation of CLOTH based on each speaker's mean formant values.

Table 37 provides an overview of the different pronunciation variants of CLOTH in the General American sub-corpus and how many of the speakers produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. CLOTH varies considerably in vowel height and slightly in vowel frontness, resulting in the pronunciation

variants open central, open back, open-mid back, mid back and even close back. Eighteen of the 20 speakers in this sub-corpus produce tokens of CLOTH. For all except one of these 18 speakers, CLOTH is a back vowel. Six speakers have an open back vowel target in CLOTH. Five speakers each have an open-mid back or mid back vowel target in this monophthong. One speaker produces a close back vowel for CLOTH, and the last speaker has an open central realisation in this monophthong.

Overall, BATH, LOT, and CLOTH display a certain variability in vowel quality in this corpus data. A closer look at how the individual textbook speakers realise these three vowels allows for a more detailed analysis of how General American is presented in the textbook materials. Table 38 below provides an overview of the pronunciation variants in BATH, LOT, and CLOTH across the individual speakers. The sequence of speakers in this table is adapted from the usual sequence introduced in Section 3.2.4 to showcase interesting pronunciation patterns across the individual speakers.

As established in Table 35, an open front realisation is the most common realisation for BATH. Of the 16 speakers that have an open front realisation for BATH, nine have an open back realisation in LOT, which is the most common realisation in this vowel in this sub-corpus (cf. Table 36). Of these nine speakers with an open front realisation in BATH and an open back realisation in LOT, four speakers have an open back realisation in CLOTH (F6E4, F13C4, M18C4, and M19C4), and three speakers have an open-mid back realisation in CLOTH (F8G4, F9G4, and F11C4). These two speaker groups only differ in their pronunciation of CLOTH as either an open back vowel in the former group or as an open-mid back vowel in the latter group of speakers. They comprise the two biggest clusters of speakers with the same realisations of the three monophthongs BATH, LOT, and CLOTH.

Speaker	BATH	LOT	CLOTH
F6E4	open front	open back	open back
F13C4	open front	open back	open back
M18C4	open front	open back	open back
M19C4	open front	open back	open back
F8G4	open front	open back	open-mid back
F9G4	open front	open back	open-mid back
F11C4	open front	open back	open-mid back
F5E4	open front	open back	mid back
F10G4	open front	open back	close back
F12C4	open front	open central	open-mid back
F7E4	open front	open central	mid back
M13E4	open front	open central	mid back
M10E4	open front	open-mid back	mid back
M12E4	open front	open-mid back	mid back
M15G4	open front	open-mid back	open-mid back
M14G4	open front	open front	open back
M11E4	open-mid front	open back	open central
M17G4	open-mid front	open back	-
M16G4	-	open back	open back
F14C4	-	open central	-

Table 38. Overview of the position of BATH, LOT, and CLOTH for the General American speakers.

To illustrate these results and the analysis further, three speakers, speakers F8G4, M13E4, and M11E4 have been selected for more detailed analysis, as they display different pronunciation patterns across the monophthongs BATH, LOT, and CLOTH. The individual vowel plots of these speakers are presented in Figure 21 – Figure 23 below and their individual realisations of the mean formant values in BATH, LOT, and CLOTH are discussed.

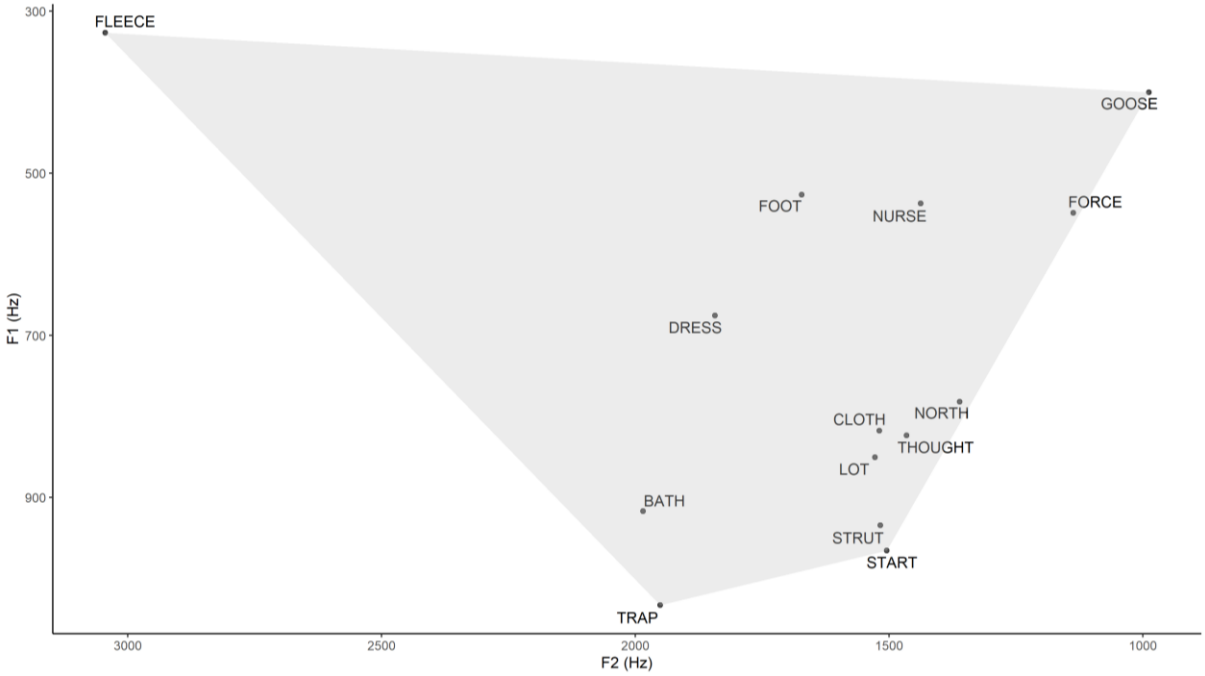


Figure 21. Averaged vowel plot for speaker F8G4.

Figure 21 displays speaker F8G4's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels BATH, LOT, and CLOTH are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker F8G4 has an open front realisation in BATH. For this speaker, BATH is as front as TRAP, but closer. LOT and CLOTH are both back vowels but vary slightly in vowel height: LOT is an open back vowel for this speaker, while CLOTH is an open-mid back vowel.

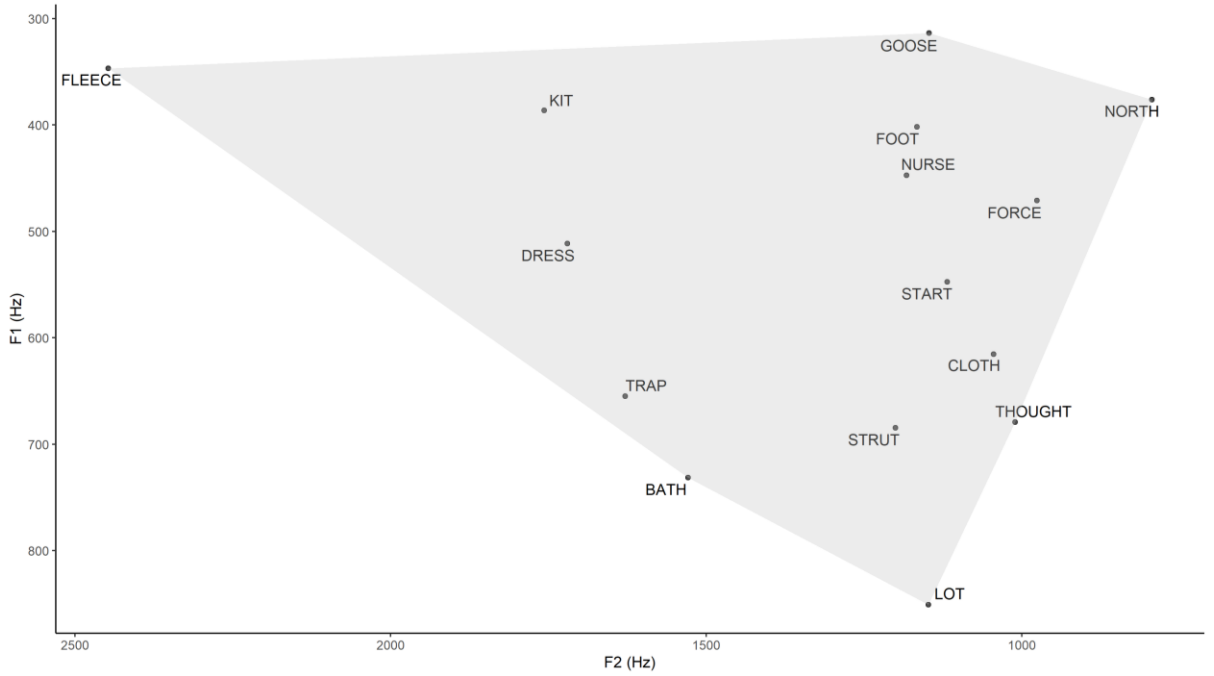


Figure 22. Averaged vowel plot for speaker M13E4.

Figure 22 displays speaker M13E4's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels BATH, LOT, and CLOTH are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker M13E4 has an open front realisation in BATH. BATH is as almost as front as TRAP, but more open than TRAP. LOT is an open central vowel for this speaker, and the vowel with the highest F1 value, so the most open vowel. CLOTH is a mid back vowel.

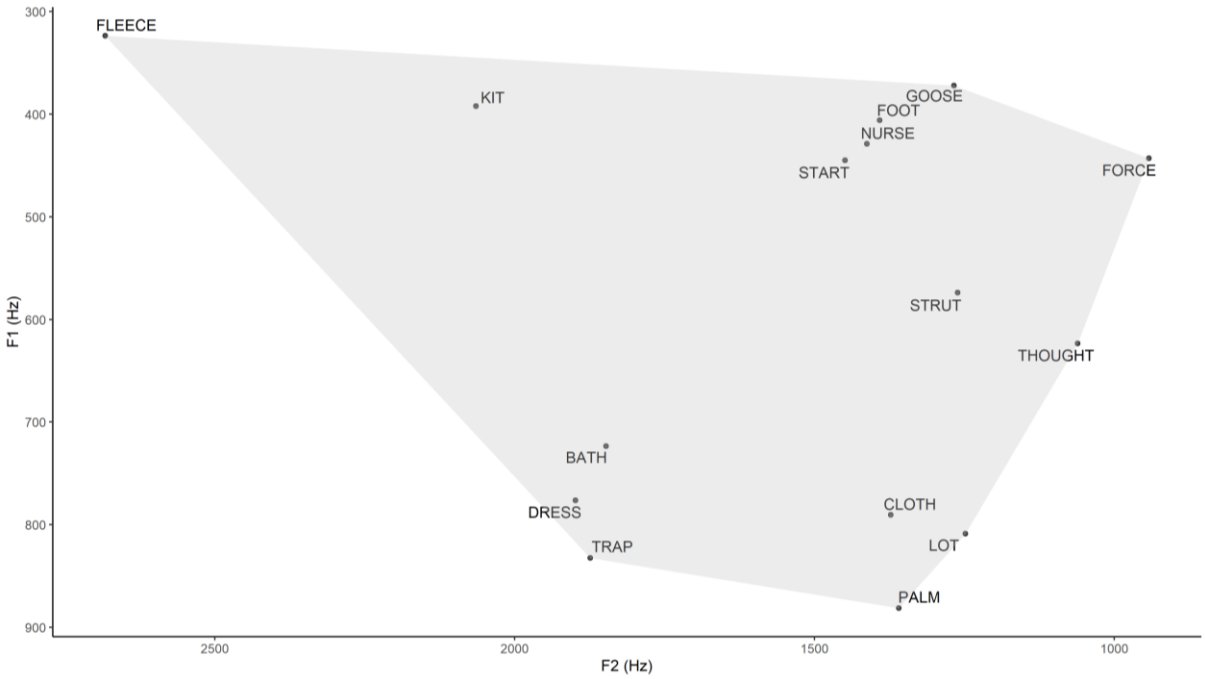


Figure 23. Averaged vowel plot for speaker M11E4.

Figure 23 displays speaker M11E4's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The vowels BATH, LOT, and CLOTH are of interest here. Based on their position in this vowel space, the vowel quality of these monophthongs can be described in the articulatory terms of tongue height and tongue advancement (see the segmented vowel plot in Appendix 2).

Speaker M11E4 has an open front realisation in BATH. BATH is as almost as front as TRAP, but less open than TRAP. LOT is an open back vowel for this speaker, and CLOTH is an open central vowel, with a similar F2 value than PALM.

The vowel plots presented in Figure 21 – Figure 23 above depict the individual vowel spaces of speakers F8G4, M13E4, and M11E4. They serve to illustrate the articulatory categorisation of the acoustic analyses presented for the individual speakers in Table 38 above exemplified by these three speakers. These results have demonstrated the variability in vowel realisations within this corpus data. While the grouped results for the textbook groups and all speakers as a cohort suggest a more homogeneous depiction of General American in the textbook audio materials, the analyses of the individual speakers display a higher degree of variability in the realisation of the monophthongs BATH, LOT, and CLOTH.

This section presented the results of the acoustic analyses of the General American corpus data to illustrate how the General American accent is depicted in English textbooks used at German secondary schools in NRW. Section 4.3.3 expands on and interprets the

findings. Section 4.3.2 introduces the results of the auditory analysis of rhoticity in the General American sub-corpus.

4.3.2 Rhoticity in the General American Sub-Corpus

Speech samples from 20 textbook speakers were analysed to investigate the representation of General American as exemplified by rhoticity in the lexical sets NEAR, SQUARE, and CURE. Seven speakers were analysed from the textbook *English G Access 4*, seven speakers were analysed from the textbook *Green Line 4*, and six speakers were analysed from the textbook *Camden Town 4*. A combined total of 212 tokens for rhoticity in the lexical sets NEAR, SQUARE, and CURE were analysed across these 20 speakers.

The tokens were marked for the phonological context of /r/. In the lexical sets NEAR, SQUARE, and CURE, /r/ can occur in seven different phonological contexts. In contexts (i) – (iv), /r/ occurs intervocalically: (i) within a word before a stressed syllable, (ii) within a word before an unstressed syllable, (iii) across a word boundary before a stressed syllable, and (iv) across a word boundary before an unstressed syllable. In contexts (v) – (vi), /r/ occurs preconsonantally: (v) within a word, and (vi) across a word boundary. In context (vii), /r/ occurs in word, phrase, or utterance final position. In these contexts, /r/ was marked as the last phoneme in the word with a definite pause or utterance break following.

Table 39 below displays the frequency of the different phonological contexts for the lexical sets NEAR, SQUARE, and CURE in the General American sub-corpus. Overall, the lexical sets NEAR and SQUARE occur with a higher frequency than the lexical set CURE: NEAR occurs 87 times in the sub-corpus, SQUARE 109 times, and CURE only 16 times.

Phonological context	NEAR	SQUARE	CURE	**Total**
(i) Intervocalically within a word before a stressed syllable	0	0	0	**0**
(ii) Intervocalically within a word before an unstressed syllable	5	2	1	**8**
(iii) Intervocalically across a word boundary before a stressed syllable	7	28	1	**36**
(iv) Intervocalically across a word boundary before an unstressed syllable	9	6	2	**17**
(v) Preconsonantally within a word	28	5	1	**34**
(vi) Preconsonantally across a word boundary	24	55	8	**87**
(vii) Word/phrase/utterance final	14	13	3	**30**
Total	**87**	**109**	**16**	**212**

Table 39. Overview of token numbers for NEAR, SQUARE, and CURE across the phonological contexts for rhoticity.

A closer look at the dispersion of the tokens across the different phonological contexts reveals that the phonological context (i) in which /r/ occurs intervocalically within a word before a stressed syllable does not appear in this data set. This context is therefore omitted from further discussion in this section. /r/ occurs in intervocalic position in 61 tokens: eight of these tokens have /r/ in intervocalic position within a word before an unstressed syllable. The remaining 53 instances are in intervocalic position across a word boundary: 36 instances are before a stressed syllable, and 17 are before an unstressed syllable. /r/ occurs in a preconsonantal position in 121 instances. In 34 of these tokens, /r/ occurs preconsonantally within a word, and in 87 tokens /r/ occurs preconsonantally across a word boundary. This makes context (vi) with /r/ in a preconsonantal position across a word boundary the most frequent phonological context in this data set.

To get a first impression of the state of rhoticity in General American as presented in German teaching materials, four speaker groups were created: a group of all speakers, all speakers from *English G Access 4*, all speakers from *Green Line 4*, and all speakers from *Camden Town 4*. For each group, the state of rhoticity marked by articulated /r/ in the different phonological context is presented in Table 40 – Table 43 below. The raw frequencies

of articulated-/r/ and not-articulated-/r/ in the different phonological contexts is presented. The relative frequency of articulated /r/ is also presented.

Phonological context	Articulated	Not Articulated	% Articulated /r/
(ii) Intervocalically within a word before an unstressed syllable	8	0	100%
(iii) Intervocalically across a word boundary before a stressed syllable	33	3	91.67%
(iv) Intervocalically across a word boundary before an unstressed syllable	17	0	100%
(v) Preconsonantally within a word	33	1	97.06%
(vi) Preconsonantally across a word boundary	68	19	78.16%
(vii) Word/phrase/utterance final	29	1	96.67%
Total	**188**	**24**	**88.68%**

Table 40. Number and percentage of articulated /r/ in different phonological contexts across all 20 speakers from the General American sub-corpus.

Table 40 displays the number and percentage of articulated /r/ in the different phonological contexts across all 20 speakers in the General American sub-corpus. Overall, 212 tokens from the lexical sets NEAR, SQUARE, and CURE are analysed regarding the articulation of post-vocalic /r/. /r/ is articulated in 88.68% of these tokens, accounting for 188 of the 212 tokens. Most instances of articulated post-vocalic /r/, 68 tokens of 188, occur in context (vi) with /r/ occurring preconsonantally across a word boundary. Within in this phonological context, /r/ was not articulated in 19 tokens. Thus, context (vi) only has a 78.16% articulation rate of post-vocalic /r/. Contexts (ii) and (iv), /r/ in intervocalic contexts before unstressed syllables within a word and across a word boundary, show full rhoticity with /r/ being articulated in 100% of the tokens. In the remaining contexts, /r/ is articulated in 91.67% in context (iii), 97.06% in context (v), and 78.16% in context (vi). Thus, /r/ in preconsonantal position across a word boundary displays the lowest rate of rhoticity.

While considering all speakers of the sub-corpus as one cohort provides a comprehensive overview of the depiction of rhoticity in General American in English textbooks used at German secondary schools, examining the speaker groups from each textbook

individually offers unique insights into pronunciation patterns. Table 41 – Table 43 below display the rates of rhoticity across the different phonological contexts for the individual textbook groups starting with *English G Access 4*.

Phonological context	Articulated	Not Articulated	% Articulated /r/
(ii) Intervocalically within a word before an unstressed syllable	3	0	100%
(iii) Intervocalically across a word boundary before a stressed syllable	7	0	100%
(iv) Intervocalically across a word boundary before an unstressed syllable	5	0	100%
(v) Preconsonantally within a word	6	0	100%
(vi) Preconsonantally across a word boundary	28	5	84.85%
(vii) Word/phrase/utterance final	10	0	100%
Total	**59**	**5**	**92.19%**

Table 41. Number and percentage of articulated /r/ in different phonological contexts across all 7 speakers from the textbook English G Access 4.

Table 41 displays the number and percentage of articulated /r/ in the different phonological contexts across the seven speakers from *English G Access 4.* This subset of the data consists of 64 tokens from the lexical sets NEAR, SQUARE, and CURE. Within this subset, /r/ is articulated in 92.19% of tokens. Thus, /r/ is not articulated in five out of 64 tokens. Taking the phonological contexts into account, /r/ is articulated consistently in five of the six different phonological contexts. /r/ is consistently articulated in all intervocalic positions either within a word (context (ii)) or across word boundaries before a stressed (context (iii)) or unstressed syllable (context (iv)). In preconsonantal position within a word, /r/ is articulated in 100% of tokens. In preconsonantal position across a word boundary, however, /r/ is only articulated in 84.85% of the tokens: in five of the 33 tokens, /r/ is not articulated. These tokens are examined more closely below when the rhoticity level of the individual speakers is presented. In word, phrase, or utterance final position, /r/ is consistently articulated in the *English G Access* subset of the data.

Phonological context	Articulated	Not Articulated	% Articulated /r/
(ii) Intervocalically within a word before an unstressed syllable	-	-	-
(iii) Intervocalically across a word boundary before a stressed syllable	16	0	100%
(iv) Intervocalically across a word boundary before an unstressed syllable	9	0	100%
(v) Preconsonantally within a word	13	0	100%
(vi) Preconsonantally across a word boundary	21	12	63.64%
(vii) Word/phrase/utterance final	13	0	100%
Total	**72**	**12**	**85.71%**

Table 42. Number and percentage of articulated /r/ in different phonological contexts across all 7 speakers from the textbook Green Line 4.

Table 42 displays the number and percentage of articulated /r/ in the different phonological contexts across the seven speakers from *Green Line 4.* This subset of the data consists of 84 tokens from the lexical sets NEAR, SQUARE, and CURE. Within this subset, /r/ is articulated in 85.71% of tokens. Thus, /r/ is not articulated in 12 out of 84 tokens. Taking the phonological contexts into account, /r/ is articulated consistently in four of the six different phonological contexts. /r/ does not occur intervocalically within a word before an unstressed syllable (context (ii)) in this subset of the data but is consistently articulated in the remaining intervocalic positions across word boundaries (contexts (iii) – (iv)). In preconsonantal position within a word, /r/ is also articulated in 100% of tokens. In preconsonantal position across a word boundary, however, /r/ is only articulated in 63.64% of the tokens: /r/ is not articulated in preconsonantal position across a word boundary in 12 of the 33 tokens. /r/ is consistently articulated in word, phrase, or utterance final position in the *Green Line* subset of the data.

Phonological context	Articulated	Not Articulated	% Articulated /r/
(ii) Intervocalically within a word before an unstressed syllable	5	0	100%
(iii) Intervocalically across a word boundary before a stressed syllable	10	3	76.92%
(iv) Intervocalically across a word boundary before an unstressed syllable	3	0	100%
(v) Preconsonantally within a word	14	1	93.33%
(vi) Preconsonantally across a word boundary	19	2	90.48%
(vii) Word/phrase/utterance final	6	1	85.71%
Total	**57**	**7**	**89.06%**

Table 43. Number and percentage of articulated /r/ in different phonological contexts across all 6 speakers from the textbook Camden Town 4.

Table 43 displays the number and percentage of articulated /r/ in the different phonological contexts across the six speakers from *Camden Town 4.* This subset of the data consists of 64 tokens from the lexical sets NEAR, SQUARE, and CURE. Within this subset, /r/ is articulated in 89.06% of tokens. Thus, /r/ is not articulated in seven out of 64 tokens. Taking the phonological contexts into account, /r/ is only articulated consistently in two of the six different phonological contexts. /r/ is consistently articulated in the phonological contexts (ii) and (iv): in intervocalic position before an unstressed syllable either within a word (ii) or across a word boundary (iv). /r/ is not articulated consistently in an intervocalic position across a word boundary before a stressed syllable (context (iii)). In this context, /r/ is articulated in only ten of 13 tokens. In the preconsonantal contexts (v) and (vi), /r/ is also articulated inconsistently, with one token in context (v) and two tokens in context (vi) showing no consonantal articulation of /r/. In word, phrase, or utterance final position, /r/ is articulated in six of seven tokens.

Table 44 below summarises the level of rhoticity per speaker group and provides an overview of the phonological contexts displaying full or variable rhoticity per speaker group in the General American sub-corpus. The speaker group from *English G Access 4* displays the highest level of rhoticity in the corpus data with 92.19% of articulated /r/ in the lexical sets NEAR, SQUARE, and CURE. The speaker group from *Green Line 4* has the lowest level of

rhoticity with only 85.71% of articulated /r/. The speaker groups from English G Access 4 and Green Line 4 only display variable rhoticity—so instances of not-articulated-/r/—in the phonological context (vi): preconsonantally across a word boundary. The speaker group from *Camden Town 4*, on the other hand, displays variable rhoticity in the phonological contexts (iii), (v), (vi), and (vii). This speaker group therefore displays variable rhoticity in intervocalic, preconsonantal, and word-final positions.

Speaker group	% articulated /r/	Phonological contexts with full rhoticity	Phonological contexts with variable rhoticity
EGA4	92.19%	(ii), (iii), (iv), (v), (vii)	(vi)
GL4	85.71%	(iii), (iv), (v), (vii)	(vi)
CT4	89.06%	(ii), (iv)	(iii), (v), (vi), (vii)
All speakers	**88.68%**	**(ii), (iv)**	**(iii), (v), (vi), (vii)**

Table 44. Summary of rhoticity levels per speaker group in the General American sub-corpus.

These analyses of the individual speaker groups highlight overall trends within the dataset. A detailed examination of the state of rhoticity in the lexical sets NEAR, SQUARE, and CURE at the speaker level offers more nuanced insights into the depiction of rhoticity in General American in the English textbooks used at German secondary schools.

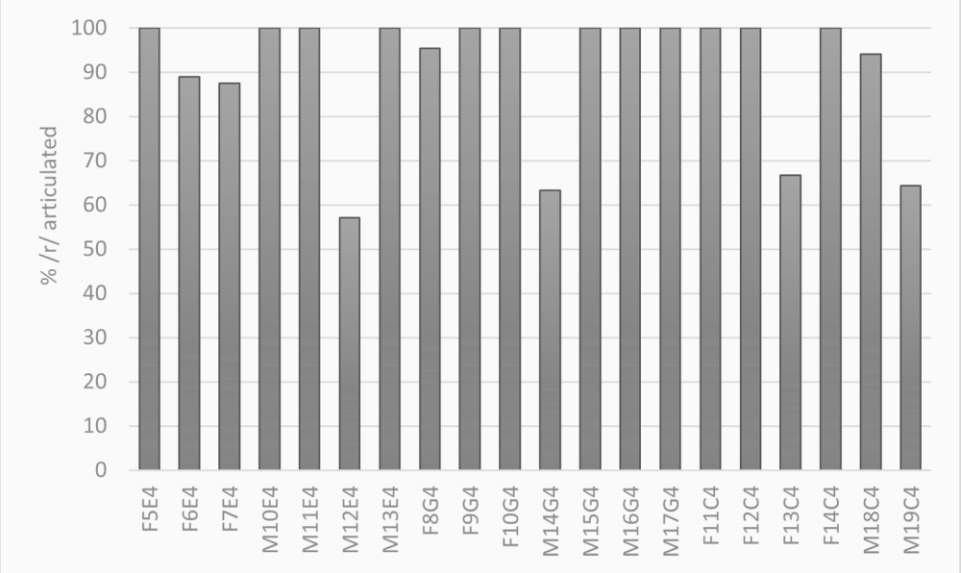


Figure 24. Level of rhoticity in % per speaker in the General American sub-corpus.

Figure 24 displays the level of rhoticity per speaker in the General American sub-corpus. Twelve of the 20 speakers in this sub-corpus articulate post-vocalic /r/ in 100% of instances. This includes four speakers from the textbook *English G Access 4* (F5E4, M10E4, M11E4, and M13E4), five speakers from the textbook *Green Line 4* (F9G4, F10G4, M15G4, M16G4, and M17G4), and three speakers from the textbook *Camden Town 4* (F11C4, F12C4, and F14C4).

Four speakers display a rate of rhoticity between 87-96%: speakers F6E4, F7E4, F8G4, and M18C4. Speaker F6E4 articulates post-vocalic /r/ in 88.89% of instances. Speaker F7E4 has an 87.5% articulation rate, speaker F8G4 has a 95.45% articulation rate, and speaker M18C4 has a 94.12% articulation rate. This relates to one token per speaker being articulated without a consonantal realisation of /r/.

Four speakers display a rate of rhoticity of below 70%: speakers M12E4, M14G4, F13C4, and M19C4. Speaker M12E4 only articulates post-vocalic /r/ in 57.14% of tokens. In three of his seven tokens, /r/ is not articulated. Speaker M14G4 articulates /r/ in 63.33% of tokens. He does not articulate /r/ in 11 of 30 tokens. Speaker F13C4 articulates /r/ in 66.67% of tokens, which accounts for two of three tokens in which /r/ is articulated. Speaker M19C4 articulates /r/ in 64.29% of tokens. He does not produce post-vocalic /r/ in five out of 14 tokens.

Speaker	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
F5E4	3	2	1	1	3	1
M10E4	0	0	0	1	5	3
M11E4	0	3	2	0	3	4
M13E4	0	1	0	0	6	0
F9G4	0	2	0	0	4	2
F10G4	0	1	0	2	1	3
M15G4	0	2	1	1	2	0
M16G4	0	0	1	1	0	0
M17G4	0	1	1	1	4	2
F11C4	0	2	0	1	3	3
F12C4	3	0	1	6	10	0
F14C4	1	0	0	0	0	0
Total	**7**	**14**	**7**	**14**	**41**	**18**

Table 45. Absolute frequencies of articulated /r/ across the phonological contexts for the speakers with 100% rhoticity.

The 12 speakers who articulate post-vocalic /r/ in all instances do so across the range of phonological contexts. Table 45 above displays the frequency of articulated /r/ across the individual phonological contexts for each of the 12 speakers who consistently articulate /r/ in this data set. In the phonological contexts (ii)-(iv), /r/ is in an intervocalic position, contexts (v) and (vi) are preconsonantal positions, and context (vii) is a word, phrase, or utterance final position (see Table 43 for the description of the individual phonological contexts).

Eleven speakers produced tokens of NEAR, SQUARE, or CURE with /r/ in an intervocalic position. Ten of these eleven speakers also produced tokens of NEAR, SQUARE, or CURE with /r/ in preconsonantal positions. Thus, these speakers not only display rhoticity in intervocalic, but also in preconsonantal contexts. Only speaker F14C4 produced only one token of NEAR, SQUARE, or CURE overall. She produced only one intervocalic token within a word before an unstressed syllable.

Seven speakers—F5E4, M10E4, M11E4, F9G4, F10G4, M17G4, and F11G4—produced tokens of NEAR, SQUARE, or CURE with /r/ in word-final position.

The most frequent phonological context of /r/ for these speakers is /r/ in preconsonantal position across a word boundary. Post-vocalic /r/ within the lexical sets NEAR, SQUARE, and CURE occurs in 41 out of 101 total instances of post-vocalic /r/ in this context. Ten of the 12 speakers produced tokens in this context. Speakers M16G4 and F14C4 did not produce any tokens in this phonological context.

Not all speakers in the General American sub-corpus display full rhoticity in this data set. Eight speakers display variable rhoticity at different levels. Table 46 below displays the number of tokens in which /r/ is not articulated relative to the absolute frequency of possible /r/ tokens in the individual phonological contexts and across speakers. The first number indicates the tokens in which /r/ has not been articulated, and the second number indicates the total number of tokens for that phonological context and speaker. Thus, for instance, speaker F6E4 produced five tokens in total for context (vi) and one of these five was articulated without /r/.

Speaker	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
F6E4	0/0	0/0	0/2	0/1	1/5	0/0
F7E4	0/0	0/0	0/2	0/1	1/5	0/0
F8G4	0/0	0/5	0/1	0/6	1/7	0/3
M18C4	0/1	0/4	0/2	0/5	0/3	1/2
M12E4	0/0	0/1	0/0	0/0	3/6	0/0
M14G4	0/0	0/5	0/5	0/2	11/15	0/3
F13C4	0/0	0/1	0/0	0/0	1/1	0/1
M19C4	0/0	3/6	0/0	1/3	1/4	0/1
Total	**0/1**	**3/17**	**0/12**	**1/18**	**19/46**	**1/10**

Table 46. Overview of not articulated /r/ tokens across the individual speakers and phonological contexts.

Overall, across the eight speakers, 24 out of 95 tokens were articulated without /r/. Nineteen of these tokens were within a preconsonantal context across a word boundary (context (vi)). Three non-/r/ tokens were within an intervocalic position across a word boundary with a succeeding stressed syllable (context (iii)). One non-/r/ token occurred in context (v) in a preconsonantal position within a word, and the remaining non-/r/ token was in word-final position (vii).

Six of the eight speakers only showed non-articulation of /r/ in context (vi): preconsonantally across a word boundary. Four of these six speakers only displayed non-articulation of /r/ in just one token: speakers F6E4, F7E4, F8G4, and F13C4. Speaker M12E4 did not articulate /r/ in half of his tokens for context (vi), while speaker M14G4 even displayed a non-rhotic articulation in 11 of 15 tokens in context (vi).

Only one speaker displayed tokens articulated without /r/ in a word-final position (context (vii)). Speaker M18C4 only had one token of not-articulated /r/ in this context. This speaker articulated /r/ in all other contexts.

These results have demonstrated the state of rhoticity in the lexical sets NEAR, SQUARE, and CURE in the General American corpus data. While the grouped results for the textbook groups show that all three textbooks show variable rhoticity in the lexical sets NEAR, SQUARE, and CURE, the analyses of the individual speakers revealed, that the majority of speakers actually displays 100% rhoticity in the lexical sets NEAR, SQUARE, and CURE. Only eight speakers display variable rhoticity in the data set.

This section presented the results of the auditory analyses of rhoticity in the General American corpus data to illustrate how rhoticity in General American is depicted in English textbooks used at German secondary schools in NRW. The findings of the vowel analysis presented in section 4.3.1 and these rhoticity results are expanded on and interpreted in the following section.

4.3.3 Synthesis—the General American Corpus Data and General American in Linguistic Literature

This section expands on the results presented in Sections 4.3.1 and 4.3.2, offering a detailed analysis of the representation of General American in English textbooks used at German secondary schools in NRW. The articulatory categorisations of the grouped and the speaker-individual vowel results, and the state of rhoticity are compared to the linguistic description of General American presented in Section 2.6. This facilitates a detailed interpretation of the representation of General American in these textbook materials.

The features of interest in this sub-corpus are the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE. The following paragraph summarises the linguistic description of these features as presented in Section 2.6.

BATH is an open front vowel in General American with the phonetic quality of TRAP (Tottie 2002: 17; Wells 1982a: 133-134). LOT is an open central to open back vowel with the phonetic quality of PALM or START (Jones, 2011: vii; ix; Wells 1982a: 130; 2008: xxiv). CLOTH is an open back to open-mid back vowel with the phonetic quality of THOUGHT (Wells 1982a: 136; Zsiga 2013: 430). As General American is a rhotic accent of English, the lexical sets NEAR, SQUARE, and CURE are produced with post-vocalic-/r/ in all phonological contexts (Tottie 2002: 18).

These descriptions of the phonetic realisations of the monophthongs BATH, LOT, and CLOTH and rhoticity in NEAR, SQUARE, and CURE are used to contextualise and interpret the results presented in Sections 4.3.1 and 4.3.2. First, the articulatory categorisation of the results for BATH, LOT, and CLOTH regarding vowel height and tongue advancement and then the state of rhoticity in the corpus data are compared to the theory presented above. The results facilitate an interpretation of the realisations of the monophthongs BATH, LOT, and CLOTH and rhoticity in NEAR, SQUARE, and CURE in the grouped and speaker-individual results regarding the representation of General American in the corpus data.

Table 47 below provides an overview of General American pronunciation patterns in the monophthongs BATH, LOT, and CLOTH for the speaker groups from the textbooks *English G Access 4*, *Green Line 4*, *and Camden Town 4*, as well as for all speakers as a group based on the articulatory classifications presented in Table 34 (page 90). These articulatory classifications are matched to the descriptions of General American presented above. Articulatory classifications of the corpus data that match General American are indicated with GA (for General American) in Table 47 below. Overall, the speaker groups display General American pronunciation patterns consistently for BATH, LOT, and CLOTH.

Speaker groups	BATH	LOT	CLOTH
EGA4	GA	GA	GA
GL4	GA	GA	GA
CT4	GA	GA	GA
All speakers	**GA**	**GA**	**GA**

Table 47. Overview of General American pronunciation patterns in BATH, LOT, and CLOTH for the different speaker groups based on articulatory descriptions.

For all speaker groups, BATH is an open front vowel, which is consistent with General American. All speaker groups display General American pronunciation patterns in LOT, represented by the use of an open back vowel. CLOTH is an open-mid back vowel for the speaker groups from *English G Access 4* and *Green Line 4*, and an open back vowel for the speakers from *Camden Town 4*, as well as for all speakers as a cohort. Both pronunciation variants display a feature of General American.

The state of rhoticity in this sub-corpus data is analysed with focus on the different phonological contexts of /r/ that can occur in the lexical sets NEAR, SQUARE, and CURE. In contexts (ii) – (iv), /r/ occurs intervocalically: (ii) within a word before an unstressed syllable, (iii) across a word boundary before a stressed syllable, and (iv) across a word boundary before an unstressed syllable. In contexts (v) and (vi), /r/ occurs preconsonantally: (v) within a word and (vi) across a word boundary. In context (vii), /r/ occurs in word, phrase, or utterance final position. Table 48 below provides an overview of the state of rhoticity in the vowels NEAR, SQUARE, and CURE across the different phonological contexts for the speaker groups from the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4*, as well as for all speakers as a group. Phonological contexts in which the speaker groups showed full rhoticity are marked with *GA*. Three dashes (---) indicate that a phonetic context

displayed variable rhoticity, so not all tokens in that category were articulated with /r/. The letter 'x' indicates that the phonetic context does not occur in that speaker group's data set. The degree of rhoticity per speaker group in the form of the relative frequency of articulated /r/ is also presented.

Speaker groups	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	% /r/ articulated
EGA4	GA	GA	GA	GA	---	GA	92.19%
GL4	x	GA	GA	GA	---	GA	85.71%
CT4	GA	---	GA	---	---	---	89.06%
All speakers	**GA**	**---**	**GA**	**---**	**---**	**---**	**88.68%**

Table 48. Overview of General American pronunciation patterns of rhoticity in NEAR, SQUARE, and CURE in different phonological contexts for the different speaker groups.

All speaker groups produce full rhoticity in the lexical sets NEAR, SQUARE, and CURE intervocalically across a word boundary before an unstressed syllable (context (iv)). All speaker groups produce variable rhoticity, so not all tokens were articulated with /r/, in a preconsonantal context across a word boundary (context (vi)). Overall, the speaker group from *English G Access 4* is most consistent in the depiction of rhoticity: they are fully rhotic in five of the six phonological contexts. They produce post-vocalic /r/ in intervocalic positions (contexts (ii)-(iv), in preconsonantal position within a word (v), and also word-finally (vii). The speaker group from *Camden Town 4* is the least consistent in presenting rhoticity in terms of phonological contexts. The *Camden Town 4* speakers only depict full rhoticity in intervocalic contexts (context (ii) and (iv)) and display variable rhoticity in intervocalic contexts across a word boundary before a stressed syllable (context (iii)), preconsonantally within a word (context (v)), and across a word boundary (context (vi)). In terms of the relative frequency of the level of rhoticity (i.e. the relative frequency of /r/ articulated per speaker group), the speaker group from *English G Access 4* displays the highest level of rhoticity with /r/ articulated in 92.19% of tokens. *Green Line 4* displays the lowest rate of rhoticity with 85.71%.

A closer look at the individual speakers is necessary to analyse the depiction of General American in these textbooks and provide a complete picture of how these textbooks portray General American. First, the depiction of the monophthongs BATH, LOT, and CLOTH is assessed. Each individual speaker's averaged vowel pronunciations of the monophthongs BATH, LOT, and CLOTH were classified into the articulatory categories of vowel height and

tongue advancement in Section 4.3.1. These articulatory classifications are matched to the descriptions of General American presented above. Articulatory classifications that match General American are indicated with GA (for General American) in Table 49 below. Three dashes (---) indicate that an articulatory classification does not match General American, and the letter 'x' indicates that a vowel is not present in that speaker's data set.

Speaker	BATH	LOT	CLOTH
F6E4	GA	GA	GA
F13C4	GA	GA	GA
M18C4	GA	GA	GA
M19C4	GA	GA	GA
F8G4	GA	GA	GA
F9G4	GA	GA	GA
F11C4	GA	GA	GA
F5E4	GA	GA	---
F10G4	GA	GA	---
F12C4	GA	GA	GA
F7E4	GA	GA	---
M13E4	GA	GA	---
M10E4	GA	---	---
M12E4	GA	---	---
M15G4	GA	---	GA
M14G4	GA	---	GA
M11E4	---	GA	---
M17G4	---	GA	x
M16G4	x	GA	GA
F14C4	x	GA	x

Table 49. Overview of General American pronunciation patterns in BATH, LOT, and CLOTH in different phonological contexts for the individual speakers.

Eight speakers (F6E4, F13C4, M18C4, M19C4, F8G4, F9G4, F11C4, and F12C4) display General American pronunciation patterns in all three monophthongs BATH, LOT, and CLOTH. Sixteen speakers have an open front vowel in BATH and thus display a General American pronunciation in this vowel. Two speakers (M16G4 and F14C4) have not produced any

tokens of this vowel, and two speakers (M11E4 and M17G4) have an open-mid front realisation of BATH, which is inconsistent with General American. Sixteen speakers, albeit different speakers, produce a LOT vowel consistent with General American pronunciation patterns. Twelve of these speakers (F6E4, F13C4, M18C4, M19C4, F8G4, F9G4, F11C4, F5E4, F10E4, M11E4, M17G4, and M16G4) have an open back realisation of LOT. The remaining four speakers (F12C4, F7E4, M13E3, and F14C4) have an open central realisation of LOT. CLOTH has a General American realisation for 11 speakers. Six speakers (F6E4, F13C4, M18C4, M19C4, M14G4, and M16G4) have an open back vowel in CLOTH, while five speakers (F8G4, F9G4, F11C4, F12C4, and M15G4) have an open-mid back realisation of CLOTH.

The state of rhoticity in the vowels NEAR, SQUARE, and CURE across the different phonological contexts for the individual speakers is presented in Table 50 below. Phonological contexts in which the speakers showed full rhoticity are marked with *GA*. Three dashes (---) indicate that a phonetic context displayed variable rhoticity, so not all tokens in that category were articulated with /r/. The letter 'x' indicates that the phonetic context does not occur in that speaker's data set. The degree of rhoticity per speaker in the form of the relative frequency of articulated /r/ is also presented. The speakers are presented in the same order as in Table 49 above.

Twelve speakers display full rhoticity in the lexical sets NEAR, SQUARE, and CURE, which is compliant with a depiction of General American. A full level of rhoticity in preconsonantal and word-final contexts is of particular importance here, as /r/ is articulated in intervocalic contexts in non-rhotic accents of English, as well. Eleven of these 12 speakers show rhoticity in preconsonantal contexts, which matches a General American pronunciation. Speaker F14C4 only produced one token of NEAR, SQUARE, or CURE and only in an intervocalic context. Thus, for this speaker, the full level of rhoticity does not necessarily support a representation of General American.

Speaker groups	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	% /r/ articulated
F6E4	x	x	GA	GA	---	x	88.89%
F13C4	x	GA	x	x	---	GA	66.67%
M18C4	GA	GA	GA	GA	GA	---	94.12%
M19C4	x	---	x	---	---	GA	64.29%
F8G4	x	GA	GA	GA	---	GA	95.45%
F9G4	x	GA	x	x	GA	GA	100%
F11C4	x	GA	x	GA	GA	GA	100%
F5E4	GA	GA	GA	GA	GA	GA	100%
F10G4	x	GA	x	GA	GA	GA	100%
F12C4	GA	x	GA	GA	GA	x	100%
F7E4	x	x	GA	GA	---	x	87.5%
M13E4	x	GA	x	x	GA	x	100%
M10E4	x	x	x	GA	GA	GA	100%
M12E4	x	GA	x	x	---	x	57.14%
M15G4	x	GA	GA	GA	GA	x	100%
M14G4	x	GA	GA	GA	---	x	63.33%
M11E4	x	GA	GA	x	GA	GA	100%
M17G4	x	GA	GA	GA	GA	GA	100%
M16G4	x	x	GA	GA	x	x	100%
F14C4	GA	x	x	x	x	x	100%

Table 50. Overview of General American pronunciation patterns of rhoticity in NEAR, SQUARE, and CURE in different phonological contexts for the individual speakers.

Six of the eight speakers displaying variable rhoticity do so only in a preconsonantal context across a word boundary (context (vi)): speakers F6E4, F13C4, F8G4, F7E4, M12E4, and M14G4. Speaker F13C4 only produces one token in context (vi) and does not articulate the /r/: *their* in the context of '*their home*'. The remaining five speakers produce non-rhotic as well as rhotic tokens in this phonological context. Interestingly, three of these speakers (F6E4, F7E4, M12E4) produce the non-rhotic tokens in context (vi) in a very specific phonological context: for these speakers, /r/, or rather the not articulated /r/, is followed by the semivowel /w/ in the following word as in *there was* (speakers F6E4 and M12E4), *sure what* (speaker M12E4), and *care what* (speaker F7E4). Acoustically, /r/ has a low F3, while /w/ has a high F3 (King and Ferragne 2020: 16). Thus, it can be assumed that the non-rhotic

articulation in these tokens is likely due to coarticulation processes with succeeding /w/. These four speakers could then still be considered fully rhotic.

Speaker M19C4 is the only speaker in this sub-corpus to display variable rhoticity in more than one phonological context. Three instances of non-rhotic pronunciations occur in intervocalic contexts across a word boundary before a stressed syllable (context (iii)): *there are* (twice), and *there is*. One instance of a non-rhotic pronunciation occurs preconsonantally within a word (context (v)) in the word *years*. One instance of a non-rhotic pronunciation occurs across a word boundary (context (vi)) in the words *where dogs*.

This section offered a detailed analysis of the representation of General American in English textbooks used at German secondary schools in NRW, based on the analysis of the monophthongs BATH, LOT, and CLOTH and the state of rhoticity in the vowels NEAR, SQUARE, and CURE. The analysis of the data at the superordinate level of the textbook speaker groups, and at the speaker-individual level revealed interesting insights into the representation of General American in the textbook audio materials. The interpretation of the acoustic vowel analyses and auditory rhoticity analysis revealed that the corpus data matches the linguistic description of General American in many, but not all, variables.

Section 4.4 presents and interprets the findings from the analyses of the Australian English sub-corpus following a similar structure.

4.4 Australian English

This section investigates how Australian English, and specifically the three diphthongs FACE, PRICE, and MOUTH, are represented in the German textbooks of English: *English G Access 5* (Rademacher 2017a) and *Green Line 5* (Weisshaar 2018a). Section 4.4.1 illustrates the results of the acoustic analyses of the corpus data to depict the Australian English accent of the textbooks. The section moves from a general overview of the pronunciation of the entire speaker group to more fine-grained analyses of the two textbook speaker groups and the individual speakers. This analysis provides a comprehensive overview of how Australian English is depicted in German textbooks of English. The results of the acoustic analyses of the Australian English sub-corpus have also been presented in Scheiwe (2022). The method of categorising the acoustic results into the articulatory categories of vowel height and tongue advancement in this study (as presented in Section 3.3.1) differs slightly from Scheiwe (2022): in the previous study, this segmentation was achieved by hand instead of

the carefully constructed segmentation employed in this study. Thus, a few individual results presented here differ slightly from the results presented in Scheiwe (2022).

Section 4.4.2 expands on these findings: a comparison with the linguistic description of Australian English, as presented in Section 2.7, facilitates an interpretation of these findings. The results of the acoustic analyses are evaluated against the descriptions of Australian English. The results presented here expand on the results presented in Scheiwe (2022) by considering not only General and Broad Australian English as the theoretical framework for the comparison, but also Cultivated Australian English.

4.4.1 The Vowels in the Australian English Sub-Corpus

Speech samples from 13 textbook speakers have been analysed to investigate the representation of Australian English as exemplified by the diphthongs FACE, PRICE, and MOUTH. Six speakers were analysed from the textbook *English G Access 5*, and seven speakers from the textbook *Green Line 5*. A combined total of 729 tokens for FACE, PRICE, and MOUTH were analysed across these 13 speakers.

Three speaker groups were created to get a first impression of the representation of Australian English in German textbooks: a group of all speakers, all speakers from *English G Access 5* (EGA5), and all speakers from *Green Line 5* (GL5). For each group, pooled averages for F1 and F2 of the first targets in FACE, PRICE, and MOUTH were calculated based on the groups' speakers' mean raw F1 and F2 values of the first targets in these vowels.

Speaker groups	FACE				PRICE				MOUTH			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
EGA5	F1	761.2	112.4	6	F1	717.3	93.2	6	F1	762.3	95.5	6
	F2	1469.5	120.2		F2	1122.9	132.2		F2	1670.1	232.0	
GL5	F1	764.6	88.3	7	F1	853.6	102.8	7	F1	797.4	163.3	7
	F2	1724.2	197.3		F2	1234.0	126.2		F2	1767.7	140.4	
All	**F1**	**763**	**95.8**	**13**	**F1**	**790.7**	**117.9**	**13**	**F1**	**781.2**	**115.8**	**13**
	F2	**1606.6**	**207.2**		**F2**	**1182.7**	**136.2**		**F2**	**1722.6**	**186.7**	

Table 51. Pooled mean formant values and standard deviations for the first targets in FACE, PRICE, and MOUTH for the speaker groups from English G Access 5, Green Line 5, and all speakers.

Table 51 displays the pooled averages in hertz of the F1 and F2 values for the first targets in the vowels FACE, PRICE, and MOUTH, and the standard deviations (SD) per speaker group per vowel. Both the mean F1 and F2 values and the SD are rounded to one decimal place. The token number (n) indicates the number of speakers per group, as these mean values were calculated based on the speaker-individual mean F1 and F2 values per vowel. The pooled averages of F1 and F2 for the diphthongs FACE, PRICE, and MOUTH are visualised in Figure 25 – Figure 27 below. Figure 25 displays the mean vowel plot for the entire speaker group, Figure 26 displays the mean vowel plot for the speakers from *English G Access 5*, and Figure 27 displays the mean vowel plot for the speakers from *Green Line 5*.

The pooled mean vowel realisations of the monophthongs for each speaker group are used as the reference points in each plot. Thus, the vowel plot for the entire speaker group also displays the pooled mean vowel realisations of the monophthongs for this speaker group. The vowel spaces for the speaker groups from *English G Access 5* and *Green Line 5* are each created from the pooled mean vowel realisations of the monophthongs for these respective speaker groups. This procedure provides a general first impression of the Australian English accent in German textbooks of English. These results are interpreted regarding their representation of Australian English, i.e. whether these features represent an Australian English pronunciation or not, in Section 4.4.2.

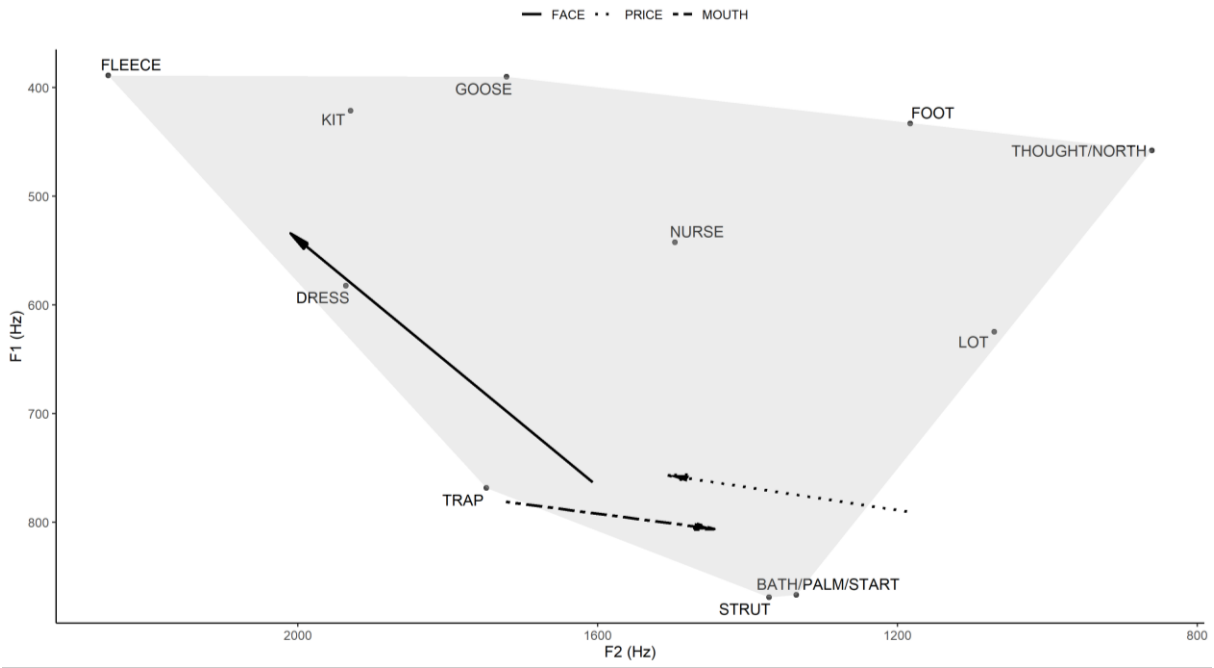


Figure 25. Averaged realisations of the diphthongs FACE, PRICE, and MOUTH for all 13 speakers from the Australian English sub-corpus.

Figure 25 displays an averaged vowel plot of all speakers' individual averaged F1 and F2 values in hertz. These mean values were calculated from each speaker's averaged F1 and F2 values in hertz for each vowel. Averaged trajectories for FACE, PRICE, and MOUTH from the first to the second target are superimposed onto the monophthongal vowel space.

As discussed above, the position of the first targets in the diphthongs FACE, PRICE, and MOUTH are of interest here. The categorisation of these vowels according to the articulatory features of vowel height and tongue advancement is based on the 'segmented' vowel plots in Appendix 2. These vowel plots include a segmentation into the articulatory categories of front, central, and back on the horizontal (F2) axis, and close, close-mid, mid, open-mid, and open on the vertical (F1) axis.

Both FACE and MOUTH have an open front first target with a vowel quality close to TRAP. The first target in FACE is retracted slightly, compared to the first target in MOUTH. PRICE has an open back first vowel target, which is illustrated by the backed position of this vowel target in comparison to the rather central position of the vowels STRUT, as well as the combined cluster of the vowels BATH, PALM, and START.

Considering all speakers of the sub-corpus as one cohort provides a comprehensive overview of the depiction of Australian English in English textbooks used in NRW, and examining the speaker groups from each textbook individually offers insights into unique pronunciation patterns. Therefore, the subsequent paragraphs delve into the results for each textbook speaker group, starting with the textbook *English G Access 5*.

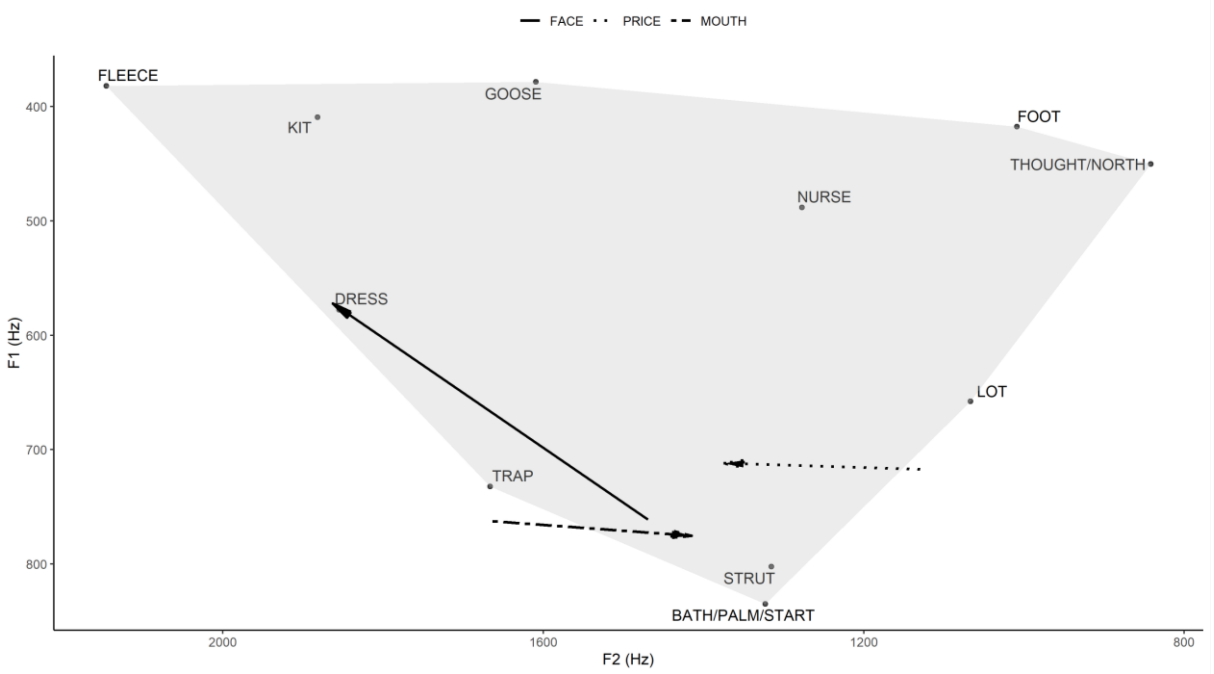


Figure 26. Averaged realisations of the diphthongs FACE, PRICE, and MOUTH for the six English G Access speakers.

Figure 26 shows the averaged vowel plot for all speakers from the textbook *English G Access 5*. The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook. This speaker group has an open front first target in both diphthongs FACE and MOUTH, with the first target of MOUTH being considerably fronter than the first target in FACE. MOUTH starts close to TRAP, whereas FACE has a starting point halfway between TRAP and STRUT. PRICE has an open back first target, which is slightly closer to the open-mid area compared to the realisation in Figure 25 for all textbook speakers. The first target in PRICE is located between START and LOT.

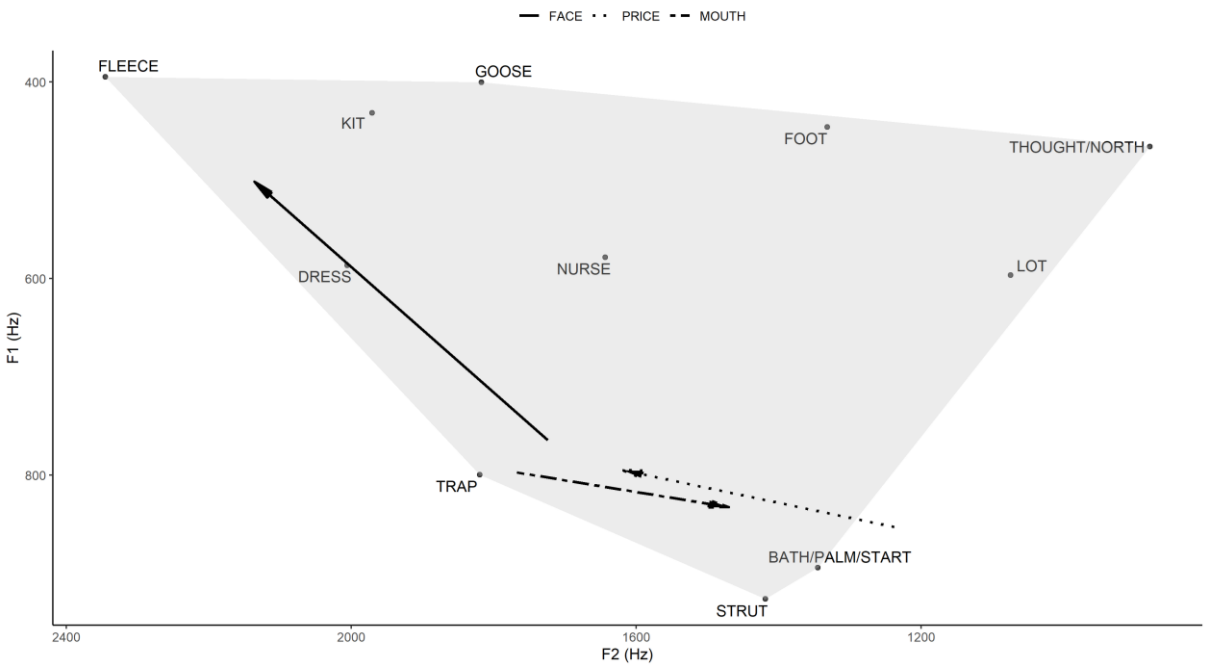


Figure 27. Averaged realisations of the diphthongs FACE, PRICE, and MOUTH for the seven Green Line speakers.

Figure 27 displays the averaged vowel plot for all speakers from the textbook *Green Line 5*. The mean values depicted in this plot are derived from calculating the mean hertz values of the averaged F1 and F2 values per speaker and vowel from the textbook. This speaker group has an open front first vowel target in the diphthong MOUTH, which starts close to TRAP. FACE has an open-mid front first vowel target that is slightly retracted and raised compared to TRAP. PRICE has an open back first vowel target for this speaker group. The first target in PRICE is located between START and LOT, although as LOT is fairly close for this speaker group, the first target in PRICE is much closer to START than to LOT.

Speaker group	FACE	PRICE	MOUTH
EGA5	open front	open back	open front
GL5	open-mid front	open back	open front
All	**open front**	**open back**	**open front**

Table 52. Overview of articulatory categorisation of the first targets in FACE, PRICE, and MOUTH for the speaker groups from English G Access 5, Green Line 5, and all speakers.

Table 52 summarises the results from the acoustic analysis and categorisation into the articulatory features of vowel height and tongue advancement for the two textbook speaker groups and all speakers as a group. The first targets in PRICE and MOUTH are

consistently categorised as open back for PRICE and open front for MOUTH across all three speaker groups: the speakers from the textbook *English G Access 5*, the textbook *Green Line 5*, as well as all speakers combined as a group. Although categorised as open back across all speaker groups, the first target in PRICE displays some variability, nonetheless. For the speakers from the textbook *English G Access 5*, the first target in PRICE is slightly raised towards the open-mid area of the vowel space. The first target in FACE displays some variability in vowel height. If all speakers are considered as a group, FACE has an open front first target. This vowel realisation is also found in the speaker group from *English G Access 5*. The speakers from *Green Line 5*, however, an open-mid front first target in FACE.

These analyses of the individual speaker groups highlight overall trends within the dataset. A detailed examination of the pronunciation patterns at the speaker level offers more nuanced insights into the depiction of Australian English in the English textbooks used at German secondary schools. The analyses at the speaker level are based on each speaker's individual vowel space (see Appendix 1), which are built from each speaker's mean F1 and F2 values of their monophthongs. Table 7 (Appendix 5) provides an overview of the speakers' mean F1 and F2 values in hertz and the standard deviation of the vowels under discussion: the first targets in the diphthongs FACE, PRICE, and MOUTH. The pronunciation of the first target in the diphthongs FACE, PRICE, and MOUTH is evaluated based on their position in the vowel space. The first targets in the diphthongs are categorised according to the articulatory features of vowel height and tongue advancement. The 'segmented' vowel plots for each speaker used for this categorisation can be found in Appendix 2. Table 53 – Table 55 below indicate the different pronunciation variants of the first targets in FACE, PRICE, and MOUTH and indicate how many speakers produce these variants.

	Position of the first target in FACE				
	Mid front	Open-mid front	Open-mid central	Open front	Open central
Number of speakers	2	3	1	6	1

Table 53. Realisation of the first target in FACE based on each speaker's mean formant values.

Table 53 provides an overview of the different pronunciation variants in the first target of FACE and how many the speakers in the Australian English sub-corpus produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The

first target in FACE has five different pronunciation variants in the Australian English sub-corpus: mid front, open-mid front, open-mid central, and open central. All 13 speakers produce tokens of FACE. The first vowel target in FACE primarily varies in vowel height ranging from open to open-mid and even mid tokens. Only two speakers display some variability in the frontness of the first vowel target with a rather central realisation of the first target in this diphthong. Six of the 13 speakers in this sub-corpus produce FACE with an open front first target. Three speakers have an open-mid front first target, while two speakers have a mid-front first target in FACE. Two speakers have a central first target in FACE: one speaker produced FACE with an open-mid central first target, and the other with an open central first target.

	Position of the first target in PRICE			
	Mid back	Open-mid back	Open-mid central	Open back
Number of speakers	1	4	1	7

Table 54. Realisation of the first target in PRICE based on each speaker's mean formant values.

Table 54 provides an overview of the different pronunciation variants in the first target of PRICE and how many of the speakers in the Australian English sub-corpus produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The first target in PRICE has four different pronunciation variants in the Australian English sub-corpus: mid back, open-mid back, open-mid central, and open back. The first vowel target in PRICE mainly varies in vowel height for the 13 speakers, ranging from open to open-mid and even mid tokens. For all speakers but one, PRICE has a back first target; the remaining speaker had a central first target for PRICE. Seven of the 13 speakers in this sub-corpus produce PRICE with an open back first target. Four speakers have an open-mid back first target, while one speaker has a mid back first target, and one has an open-mid central first target.

	Position of the first target in MOUTH		
	Mid front	Open-mid front	Open front
Number of speakers	2	4	7

Table 55. Realisation of the first target in MOUTH based on each speaker's mean formant values.

Table 55 provides an overview of the different pronunciation variants in the first target of MOUTH and how many of the speakers in the Australian English sub-corpus produce these variants based on each speaker's mean F1 and F2 formant values for this vowel. The first target in MOUTH has three different pronunciation variants in the Australian English sub-corpus: mid front, open-mid front, and open front. The first vowel target in MOUTH is a front vowel for all 13 speakers and only varies in vowel height, ranging from a mid front realisation to an open-mid front and an open front realisation. The majority of speakers—seven of the 13—have an open front first target in this diphthong. An open-mid front realisation is the second most frequent realisation, with four speakers showing this pronunciation. Two of the 13 speakers have a mid front first target in MOUTH.

Overall, FACE, PRICE, and MOUTH display a certain variability in vowel quality in this corpus data. A closer look at how the individual textbook speakers realise these three vowels allows for a more detailed analysis of how Australian English is presented in the textbook materials. Table 56 below provides an overview of the distribution of the position of the first targets in FACE, PRICE, and MOUTH across the individual speakers. The sequence of speakers in this table is adapted from the usual sequence introduced in Section 3.2.5 to showcase interesting pronunciation patterns across several speakers.

Speaker	FACE	PRICE	MOUTH
F1E5	open front	open back	open front
M4E5	open front	open back	open front
M5E5	open front	open back	open front
M2E5	open front	open-mid back	open front
F4G5	open front	open central	open front
M6G5	open front	open back	open-mid front
M8G5	open-mid front	open back	mid front
M9G5	open-mid front	open back	mid front
F2G5	open-mid front	open-mid back	open-mid front
M7G5	mid front	open back	open front
F3G5	mid front	open-mid back	open-mid front
M1E5	open central	open-mid back	open front
M3E5	open-mid central	mid back	open-mid front

Table 56. Overview of the position of the first target in FACE, PRICE, and MOUTH for the Australian English speakers.

As established in Table 53 above, an open front first target is the most common realisation for FACE. Of the six speakers that have an open front realisation in the first target of FACE, five also have an open front realisation in the first target of MOUTH: speakers F1E5, M4E5, M5E5, M2E5, and F4G5. Three of these five speakers—F1E5, M4E5, and M5E5—have an open back first target in PRICE. These three speakers thus constitute the biggest cluster of speakers with the same realisation of the first target in each respective diphthong.

To illustrate these results and the analysis further, three speakers, speakers M4E5, F2G5, and M3E5 have been selected for more detailed analysis. These speakers display different combinations of first target realisations for FACE, PRICE, and MOUTH. The individual vowel plots of these speakers are presented in Figure 28–Figure 30 below and their individual realisations of the mean formant values for the first target in FACE, PRICE, and MOUTH are discussed.

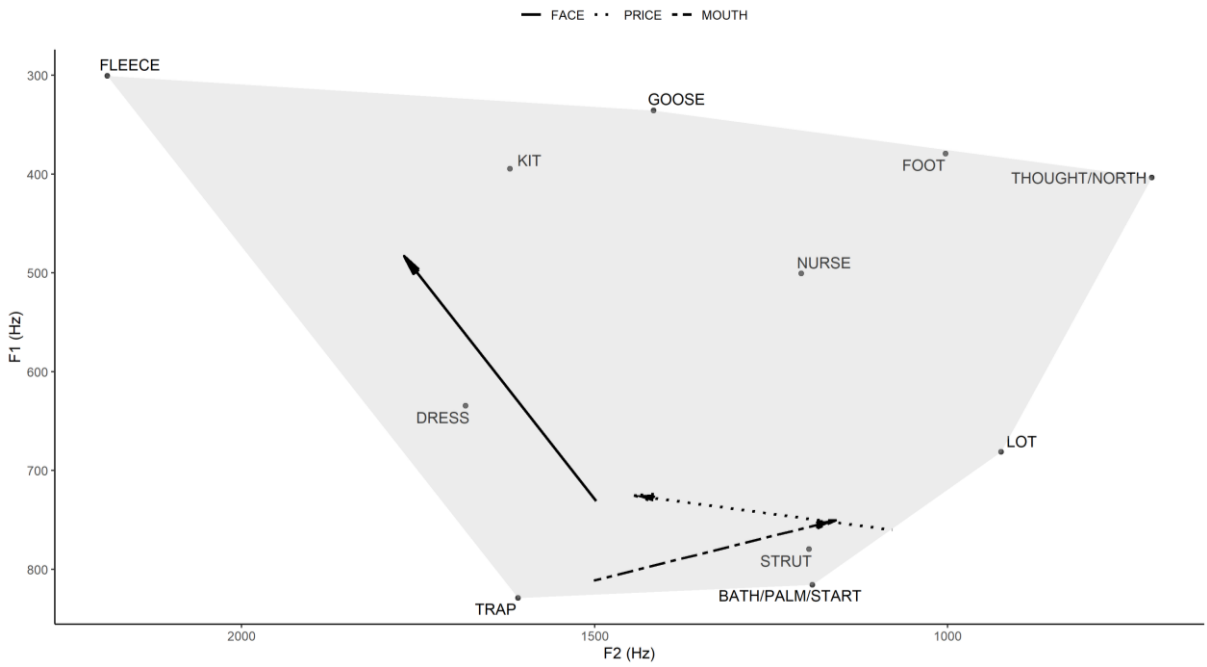


Figure 28. Vowel plot of mean formant values for speaker M4E5 (This vowel plot has been previously published in Scheiwe (2022: 68)).

Figure 28 displays speaker M4E5's vowel space generated by the mean F1 and F2 formant values of the stressed monophthongs available. The diphthong trajectories from the mean values of the first to the second vowel target for the diphthongs FACE, PRICE, and MOUTH are superimposed onto the monophthongal vowel space. Based on their position in the monophthongal vowel space, the vowel quality of the first target of these diphthongs can

be described in the articulatory terms of tongue height and tongue advancement (see segmented vowel plot in Appendix 2).

Speaker M4E5 produces an open front first target in both FACE and MOUTH. The first target in MOUTH is more open and slightly more fronted than the first target in FACE. PRICE has an open back first target situated halfway between the open-mid back vowel LOT and the open central vowels in STRUT, as well as the combined cluster of the vowels BATH, PALM, and START.

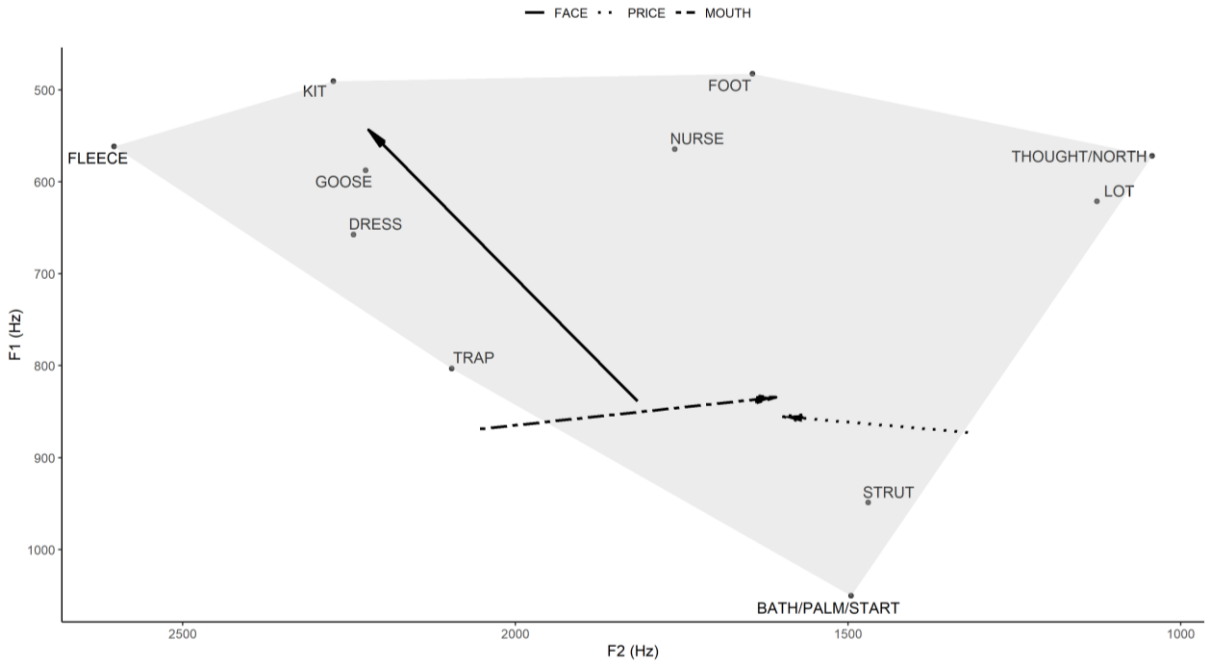


Figure 29. Vowel plot of mean formant values for speaker F2G5 (This vowel plot has been previously published in Scheiwe (2022: 69)).

Figure 29 displays the trajectories from the first to the second vowel target of the mean realisations of speaker F2G5's diphthongs FACE, PRICE, and MOUTH superimposed onto this speaker's monophthongal vowel space. This speaker's vowel space clearly resembles a triangular shape as opposed to the traditional vowel quadrilateral. The first targets in all three diphthongs—FACE, PRICE, and MOUTH—can be described as open-mid: FACE and MOUTH both have an open-mid front first target, while PRICE has an open-mid back first target. The first target in FACE is slightly retracted compared to the first target in MOUTH. Both first targets in FACE and MOUTH have a similar openness as the vowel TRAP, which is an open-mid front vowel for this speaker. The first target in PRICE is located not quite halfway between the cluster of BATH, PALM, and START and the very close realisation of LOT.

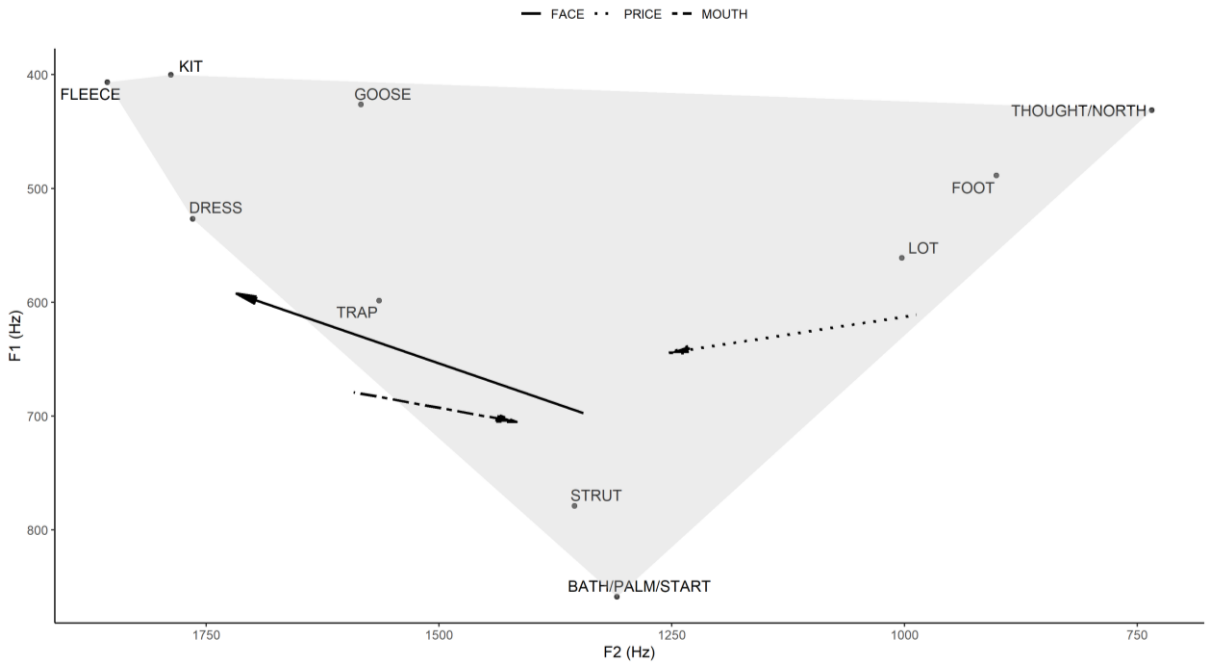


Figure 30. Vowel plot of mean formant values for speaker M3E5.

The last exemplary plot of a speaker's vowel space is depicted in Figure 30. This plot shows speaker M3E5's vowel space and superimposed the averaged diphthong trajectories of FACE, PRICE, and MOUTH. Like speaker F2G5's vowel space, this vowel space is triangular in shape as opposed to the familiar vowel quadrilateral. Both first targets in FACE and MOUTH have an open-mid quality for this speaker. While MOUTH has an open-mid front first target, FACE has an open-mid central first target. PRICE has a mid back first target for this speaker, which is slightly more open than this speaker's realisation of the LOT vowel.

The vowel plots presented in Figure 28 – Figure 30 above depict the individual vowel spaces of speakers M4E5, F2G5, and M3E5. They serve to illustrate the articulatory categorisations of the acoustic analyses presented for the individual speakers in Table 56 above exemplified by these three speakers. These results have demonstrated the variability in vowel realisations within the corpus data. While the grouped results for the textbook groups and all speakers as a cohort suggest a more homogenous depiction of the Australian English features FACE, PRICE, and MOUTH in the textbook audio materials, the analyses of the individual speakers display a higher degree of variability in the realisation of the first targets in the diphthongs FACE, PRICE, and MOUTH.

This section presented the results of the acoustic analyses of the Australian English corpus data to illustrate how the Australian English accent is depicted in English textbooks

used at German secondary schools in NRW. These findings are expanded on and interpreted in the following section.

4.4.2 Synthesis—the Australian English Corpus Data and Australian English in Linguistic Literature

This section expands on the results presented in Section 4.4.1 and offers a detailed analysis of the representation of Australian English in English textbooks used at German secondary schools in NRW. The articulatory categorisations of the grouped and the speaker-individual results are compared to the linguistic description of Australian English presented in Section 2.7. This facilitates a detailed interpretation of the representation of Australian English in these textbook materials.

The features of interest in this sub-corpus are the diphthongs FACE, PRICE, and MOUTH. The diphthongs are salient features of Australian English differentiating this accent from other accents of English. FACE, PRICE, and MOUTH are also markers for the accents on the socio-stylistic continuum within Australian English: Cultivated Australian English, General Australian English, and Broad Australian English (Cox 2008: 329; Harrington, Cox, and Evans 1997: 179). As the majority of speakers of Australian English speak General Australian English (Horvath 2008: 89-90; Harrington, Cox, and Evans 1997: 156), this accent type serves as the first point of reference in comparing the results from the textbook corpus to the linguistic description of Australian English. Deviations from General Australian English are then compared to descriptions of Cultivated and Broad Australian English.

As presented in Section 2.7, FACE and MOUTH both have an open front first target in General Australian English. Male speakers have a first target between TRAP and START in both FACE and MOUTH, whereas the first target in FACE is closer to TRAP for female speakers (Cox 2006: 158; Cox 2008: 331). FACE mostly displays accent differences for speakers of Cultivated Australian English with a more fronted first target compared to speakers of General or Broad Australian English (Harrington, Cox, and Evans 1997: 174). MOUTH shows increasing lowering and retraction in the first target with increasing cultivation. The first target in MOUTH is an open-mid front vowel for speakers of Broad Australian English, which is situated between TRAP and DRESS (Harrington, Cox, and Evans 1997: 179). In Cultivated Australian English, MOUTH has a more open and retracted first target that can be described as an open front to open central vowel, which is slightly fronted compared to START (Cox 1998: 40-41). PRICE has an open back first target for speakers of General Australian English,

situated between START and LOT (Cox 2006: 159; Cox 2008: 331). For speakers of Broad Australian English, PRICE has a more retracted and raised target closer to an open-mid back position (Cox and Palethorpe 2010: 176; Harrington, Cox, and Evans 1997: 172-173). In Broad Australian English, the first target in PRICE is situated close to LOT (Cox 1998: 40). In Cultivated Australian English, the first target in PRICE has an open central position just slightly retracted from START (Cox 1998: 40).

These descriptions of the phonetic realizations of the first targets in FACE, PRICE, and MOUTH are used to contextualize and interpret the results presented in Section 4.4.1. First, the articulatory categorisation of the results regarding vowel height and tongue advancement is compared to the theory presented above. This facilitates an interpretation of the realisations of the first targets in FACE, PRICE, and MOUTH of the grouped and speaker-individual results regarding the representation of Cultivated, General, and Broad Australian English in the corpus data. The representation of Australian English in German textbooks of English is evaluated further based on the relation between the first targets in FACE, PRICE, and MOUTH and the relevant monophthongs in the vowel plots for the grouped and speaker-individual results.

Table 57 below provides an overview of Australian English pronunciation patterns in the first targets of FACE, PRICE, and MOUTH for the speaker groups from the textbooks *English G Access 5* and *Green Line 5*, as well as for all speakers as a group based on the articulatory classifications presented in Table 52 above (page 119). These articulatory classifications are matched to the Cultivated, General, and Broad Australian English accent types as presented above. Three dashes (---) indicate that an articulatory classification does not match any of the three accent types.

Speaker group	FACE	PRICE	MOUTH
EGA5	General Australian	General Australian	General Australian
GL5	---	General Australian	General Australian
All	**General Australian**	**General Australian**	**General Australian**

Table 57. Overview of Australian English pronunciation patterns in the first targets in FACE, PRICE, and MOUTH for the different speaker groups based on articulatory descriptions.

Overall, the speaker groups strongly display General Australian English in the first targets of the diphthongs FACE, PRICE, and MOUTH. Apart from the first target in FACE for the speakers from *Green Line 5*, all vowel realisations represent General Australian English. Open front first targets in FACE and MOUTH and an open back first target in PRICE mark a General Australian accent (Cox 2006: 161). The two textbook speaker groups thus display a General Australian accent in the first targets of PRICE and MOUTH. FACE only represents a General Australian accent for the *English G Access 5* speakers. The speaker group from *Green Line 5* does not display an Australian English pronunciation pattern in FACE, as their open-mid front first target in FACE is not consistent with Cultivated, General, or Broad Australian English. FACE only displays accent differences in Cultivated Australian English compared to General or Broad Australian English. A Cultivated Australian first target in FACE is fronted compared to its counterparts in General or Broad Australian English (Harrington, Cox, and Evans 1997: 174). As the speaker group from *Green Line 5* displays an open-mid front—and thus not just a fronted, but also raised—first target in FACE, this realisation is not consistent with any Australian English pronunciation pattern.

Australian English diphthongs are described according to their articulatory features and, commonly, according to their relation to the Australian English monophthongs (cf. Butcher 2006; Cox 1998; 2006; 2008). In General Australian English, the first targets in FACE and MOUTH are both situated close to TRAP, between TRAP and START, with the first target in FACE being closer to TRAP for female speakers than for male speakers (Cox 2006: 158-159). As FACE is fronter for Cultivated speakers (Harrington, Cox, and Evans 1997: 174), the first target in FACE can be expected to be fronter than TRAP for these speakers. MOUTH shows increasing lowering and retraction in the first target with increasing cultivation. The first target in MOUTH is situated between TRAP and DRESS for Broad Australian English speakers (Harrington, Cox, and Evans 1997: 179). In Cultivated Australian English, MOUTH has a more open and retracted first target, which is slightly fronted compared to START (Cox 1998: 40-41). PRICE has an open back first target for speakers of General Australian English which is situated between START and LOT (Cox 2006: 159; Cox 2008: 331). For speakers of Broad Australian English, PRICE has a more retracted and raised target close to LOT (Cox 1998: 40). In Cultivated Australian English, the first target in PRICE has an open central position just slightly retracted from START (Cox 1998: 40).

If these relations between the position of the first targets in FACE, PRICE, and MOUTH in the vowel space compared to the Australian English monophthongs are considered, the results presented in Table 57 above can be confirmed. Apart from the first target in FACE for

the speakers from *Green Line 5*—which does not display Australian English pronunciation patterns—the remaining vowel realisations in the first targets of FACE, PRICE, and MOUTH in comparison to the respective monophthongal vowel spaces confirm the categorisation into General Australian English. For all speaker groups, the first target in PRICE is situated between START and LOT, and the first target in MOUTH is situated either close to TRAP or between TRAP and START, thus confirming the interpretation as General Australian. The first target in FACE for the *Green Line 5* speaker group is closer than and more retracted than TRAP. This is also inconsistent with the expected Australian English pronunciation patterns. The comparison of the first targets in FACE, PRICE, and MOUTH to the monophthongs in the respective vowel spaces confirms the classification based on the articulatory features of vowel height and tongue advancement: All speaker groups display General Australian pronunciation patterns in PRICE and MOUTH, and the entire speaker groups, as well as the *English G Access* speakers display General Australian pronunciation patterns in FACE.

A closer look at the individual speakers is necessary to analyse the depiction of Australian English in these textbooks and provide a complete picture of how they portray Australian English. If each individual speaker's averaged vowel pronunciations are considered, all speakers display some form of Australian English pronunciation variants in at least one of the three diphthongs. Table 58 below displays whether each speaker's averaged pronunciation of the first targets in FACE, PRICE, and MOUTH is consistent with Cultivated, General, or Broad Australian English based on the articulatory features of vowel height and tongue advancement. Pronunciation variants that do not match any of these three sociolects of Australian English are indicated by three dashes (---).

All speakers display some form of Australian English pronunciation in at least one of the first targets of FACE, PRICE, and MOUTH. Six speakers—M2E5, M4E5, M5E5, M6G5, F1E5, and F4G5—display an Australian English pronunciation pattern in all three diphthongs, albeit to varying degrees of broadness across speakers and speaker-internally across the three diphthongs. The remaining seven speakers only show Australian English pronunciation patterns in one or two of the diphthongs. Overall, General Australian is the dominant sociolect across the 13 speakers.

Speaker	FACE	PRICE	MOUTH
M1E5	---	Broad Australian	General Australian
M2E5	General Australian	Broad Australian	General Australian
M3E5	---	---	Broad Australian
M4E5	General Australian	General Australian	General Australian
M5E5	General Australian	General Australian	General Australian
M6G5	General Australian	General Australian	Broad Australian
M7G5	---	General Australian	General Australian
M8G5	---	General Australian	---
M9G5	---	General Australian	---
F1E5	General Australian	General Australian	General Australian
F2G5	---	Broad Australian	Broad Australian
F3G5	---	Broad Australian	Broad Australian
F4G5	Cultivated Australian	Cultivated Australian	General Australian

Table 58. Overview of Australian English pronunciation patterns in FACE, PRICE, and MOUTH across the individual speakers.

Three speakers—M4E5, M5E5, and F1E5—show a General Australian pronunciation in all three first targets. For these speakers, FACE and MOUTH have an open front first target, and PRICE has an open back first target. To illustrate this further, Figure 31 below shows the averaged vowel plot for speaker M5E5.

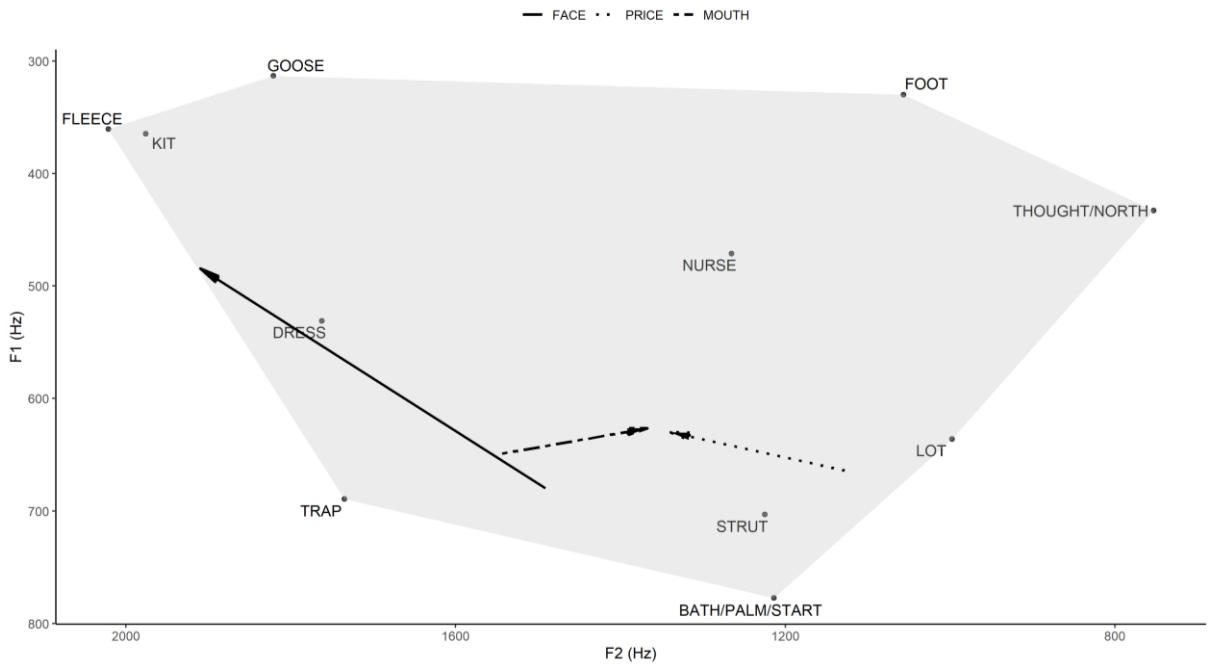


Figure 31. Averaged vowel plot for speaker M5E5 (This vowel plot has been previously published in Scheiwe (2022: 69)).

For speaker M5E5, FACE and MOUTH both have a first target between TRAP and START, although the first target in MOUTH is slightly raised compared to TRAP. Nonetheless, the position of the first targets in these two diphthongs in relation to the monophthongal vowel space confirms the General Australian English pronunciation in these two diphthongs. PRICE has a first target situated between START and LOT, also consistent with a General Australian English pronunciation.

Two speakers, F2G5 and F3G5, display a Broad Australian accent in the first targets of PRICE and MOUTH. Their pronunciation in the first target in FACE does not match any Australian English sociolect. These speakers produce PRICE with an open-mid back first target and MOUTH with an open-mid front first target. FACE has an open-mid front first target for speaker F2G5 and a mid front first target for speaker F3G5. Both pronunciation variants of the first target in FACE are inconsistent with Australian English. Figure 32 below illustrates such a Broad Australian English pronunciation.

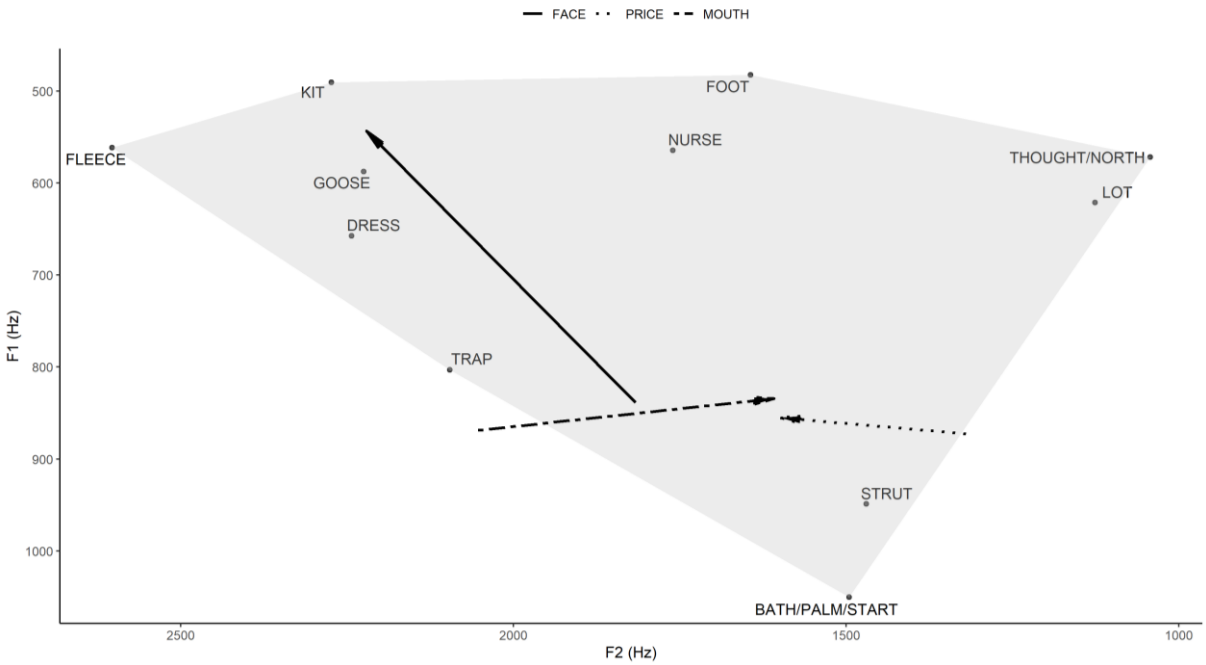


Figure 32. Averaged vowel plot for speaker F2G5 (This vowel plot has also been published in Scheiwe (2022: 69)).

Figure 32 displays speaker F2G5's averaged vowel plot. The first targets in MOUTH and PRICE are clearly raised to an open-mid position in this vowel space (also compare the 'categorised' vowel plot in Appendix 2), indicating a Broad Australian pronunciation. Considering the relation between the first targets in MOUTH and PRICE, and this speaker's monophthongs, however, this interpretation has to be relativized slightly. In Broad Australian English, the first target in MOUTH is situated between DRESS and TRAP (Harrington, Cox, and Evans 1997: 179). For speaker F2G5, however, the first target in MOUTH is situated between TRAP and START—an indicator for General Australian English (Cox 2008: 331). In a similar fashion, the first target in PRICE is usually situated close to LOT in Broad Australian English (Cox 1998: 40). For this speaker, however, the first target in PRICE is situated between START and LOT, an indicator of General Australian English (Cox 2008: 331). Admittedly, LOT has an unusually close realisation for this speaker. Nonetheless, in relation to her monophthongs, speaker F2G5 displays General Australian English features in PRICE and MOUTH.

A closer inspection of speaker F3G5's vowel space in Figure 33 below reveals a similar pattern. For this speaker, TRAP and LOT are both raised: TRAP to an open-mid front position, and LOT to a close back position. Thus, even though the articulatory categorisation of PRICE and MOUTH indicate a Broad Australian accent, the relationship to this speaker's monophthongs rather confirms a General Australian accent. The first target in MOUTH starts

close to TRAP, while the first target in PRICE starts between START and LOT. These are indicators of General rather than Broad Australian English (Cox 1998: 40; Cox 2006: 158-159).

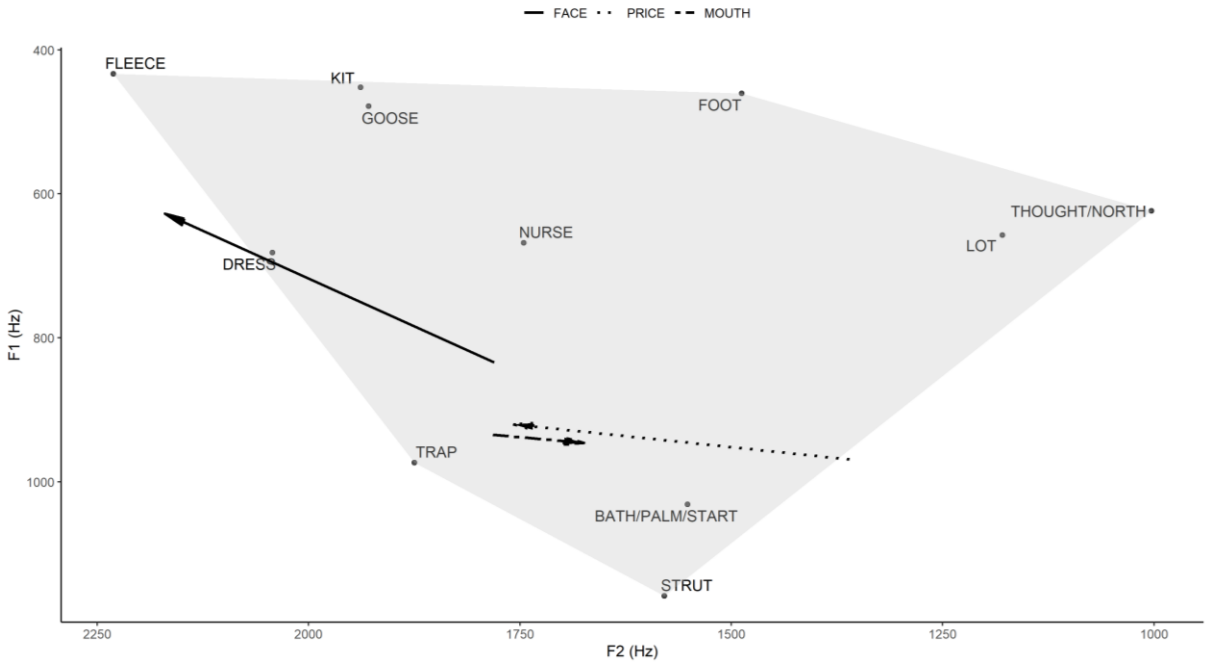


Figure 33. Averaged vowel plot for speaker F3G5.

Overall, these results display that English textbooks used in NRW depict Australian English accent features quite well. At a superordinate level, Australian English is presented quite homogeneously as General Australian English. The entire speaker group, as well as the individual textbook groups. strongly favour General Australian English. An analysis at the speaker level revealed slightly more variation, surfacing features of Cultivated and Broad Australian English. The presence of Broad Australian English, however, cannot be unconditionally accepted, as the relation to the individual speakers' monophthongs did not confirm the categorisation into Broad Australian English based on the articulatory features.

The goal of this section was to expand on the results presented in Section 4.4.1 and offer a detailed analysis of the representation of Australian English in English textbooks used at German secondary schools in NRW. The articulatory categorisations of the grouped and the speaker-individual results were compared to the linguistic description of Australian English presented in Section 2.7. This allowed for a detailed interpretation of the representation of Australian English in these textbook materials.

4.5 Summary

This Chapter presented the findings of the acoustic and auditory analyses of the textbook audio materials regarding the representation of Received Pronunciation, General American, and Australian English in English textbooks used at secondary schools in NRW. Section 4.2 presented the results of the acoustic analysis of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START in the Received Pronunciation sub-corpus and compared these findings to the linguistic description of Received Pronunciation presented in Section 2.5. Section 4.3 presented the results of the acoustic analysis of the monophthongs BATH, LOT, and CLOTH in the General American sub-corpus and compared these findings to the linguistic description of General American presented in Section 2.6. Section 4.4 presented the results of the acoustic analysis of the first targets of the diphthongs FACE, PRICE, and MOUTH in the Australian English sub-corpus and compared these findings to the linguistic description of Australian English presented in Section 2.7.

5 Discussion

5.1 Introduction

This section discusses the findings of the results presented in Chapter 4. The advantages of the corpus-phonological approach are discussed, and the findings are contextualised within the wider discourse on varieties of English and ELT.

5.2 The Accents

This section discusses how Received Pronunciation, General American, and Australian English are represented in the audio materials of the textbook series *English G Access, Green Line,* and *Camden Town*.

5.2.1 Received Pronunciation

This section answers Research Question 1 as presented in Section 1.4: How is Received Pronunciation represented in the audio materials of the textbooks *English G Access 2, Green Line 2,* and *Camden Town 2,* exemplified by the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START? The hypothesis was that the textbooks *English G Access 2, Green Line 2,* and *Camden Town 2* consistently present pronunciation variants consistent with Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START.

The results of the acoustic analyses presented in Section 4.2 are differentiated at two different levels: the textbook level and the speaker level. The analyses at the textbook level are based on the analysis and interpretation of the grouped vowel plots presented in Section 4.2. The analyses at the speaker level are based on the analysis and interpretation of the speaker-individual vowel plots (Appendix 1 and Appendix 2).

At the textbook level, the representation of the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START only partially matches the descriptions of Received Pronunciation presented in Section 2.5. The vowels GOOSE and TRAP are consistent with Received Pronunciation in all three textbooks. BATH is consistent with Received Pronunciation in the textbooks *English G Access 2* and *Camden Town 2,* but not in the textbook *Green Line 2*. The vowels LOT, CLOTH are inconsistent with Received Pronunciation in all three textbooks. START is only consistent with Received Pronunciation in the textbook *Camden Town 2*. Thus, from this more global textbook perspective, the representation of Received Pronunciation

differs only slightly across the three textbooks: all three textbooks display pronunciation patterns consistent with Received Pronunciation only in the vowels GOOSE and TRAP. Two textbooks display pronunciation patterns consistent with Received Pronunciation in BATH, and only one textbook displays a pronunciation pattern consistent with Received Pronunciation in START. All three textbooks show pronunciation patterns inconsistent with Received Pronunciation in LOT and CLOTH.

The analysis at the speaker level revealed similar patterns and slightly more variation in the data. The vowel LOT is inconsistent with Received Pronunciation across all speakers supporting the textbook-level observation. Only one speaker from *Camden Town 2* displays a pronunciation pattern consistent with Received Pronunciation in CLOTH. The individual speakers display varying degrees of Received Pronunciation in the remaining four vowels (GOOSE, TRAP, BATH, and START) without any clear pattern. Across the three textbooks, some speakers display pronunciation patterns consistent with Received Pronunciation in the four vowels GOOSE, TRAP, BATH, and START. Other speakers display pronunciation patterns consistent with Received Pronunciation in one to three vowels.

The hypothesis that the textbooks *English G Access 2*, *Green Line 2*, and *Camden Town 2* consistently present pronunciation variants consistent with Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START has to be refuted. While all three textbooks do show pronunciation patterns of Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START, they do not do it consistently.

5.2.2 General American

This section answers Research Question 2 as presented in Section 1.4: How is General American represented in the audio materials of the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4*, exemplified by the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE? The hypothesis was that the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4* consistently present a General American accent in the pronunciation of the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE.

The results of the acoustic analyses presented in Section 4.3 are differentiated at two different levels: the textbook level and the speaker level. The analyses at the textbook level are based on the analysis and interpretation of the grouped vowel plots in Section 4.3. The

analyses at the speaker level are based on the analysis and interpretation of the speaker-individual vowel plots (Appendix 1 and Appendix 2).

At the textbook level, the representation of the monophthongs BATH, LOT, and CLOTH and the state of rhoticity in the vowels NEAR, SQUARE, and CURE mostly matches the descriptions of General American presented in Section 2.6. The representation of the monophthongs BATH, LOT, and CLOTH matches General American pronunciation patterns in the three textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4*. The representation of rhoticity in the vowels NEAR, SQUARE, and CURE, however, does not fully match the expectations of a General American accent, as full rhoticity would be expected here, and all three textbooks display partial rhoticity in at least one phonological context.

The analysis at the speaker level revealed more variation in the data. The individual speakers display varying degrees of representing General American pronunciation patterns in the vowels BATH, LOT, and CLOTH. Across the three textbooks, several speakers display General American pronunciation patterns in all three monophthongs. Some speakers only display General American pronunciation patterns in one or two monophthongs. Regarding rhoticity, several speakers per textbook display full rhoticity in 100% of tokens in the lexical sets NEAR, SQUARE, and CURE. However, all textbooks also have speakers with variable rhoticity. The speakers from *English G Access 4*, however, only have variable rhoticity in a very specific phonetic context which could quite possibly for account for the non-rhotic tokens. In that case, the speakers from *English G Access 4* can be considered fully rhotic.

The hypothesis that the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4* consistently present a General American accent in the pronunciation of the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE, has to be refuted. While all three textbooks show pronunciation patterns of General American in the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE, they do not do it consistently.

5.2.3 Australian English

This section answers Research Question 3 as presented in Section 1.4: How is Australian English represented in the audio materials of the textbooks *English G Access 5* and *Green Line 5*, exemplified by the diphthongs FACE, PRICE, and MOUTH? The hypothesis was that the textbooks *English G Access 5* and *Green Line 5* display features of Australian English, but not consistently and with a focus on General Australian English.

The results of the acoustic analyses presented in Section 4.4 are differentiated at two different levels: the textbook level and the speaker level. The analyses at the textbook level are based on the analysis and interpretation of the grouped vowel plots presented in Section 4.4. The analyses at the speaker level are based on the analysis and interpretation of the speaker-individual vowel plots (Appendix 1 and Appendix 2).

At the textbook level, the representation of the first targets in the diphthongs FACE, PRICE, and MOUTH completely fits the descriptions of Australian English presented in Section 2.7 only for the textbook *English G Access 5*. The speaker group from *Green Line 5* only depicts Australian English pronunciation patterns in the first targets in PRICE and MOUTH, but not in the first target in FACE. Both textbooks portray features of General Australian English.

The analysis at the speaker level revealed a similar focus on General Australian English, but also slightly more variation in the data. The individual speakers display varying degrees of representing Australian English pronunciation patterns in the first targets of the diphthongs FACE, PRICE, and MOUTH. Across the two textbooks, several speakers display Australian English pronunciation patterns in all three diphthongs with varying degrees of broadness. In the textbook *English G Access 5*, most speakers with Australian realisations in the diphthongs FACE, PRICE, and MOUTH have pronunciation patterns consistent with General Australian English in all three diphthongs. One speaker displays General Australian in FACE and MOUTH, but Broad Australian in PRICE. The speakers from *Green Line 5* only display pronunciation patterns with varying broadness degrees in all three diphthongs: either General Australian in FACE and PRICE and Broad Australian in MOUTH, or Cultivated Australian in FACE and PRICE, and General Australian in MOUTH. Some speakers in both textbooks only display Australian English pronunciation patterns to varying degrees in one or two diphthongs.

The hypothesis that the textbooks *English G Access 5* and *Green Line 5* display features of Australian English, but not consistently and with a focus on General Australian English can be confirmed. The textbooks both display features of Australian English which match General Australian in most cases. They are, however, not consistent in portraying Australian English in all features.

5.3 Implications for English Language Teaching

This section adopts a broader perspective to contextualise the results within the wider discourse on varieties of English and ELT. Many researchers have acknowledged that learners of English should be made aware of the diversity within the English language (Bieswanger 2012: 362; Matsuda 2013: 1; Matsuda and Friedrich 2012: 23). To raise the learners' awareness, they should be exposed to different varieties of English (Matsuda and Matsuda 2018: 67, Sung 2016: 191) through, for instance, teaching materials (Matsuda 2013: 2; Matsuda and Friedrich 2012: 24). Such exposure is crucial in the development of communicative competence (Bieswanger 2012: 363), as ultimately, learners need to be able to communicate with speakers of different varieties of English.

This study investigated how different accents of English are represented in the textbook audio materials of German textbooks of English. While previous research on the inclusion of different varieties into teaching materials rather focused on countries included, this study conducted a detailed linguistic analysis of three select accents in the audio materials of three textbook series used at secondary schools in NRW.

Ultimately, the three textbook series *English G Access*, *Green Line*, and *Camden Town* display salient features of the accents Received Pronunciation, General American, and Australian English. While none of these accents are portrayed consistently in either textbook, the key conclusion to be drawn is that all three textbook series do include varieties of English in their materials and present salient features of each variety. Therefore, these textbooks and accompanying audio materials provide an effective and practical method for introducing learners to the diversity of the English language. The audio materials expose learners to different varieties of English by presenting features representative of these accents. They also illustrate subtle variations within these accents at the speaker level, thereby exposing the learners to some linguistic diversity of the English language.

5.4 Benefits of a Corpus-Phonological Approach to Textbook Analysis

As presented in Section 1.3, previous research on the inclusion of different varieties of English mostly focused on whether different varieties of English are included in teaching materials. This study introduced a new approach to a variety-centred textbook analysis. The corpus-phonological approach employed in this study included detailed acoustic and auditory analyses of select phonological features to investigate the representation of Received

Pronunciation, General American, and Australian English in textbook audio materials used at NRW secondary schools.

This methodological approach to investigating the representation of English varieties in teaching materials enriches the discourse around varieties in ELT and offers several key advantages. The theory-based assessment of the acoustic and auditory analyses provided valuable insights into the linguistic diversity present in the teaching materials. The meticulous selection and analysis of the data contributed to the reliability of the findings. To ensure the validity of the findings, the results from the automated acoustic analysis was validated manually by the researcher. This process involved systematically reviewing the results of the automated extraction of the formant values to confirm their accuracy. For the auditory rhoticity analysis, the inter-rater reliability was assessed.

5.5 Summary

This section discussed the findings of the results presented in Chapter 4. The advantages of the corpus-phonological approach were discussed, and the findings were contextualised within the wider discourse on varieties of English and ELT.

6 Conclusion

This dissertation set out to answer the research question: How are different varieties of English represented in the audio materials of German textbooks for English? This research question was further broken down into the three accent-specific research questions:

Research Question 1: How is Received Pronunciation represented in the audio materials of the textbooks *English G Access 2*, *Green Line 2*, and *Camden Town 2*, exemplified by the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START?

Research Question 2: How is General American represented in the audio materials of the textbooks *English G Access 4*, *Green Line 4*, and *Camden Town 4*, exemplified by the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE?

Research Question 3: How is Australian English represented in the audio materials of the textbooks *English G Access 5* and *Green Line 5*, exemplified by the diphthongs FACE, PRICE, and MOUTH?

To answer these research questions, this study employed a corpus-phonological approach to textbook analysis. The phonological textbook corpus was comprised of the three accents Received Pronunciation, General American, and Australian English and the volumes two, four, and five of the textbook series' *English G Access*, *Green Line*, and *Camden Town*. Volume five was only included for the textbooks *English G Access* and *Green Line*.

Select phonological features were analysed for the three accents. The monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START were analysed for Received Pronunciation, the monophthongs BATH, LOT, and CLOTH and rhoticity in the vowels NEAR, SQUARE, and CURE were analysed for General American, and the first targets in the diphthongs FACE, PRICE, and MOUTH were analysed for Australian English. Acoustic analyses of the vowels and an auditory analysis of rhoticity in General American provided a comprehensive overview of how these features are represented in the textbook audio materials.

Research Question 1 can be answered as follows: Prominent features of Received Pronunciation are not consistently represented in the textbook audio materials. At the textbook level, the three textbooks display pronunciation features consistent with Received Pronunciation only in the vowels GOOSE and TRAP. The textbooks *English G Access 2* and *Camden Town 2* also show a Received Pronunciation articulation in BATH. START is only consistent with Received Pronunciation in the textbook *Camden Town 2*. The pronunciation of the vowels LOT and CLOTH is inconsistent with Received Pronunciation in all three textbooks.

The analysis at the individual speaker level provided a more detailed investigation of the representation of Received Pronunciation: similar patterns to the textbook level were revealed. While LOT and CLOTH are mostly inconsistent with Received Pronunciation for all speakers, the remaining four vowels (GOOSE, TRAP, BATH, and START) are consistent with Received Pronunciation to varying degrees across the speakers. Overall, the three textbooks display some features of Received Pronunciation in the monophthongs GOOSE, LOT, CLOTH, TRAP, BATH, and START, but not consistently.

Research Question 2 can be answered as follows: Prominent features of General American are not consistently represented in the textbook audio materials. At the textbook level, the three textbooks display pronunciation features consistent with General American in all three vowels BATH, LOT, and CLOTH, but do not show full rhoticity in the vowels NEAR, SQUARE, and CURE.

The analysis at the individual speaker level provided a more detailed investigation of the representation of General American. The analysis at the speaker level revealed more variation in the data: Several speakers from all three textbooks produce BATH, LOT, and CLOTH consistent with General American pronunciation patterns. The speakers *from English G Access 4* are also fully rhotic, supporting the impression of a consistent General American accent. Several speakers from the other two textbooks also display full rhoticity in the lexical sets NEAR, SQUARE, and CURE, while others display variable rhoticity, which is not expected in a General American accent. Overall, the three textbooks display some features of General American in the monophthongs BATH, LOT, and CLOTH, and rhoticity in NEAR, SQUARE, and CURE, but, again, not consistently.

Research Question 3 can be answered as follows: Prominent features of Australian English are not consistently represented in the textbook audio materials. At the textbook level, the two textbooks display pronunciation features consistent with Australian English to varying degrees in the first targets of the diphthongs FACE, PRICE, and MOUTH: the textbook *English G Access 5* consistently displays features of General Australian English in the three diphthongs, while the textbook *Green Line 5* only does so in the first targets of the vowels PRICE and MOUTH.

The analysis at the individual speaker level provided a more detailed investigation of the representation of Australian English. The analysis at the speaker level revealed more variation in the data: the textbook speakers from both textbooks display a strong tendency toward General Australian English, but also display features of Broad Australian English or, in the case of one speaker from *Green Line 5*, also Cultivated Australian English. Some

speakers from both textbooks produce all three FACE, PRICE, and MOUTH consistent with Australian English pronunciation patterns. Other speakers display an Australian English accent feature in only one or two of the features. Overall, the two textbooks display some features of Australian English in the first targets of the diphthongs FACE, PRICE, and MOUTH, but not consistently. Both textbooks rather present General Australian English.

Overall, the three textbook series' display some features of the accent under discussion, but not consistently across all speakers. At a more global, textbook-level, all textbooks seem to be more consistent at presenting salient features of the three accents Received Pronunciation, General American, and Australian English as the analysis at the speaker level would suggest. However, no accent is presented consistently as would be expected based on the linguistic descriptions presented in Chapter 2. The corpus-phonological approach to textbook analysis proved beneficial in answering the research questions.

Further research on this topic could focus on (i) investigating the representation of other varieties in these textbook series, (ii) investigate the representation of the same varieties (or also others) in the textbook series for the other secondary school forms in Germany (e.g *Realschule* or *Gesamtschule*), and, importantly, (iii) the perception of learners of these varieties based on the textbook materials to investigate in how far learners perceive differences between but also within these accents.

References

Ashby, P. (2011) *Understanding Phonetics*. (Understanding Language Series). London: Hodder Education.

Asprey, E. (2007) 'Investigating residual rhoticity in a non-rhotic accent', *Leeds Working Papers in Linguistics and Phonetics*, 12, pp. 78–101. Available at: https://www.latl.leeds.ac.uk/article/investigating-residual-rhoticity-in-a-non-rhotic-accent/ (Accessed: 18 September 2023).

Bartlett, C.M. (2002) *The Southland Variety of New Zealand English: Postvocalic /r/ and the BATH vowel*. Thesis, Doctor of Philosophy. University of Otago. Available at: http://hdl.handle.net/10523/3621 (Accessed: 5 December 2023).

Bayyurt, Y. (2018) 'Issues of intelligibility in world Englishes and EIL contexts', *World Englishes*, 37, pp. 407–415. doi: 10.1111/weng.12327

Becker, K. (2014) '(r) we there yet? The change to rhoticity in New York City English', *Language Variation and Change*, 26(02), pp. 141–168. doi: 10.1017/S0954394514000064

Bhowmik, S.K. (2015) 'Wold Englishes and English Language Teaching: A pragmatic and humanistic approach', *Colombian Applied Linguistics Journal*, 17(1), pp. 142–157. doi: 10.14483/udistrital.jour.calj.2015.1.a10

Bieswanger, M. (2012) 'Varieties of English in the Curriculum', in Schröder, A., Busse, U. and Schneider, R. (eds.) *Codification, Canons, and Curricula: Description and Prescription in Language and Literature.* Bielefeld: Aisthesis, pp. 359–371.

Boersma, P. (2020) *Sound: To Formant (burg)…* Available at: https://www.fon.hum.uva.nl/praat/manual/Sound__To_Formant__burg____.html (Accessed: 14 June 2022).

Boersma, P. and Weenink, D. (2020) *Praat: Doing Phonetics by Computer* (Version 6.1.16). Amsterdam. Available at: www.praat.org (Accessed: 11 December 2023).

Borowsky, T. (2001) 'The vocalisation of dark *l* in Australian English', in Blair, D. and Collins, P. (eds.) *English in Australia.* (Varieties of English Around the World, 26). Amsterdam: John Benjamins Publishing Company, pp. 69–87.

Boyce, S. and Espy-Wilson, C.Y. (1997) 'Coarticulatory stability in American English /r/', *The Journal of the Acoustical Society of America*, 101(6), pp. 3741–3753. doi: 10.1121/1.418333

Bradley, D. (2008) 'Regional characterisctics of Australian English: Phonology', in Burridge, K. and Kortmann, B. (eds.) *The Pacific and Australasia.* (Varieties of English, 3). Berlin: Mouton de Gruyter, pp. 111–123.

Bradley, D. and Bradley, M. (2001) 'Changing Attitudes to Australian English', in Blair, D. and Collins, P. (eds.) *English in Australia.* (Varieties of English Around the World, 26). Amsterdam: John Benjamins Publishing Company, pp. 271–285.

Brinckmann, C. (2014) 'Praat Scripting', in Durand, J., Gut, U. and Kristoffersen, G. (eds.) *The Oxford Handbook of Corpus Phonology.* (Oxford Handbooks in Linguistics). Oxford: Oxford University Press, pp. 361–379.

Bruthiaux, P. (2003) 'Squaring the circles: Issues in modeling English worldwide', *International Journal of Applied Linguistics*, 13(2), pp. 159–178. doi: 10.1111/1473-4192.00042

Buschfeld, S. and Kautzsch, A. (2020) 'Theoretical Models of English as World Language', in Schreier, D., Hundt, M. and Schneider, E.W. (eds.) *The Cambridge Handbook of World*

Englishes. (Cambridge handbooks in language and linguistics). Cambridge: Cambridge University Press, pp. 51–71.

Butcher, A. (2006) 'Formant frequencies of /hVd/ vowels in the speech of South Australian females', *Proceedings of the 11th Australian International Conference on Speech Science & Technology*, University of Auckland, New Zealand, December 6-8, pp. 449–453.

Butcher, A. (2008) 'Linguistic Aspects of Australian Aboriginal English', *Clinical Linguistics and Phonetics*, 22(8), pp. 625–642.

Chan, J.Y.H. (2020) 'Towards English as an international language: The evolving ELT curricula and textbooks in Hong Kong', *International Journal of Applied Linguistics*, 30, pp. 244–263.

Claussen, S. *et al.* (eds.) (2015a) *Camden Town 4*. Braunschweig: Diesterweg (Camden Town Gymnasium).

Claussen, S. *et al.* (2015b) 'Audio-CD mit DVD für Lehrkräfte', *Camden Town 4* [CD]. Braunschweig: Diesterweg.

Collins, B. and Mees, I.M. (2006) *Practical Phonetics and Phonology: A Resource Book for Students.* (Routledge English Language Introductions). London: Routledge.

Cox, F. (1998) 'The Bernard data revisited∗', *Australian Journal of Linguistics*, 18(1), pp. 29–55. doi: 10.1080/07268609808599557

Cox, F. (2006) 'The Acoustic Characteristics of /hVd/ Vowels in the Speech of some Australian Teenagers', *Australian Journal of Linguistics*, 26(2), pp. 147–179. doi: 10.1080/07268600600885494

Cox, F. (2008) 'Vowel transcription systems: An Australian perspective', *International Journal of Speech-Language Pathology*, 10(5), pp. 327–333. doi: 10.1080/17549500701855133

Cox, F. and Fletcher, J. (2017) *Australian English Pronunciation and Transcription*. 2nd edn. Cambridge: Cambridge University Press.

Cox, F. and Palethorpe, S. (2007) 'Illustrations of the IPA: Australian English', *Journal of the International Phonetic Association*, 37(3), pp. 341–350.

Cox, F. and Palethorpe, S. (2008) 'Reversal of short front vowel raising in Australian English', *Interspeech 2008: 9th Annual Conference of the International Speech Communication Association*, Brisbane, 22-26 September 2008, pp. 342–345.

Cox, F. and Palethorpe, S. (2010) 'Broadness variation in Australian English speaking females', *Proceedings of the 13th Australasian International Conference on Speech Science and Technology,* Melbourne, Dec. 14, 2010: Australasian Speech Science and Technology Association (ASSTA), pp. 175–178.

Cox, F. and Palethorpe, S. (2012) 'Standard Australian English: the sociostylistic broadness continuum', in Hickey, R. (ed.) *Standards of English: Codified Varieties Around the World.* (Studies in English Language). Cambridge: Cambridge University Press, pp. 294–317.

Cruttenden, A. (2014) *Gimson's Pronunciation of English*. 8th edn. London: Routledge.

D'Souza, J. (1999) 'Afterword', *World Englishes*, 18(2), pp. 271–274.

Davydova, J. (2015) 'A Study in the Perception of Native and Non-Native Englishes by German Learners', *Journal of Linguistics and Language Teaching*, 6(1), pp. 89–117.

Delais-Roussarie, E. and Post, B. (2014) ‘Corpus Annotation’, in Durand, J., Gut, U. and Kristoffersen, G. (eds.) *The Oxford Handbook of Corpus Phonology.* (Oxford Handbooks in Linguistics). Oxford: Oxford University Press, pp. 46–88.

Dickson, V. and Hall-Lew, L. (2017) ‘Class, Gender, and Rhoticity: The Social Stratification of Non-Prevocalic /r/ in Edinburgh Speech’, *Journal of English Linguistics*, 45(3), pp. 229–259. doi: 10.1177/0075424217718024

Fabricius, A. (2002) ‘Weak vowels in modern RP: An acoustic study of happ y -tensing and kit /schwa shift’, *Language Variation and Change*, 14(2), pp. 211–237. doi: 10.1017/S0954394502142037

Fabricius, A. (2007) ‘Variation and change in the TRAP and STRUT vowels of RP: A real time comparison of five acoustic data sets’, *Journal of the International Phonetic Association*, 37(3), pp. 293–320. doi: 10.1017/S002510030700312X

Fabricius, A. and Mortensen, J. (2013) ‘Language ideology and the notion of construct resources: a case study of modern RP’, in Kristiansen, T. and Grondelaars, S. (eds.) *Language (de)standardisation in late modern Europe: Experimental studies.* (Standard language ideology in contemporary Europe, 2). Oslo: Novus Press.

Galloway, N. (2017) *Global Englishes and Change in English Language Teaching: Attitudes and Impact*. Abingdon, Oxon: Routledge.

Galloway, N. and Rose, H. (2015) *Introducing Global Englishes*. London: Routledge.

GeoGebra® (2024) *GeoGebra Classic* (Version 6.0.843.0). Available at: www.geogebra.org.

Gibson, A.M. (2005) ‘Non-prevocalic /r/ in New Zealand hip-hop’, *New Zealand English Journal*, 19, pp. 5–14.

Gut, U. and Voormann, H. (2014) ‘Corpus Design’, in Durand, J., Gut, U. and Kristoffersen, G. (eds.) *The Oxford Handbook of Corpus Phonology.* (Oxford Handbooks in Linguistics). Oxford: Oxford University Press, pp. 13–26.

Guy, G.R. (2018) ‘Saks vs. Macys: (r-1) marches on in New York City department stores’, *University of Pennsylvania Working Papers in Linguistics*, 24(2), pp. 49–55.

Hanus, P. *et al.* (eds.) (2013a) *Camden Town 2*. Braunschweig: Diesterweg (Camden Town Gymnasium).

Hanus, P. *et al.* (2013b) ‘Audio-CD mit DVD für Lehrkräfte’, *Camden Town 2* [CD]. Braunschweig: Diesterweg.

Harrington, J. (2010) *Phonetic Analysis of Speech Corpora*. Chichester, West Sussex: Wiley Blackwell.

Harrington, J. and Cassidy, S. (1999) *Techniques in Speech Acoustics*. Dordrecht: Springer Science + Business Media.

Harrington, J., Cox, F. and Evans, Z. (1997) ‘An acoustic phonetic study of broad, general, and cultivated Australian English vowels’, *Australian Journal of Linguistics*, 17(2), pp. 155–184. doi: 10.1080/07268609708599550

Harrington, J., Kleber, F. and Reubold, U. (2011) ‘The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English’, *Journal of the International Phonetic Association*, 41(2), pp. 137–156. doi: 10.1017/S0025100310000265

Harris, J. (2013) ‘Wide-domain r -effects in English’, *Journal of Linguistics*, 49(2), pp. 329–365. doi: 10.1017/S0022226712000369

Hartmann, J. (2021) ‘Tomorrow's teacher's perceptions of Global Englishes’, in Callies, M. *et al.* (eds.) *Glocalising Teaching English as an International Language: New Perspectives for Teaching and Teacher Education in Germany.* London: Routledge, pp. 46–62.

Henry, L. and Wickham, H. (2020) *_purrr: Functional Programming Tools_* (Version 0.3.4). Available at: https://cran.r-project.org/package=purrr.

Heselwood, B. and Plug, L. (2011) ‘The role of F2 and F3 in the perception of rhoticity: Evidence from listening experiments’, *Proceedings of the 17th International Concress of Phonetic Sciences*, City University of Hong Kong. Available at: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Heselwood/Heselwood.pdf (Accessed: 18 September 2023).

Hinton, M. (2015) ‘Changes in Received Pronunciation: Diachronic Case Studies’, *Research in Language*, 13(1), pp. 21–37. doi: 10.1515/rela-2015-0010

Horvath, B.M. (2008) ‘Australian English: phonology’, in Burridge, K. and Kortmann, B. (eds.) *The Pacific and Australasia.* (Varieties of English, 3). Berlin: Mouton de Gruyter, pp. 89–110.

Horvath, B.M. and Horvath, R.J. (2002) ‘The geolinguistics of /l/ vocalization in Australia and New Zealand’, *Journal of Sociolinguistics*, 6(3), pp. 319–346.

IPA (1999) ‘see: The International Phonetic Association’.

Jansen, S. and Mompean, J.A. (2023) ‘GOOSE-fronting in Received Pronunciation across time: A trend study’, *Language Variation and Change*, 35(1), pp. 55–77. doi: 10.1017/S0954394523000017

Jauriberry, T., Sock, R. and Hamm, A. (2015) ‘Phonetic Variation in Standard Scottish English: Rhotics in Dundee’, *ICPhS, 18th International Congress of Phonetic Sciences.* International Congress of Phonetic Sciences, Glasgow, 10-14 August 2015: University of Glasgow.

Jenkins, J. (2015) *Global Englishes: A resource book for students*. London: Routledge.

Johnson, K. (2012) *Acoustic and Auditory Phonetics*. 3rd edn. Chichester, West Sussex: Wiley-Blackwell.

Jones, D. (2011) *Cambridge English Pronouncing Dictionary*. Roach, Peter; Setter, Jane, and John Esling (eds.). 18th edn. Cambridge: Cambridge University Press.

Kachru, B.B. (1985) ‘Standards, codification and sociolinguistics realism: The English language in the outer circle.’, in Quirk, R. and Widdowson, H.G. (eds.) *English in the world: Teaching and learning the language and literatures.* Cambridge: Cambridge University Press, pp. 11–30.

Kachru, B.B. (1992) ‘World Englishes: Approaches, issues and resources’, *Language Teaching*, 25(1), pp. 1–14. doi: 10.1017/S0261444800006583

Kent, R.D. and Read, C. (1992) *The Acoustic Analysis of Speech*. San Diego, California: Singular Publishing Group, Inc.

King, H. and Ferragne, E. (2020) ‘Loose lips and tongue tips: The central role of the /r/-typical labial gesture in Anglo-English’, *Journal of Phonetics*, 80, pp. 1–19. doi: 10.1016/j.wocn.2020.100978

Kopperoinen, A. (2011) ‘Accents of English as a lingua franca: a study of Finnish textbooks’, *International Journal of Applied Linguistics*, 21(1), pp. 71–93.

Kortmann, B. and Upton, C. (2004) ‘Introduction: Varieties of English in the British Isles’, in Schneider, E.W. (ed.) *A Handbook of Varieties of English: Phonology.* (A Handbook of Varieties of English, 1). Berlin: Mouton de Gruyter, pp. 25–33.

Kövecses, Z. (2000) *American English: An Introduction*. Peterborough, Ontario: Broadview Press.

Kretzschmar, W.A., Jr. (2008) ‘Standard American English pronunciation’, in Schneider, E.W. (ed.) *The Americas and the Caribbean.* (Varieties of English, 2). Berlin: Mouton de Gruyter, pp. 37–51.

Kuecker, K., Lockenvitz, S. and Müller, N. (2015) ‘Amount of rhoticity in schwar and in vowel+/r/ in American English’, *Clinical Linguistics & Phonetics*, 29(8-10), pp. 623–629. doi: 10.3109/02699206.2015.1044674

Labov, W. (1997) ‘The Social Stratification of (r) in New York City Department Stores’, in Coupland, N. and Jaworski, A. (eds.) *Sociolinguistics: A Reader.* (Modern Linguistics Series). New York: Macmillan Education, pp. 168–178.

Ladefoged, P. and Johnson, K. (2011) *A Course in Phonetics: International Edition*. 6th edn. Boston, MA: Wadsworth Cengage Learning.

Ladefoged, P. and Maddieson, I. (1996) *The Sounds of the World's Languages*. Oxford, UK: Blackwell.

Li, Z. and Kabak, B. (2017) ‘Rhoticity in Chinese English: An experimental investigation on the realization of the variant (r) in an Expanding Circle variety’, *Alicante Journal of English Studies*, 30, pp. 61–92.

Lieberman, P. and Blumstein, S.E. (1988) *Speech Physiology, Speech Perception, and Acoustic Phonetics*. (Cambridge Studies in Speech Science and Communication). Cambridge: Cambridge University Press.

Lindqvist, N. and Soler, J. (2022) ‘World Englishes in ELT textbooks in Swedish upper-secondary schools’, *World Englishes*, pp. 1–16. doi: 10.1111/weng.12599

Lonergan, J. and Cox, F. (2010) ‘Is there any Evidence of Rhoticity in Historical Australian English?’ *Proceedings of the Australian Linguistic Society Conference 2008.* Australian Linguistic Society, University of Sydney, Sydney, 2 Jul 2008 - 4 Jul 2008, pp. 1–17.

Mair, C. (2008) *English linguistics: An Introduction*. (Bachelor-Wissen). Tübingen: Narr.

Mair, C. (2015) *English Linguistics: An Introduction*. 3rd edn. (Bachelor-Wissen). Tübingen: Narr Francke Attempto.

Marsden, S. (2017) ‘Are New Zealanders “rhotic”? The dynamics of rhoticity in New Zealand's small towns’, *English World-Wide*, 38(3), pp. 275–304. doi: 10.1075/eww.38.3.02mar

Martin, C. (2017) *_ggConvexHull: Add a convex hull geom to ggplot2_* (Version 0.1.0). Available at: http://github.com/cmartin/ggConvexHull.

Mather, P.-A. (2012) ‘The Social Stratification of /r/ in New York City: Labov's Department Store Study Revisited’, *Journal of English Linguistics*, 40(4), pp. 338–356.

Matsuda, A. (2002) ‘Representation of Users and Uses of English in Beginning Japanese EFL Textbooks’, *JALT Journal*, 24(2), pp. 182–200. doi: 10.37546/JALTJJ24.2-5

Matsuda, A. (2012) ‘Teaching Materials in EIL’, in Alsagoff, L. *et al.* (eds.) *Principles and Practices for Teaching English as an International Language.* New York: Routledge, pp. 168–185.

Matsuda, A. (2013) ‘World Englishes and Language Pedagogy’, in Chapelle, C. (ed.) *The Encyclopedia of Applied Linguistics.* Malden, MA: Blackwell Publishing Ltd, pp. 1–7. Available at: https://onlinelibrary.wiley.com/doi/10.1002/9781405198431.wbeal1291 (Accessed: 4 April 2022).

Matsuda, A. and Friedrich, P. (2012) ‘Selecting an instructional variety for an EIL curriculum’, in Matsuda, A. (ed.) *Principles and Practices of Teaching English as an International Language.* Bristol: Multilingual Matters, pp. 17–27.

Matsuda, A. and Matsuda, P.K. (2018) ‘Teaching English as an international language: a WE-informed paradigm for English langauge teaching’, in Low, E.L. and Pakir, A. (eds.) *World Englishes: Rethinking Paradigms.* (Routledge Studies in World Englishes). London: Routledge, pp. 64–77.

Matthews, P.H. (2014a) ‘Accent’. *The Concise Oxford Dictionary of Linguistics*. Available at: https://www.oxfordreference.com/display/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-31?rskey=En5PDj&result=5 (Accessed: 11 February 2025).

Matthews, P.H. (2014b) ‘Dialect’. *The Concise Oxford Dictionary of Linguistics*. Available at: https://www.oxfordreference.com/display/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-893?rskey=HEXM1q&result=1 (Accessed: 11 February 2025).

Matthews, P.H. (2014c) ‘Variety’. *The Concise Oxford Dictionary of Linguistics*. Available at: https://www.oxfordreference.com/display/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-3583?rskey=d7EB7t&result=1 (Accessed: 11 February 2025).

Meer, P. (2021) ‘Global Englishes in the secondary school curriculum in Germany: A comparative analysis of the English language curricula of the federal states’, in Callies, M. *et al.* (eds.) *Glocalising Teaching English as an International Language: New Perspectives for Teaching and Teacher Education in Germany.* London: Routledge, pp. 85–103.

Meer, P. *et al.* (2021) ‘Rhotics in Standard Scottish English’, *English World-Wide*, 42(2), pp. 121–144. doi: 10.1075/eww.00070.mee

Meer, P., Hartmann, J. and Rumlich, D. (2021) ‘Attitudes of German high school students toward different varieties of English’, *Applied Linguistics*, 48, pp. 1–26. doi: 10.1093/applin/amab046

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (MSW NRW) (2007) *Kernlehrplan für den verkürzten Bildungsgang des Gymnasiums - Sekundarstufe I (G8) in Nordrhein-Westfalen: Englisch*. Frechen: Ritterbach Verlag.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (MSW NRW) (2017) *Verzeichnis der zugelassenen Lernmittel: Gymnasium Sekundarstufe I*. Available at: https://www.schulministerium.nrw.de/docs/Schulsystem/Unterricht/Lernmittel/GymnasiumG8/index.html (Accessed: 4 January 2019).

Modiano, M. (1999) ‘International English in the global village’, *English Today*, 15(2), pp. 22–28. doi: 10.1017/S026607840001083X

MSW NRW (2007) ‘see Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen’.

Mugglestone, L. (2017) ‘Received Pronunciation’, in Brinton, L.J. and Bergs, A. (eds.) *The History of English: Varieties of English.* (Mouton Reader, 5). Berlin: De Gruyter Mouton, pp. 151–168.

Müller, K. and Wickham, H. (2020) *_tibble: Simple Data Frames_* (Version 3.0.3). Available at: https://cran.r-project.org/package=tibble.

Murphy, M.L. (2016) ‘British English? American English? Are there such things?’ *English Today*, 32(2), pp. 4–7. doi: 10.1017/S0266078416000067

Naji Meidani, E. and Pishghadam, R. (2013) ‘Analysis of English language textbooks in the light of English as an International Language (EIL): A comparative study’, *International Journal of Research Studies in Language Learning*, 2(2), pp. 83–96.

Nelson, C.L. (2011) *Intelligibility in World Englishes: Theory and Application*. (ESL and applied linguistics professional series). New York: Routledge.

Nordlund, M. (2016) ‘EFL textbooks for young learners: A comparative analysis of vocabulary’, *Education Inquiry*, 7(1), pp. 47–68. doi: 10.3402/edui.v7.27764

R Core Team (2020) *R: A language and environment for statistical computing* (Version 4.0.2). Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.

Rademacher, J. (ed.) (2014a) *Access 2*. Berlin: Cornelsen (English G).

Rademacher, J. (2014b) ‘Audio-CDs’, *Access 2* [CD]. Berlin: Cornelsen.

Rademacher, J. (ed.) (2016a) *Access 4*. Berlin: Cornelsen (English G).

Rademacher, J. (2016b) ‘Audio-CDs’, *Access 4* [CD]. Berlin: Cornelsen.

Rademacher, J. (ed.) (2017a) *Access 5: Abschlussband*. Berlin: Cornelsen (English G).

Rademacher, J. (2017b) ‘Audio-CDs’, *Access 5*. Berlin: Cornelsen.

Redzwan, S. (2016) ‘Rhoticity in Brunei and Singapore English’, *Southeast Asia: A Multidisciplinary Journal*, 16, pp. 129–137.

Richards, J.C. (2014) ‘The ELT Textbook’, in Garton, S. and Graves, K. (eds.) *International Perspectives on Materials in ELT.* London: Palgrave Macmillan UK, pp. 19–36.

Roach, P. (2001) *Phonetics*. (Oxford Introductions to Language Study). Oxford: Oxford University Press.

Roach, P. (2004) ‘Illustrations of the IPA: British English: Received Pronunciation’, *Journal of the International Phonetic Association*, 34(2), pp. 239–245.

Robb, M.P. and Gillon, G.T. (2007) ‘Speech rates of New Zealand English- and American English-speaking children’, *Advances in Speech Language Pathology*, 9(2), pp. 1–8. doi: 10.1080/14417040601013695

Rosner, B.S. and Pickering, J.B. (1994) *Vowel Perception and Production*. (Oxford Psychology Series, 23). Oxford: Oxford University Press.

Scheiwe, L. (2022) ‘Evaluating Accents of English in ELT Textbooks Used at German Secondary Schools’, *Anglistik*, 33(2), pp. 61–75. doi: 10.33675/ANGL/2022/2/8

Schildhauer, P., Schulte, M. and Zehne, C. (2020) ‘Global Englishes in the Classroom: From Theory to Practice’, *PFLB*, 2(4), pp. 26–40. doi: 10.4119/pflb-3435

Schilk, M. and Pickert, L. (2022) ‘Rhoticity in Southern New Zealand English: An acoustic analysis of the QuakeBox database’, in Flach, S. and Hilpert, M. (eds.) *Broadening the Spectrum of Corpus Linguistics: New approaches to variability and change.* (Studies in Corpus Linguistics, 105). Amsterdam: John Benjamins Publishing Company, pp. 69–89.

Schmitt, H. (2016) *Teaching English Pronunciation: A textbook for the German-speaking countries*. Heidelberg: Universitätsverlag Winter.

Schneider, E.W. (2011) *English Around the World: An Introduction*. (Introductions to the English Language). Cambridge: Cambridge University Press.

Seargeant, P. (2010) ‘Naming and defining in world Englishes’, *World Englishes*, 29(1), pp. 97–113. doi: 10.1111/j.1467-971X.2009.01627.x

Seargeant, P. (2012) *Exploring World Englishes: Language in a Global Context*. (Routledge Introductions to Applied Linguistics). London: Routledge.

Slowikowski, K. (2020) *_ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'_* (Version 0.8.2). Available at: https://cran.r-project.org/package=ggrepel.

Stuart-Smith, J., Lawson, E. and Scobbie, J.M. (2014) ‘Derhoticisation in Scottish English: A sociophonetic journey’, in Celata, C. and Calamai, S. (eds.) *Advances in Sociophonetics.* (Studies in Language Variation, 15). Amsterdam: John Benjamins, pp. 59–96.

Sung, C.C.M. (2016) ‘Exposure to multiple accents of English in the English Language Teaching classroom: from second language learners’perspectives’, *Innovation in Language Learning and Teaching*, 10(3), pp. 190–205. doi: 10.1080/17501229.2014.936869

Syrbe, M. and Rose, H. (2018) ‘An Evaluation of the Global Orientation of English Textbooks in Germany’, *Innovation in Language Learning and Teaching*, 12(2), pp. 152–163. doi: 10.1080/17501229.2015.1120736

Tan, Y.-Y. (2012) ‘To r or not to r: Social correlates of /ɹ/ in Singapore English’, *International Journal of the Sociology of Language*, 2012(218), pp. 1–24. doi: 10.1515/ijsl-2012-0057

The International Phonetic Association (ed.) (1999) *A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Thomas, E.R. (2011) *Sociophonetics: An Introduction*. Basingstoke Hampshire: Palgrave Macmillan.

Tollfree, L. (2001) ‘Variation and change in Australian English consonants’, in Blair, D. and Collins, P. (eds.) *English in Australia.* (Varieties of English Around the World, 26). Amsterdam: John Benjamins Publishing Company, pp. 45–67.

Tottie, G. (2002) *An Introduction to American English*. Malden, MA: Blackwell.

Trudgill, P. (1999) ‘Standard English: What it isn't’, in Bex, T. and Watts, R.J. (eds.) *Standard English: The Widening Debate.* London: Routledge, pp. 117–128.

Trudgill, P. and Hannah, J. (2008) *International English: A Guide to the Varieties of Standard English*. 5th edn. London: Hodder Education.

Upton, C. (2008) ‘Received Pronunciation’, in Kortmann, B. and Upton, C. (eds.) *The British Isles.* (Varieties of English, 1). Berlin: Mouton de Gruyter, pp. 237–252.

van Son, R.J.J.H. and Pols, L.C.W. (1990) ‘Formant frequencies of Dutch vowels in a text, read at normal and fast rate’, *The Journal of the Acoustical Society of America*, 88(4), pp. 1683–1693. doi: 10.1121/1.400243

Vettorel, P. and Lopriore, L. (2013) ‘Is there ELF in ELT coursebooks?’ *Studies in Second Language Learning and Teaching*, 3(4), pp. 483–540. doi: 10.14746/ssllt.2013.3.4.3

Watson, C.I., Harrington, J. and Evans, Z. (1998) ‘An acoustic comparison between New Zealand and Australian English vowels’, *Australian Journal of Linguistics*, 18(2), pp. 185–207. doi: 10.1080/07268609808599567

Weisshaar, H. (ed.) (2015a) *Green Line 2*. Stuttgart: Ernst Klett Verlag.

Weisshaar, H. (2015b) ‘4 Lehrer-Audio-CDs zum Hörverstehen’, *Green Line 2* [CD]. Stuttgart: Ernst Klett Verlag.

Weisshaar, H. (ed.) (2017a) *Green Line 4*. Stuttgart: Ernst Klett Verlag.

Weisshaar, H. (2017b) '3 Lehrer-Audio-CDs zum Hörverstehen', *Green Line 4* [CD]. Stuttgart: Ernst Klett Verlag.

Weisshaar, H. (ed.) (2018a) *Green Line 5*. Stuttgart: Ernst Klett Verlag.

Weisshaar, H. (2018b) '3 Lehrer-Audio-CDs zum Hörverstehen', *Green Line 5* [CD]. Stuttgart: Ernst Klett Verlag.

Wells, J.C. (1982a) *An Introduction*. (Accents of English, 1). Cambridge: Cambridge University Press.

Wells, J.C. (1982c) *Beyond the British Isles*. (Accents of English, 3). Cambridge: Cambridge University Press.

Wells, J.C. (2008) *Longman Pronunciation Dictionary*. 3rd edn. Harlow: Pearson Longman.

Wickham, H. (2020a) *_forcats: Tools for Working with Categorical Variables (Factors)_* (Version 0.5.0). Available at: https://cran.r-project.org/package=forcats.

Wickham, H. (2020b) *_tidyr: Tidy Messy Data_* (Version 1.1.2). Available at: https://cran.r-project.org/package=tidyr.

Wickham, H. (2016) *_ggplot2: Elegant Graphics for Data Analysis*. New York: Springer. Available at: https://ggplot2.tidyverse.org/.

Wickham, H. (2019) *_stringr: Simple, Consistent Wrappers for Common String Operations_* (Version 1.4.0). Available at: https://cran.r-project.org/package=stringr.

Wickham, H. *et al.* (2019) 'Welcome to the Tidyverse', *Journal of Open Source Software*, 4(43), p. 1686. doi: 10.21105/joss.01686

Wickham, H. and Bryan, J. (2019) *_readxl: Read Excel Files_* (Version 1.3.1). Available at: https://cran.r-project.org/package=readxl.

Wickham, H. *et al.* (2020) *_dplyr: A Grammar of Data Manipulation* (Version 1.0.2). Available at: https://cran.r-project.org/package=dplyr.

Wickham, H., Hester, J. and François, R. (2018) *_readr: Read Rectangular Text Data_* (Version 1.3.1). Available at: https://cran.r-project.org/package=readr.

Wolters, M. (2015) *ICPhS. 18th International Congress of Phonetic Sciences*. Glasgow, 10-14 August 2015: University of Glasgow.

Zsiga, E.C. (2013) *The Sounds of Language: An Introduction to Phonetics and Phonology*. (Linguistics in the World). Chichester, West Sussex: Wiley-Blackwell.

Appendices

Appendix 1

Vowel plots

Appendix 2

Vowel plots with articulatory categorisation

Appendix 3

Praat scripts

Appendix 4

R Scripts

Appendix 5

Quantitative Data – features of the analysis

Appendix 6

Quantitative Data – additional monophthongs

Appendix 7

Overview of the digital Appendix

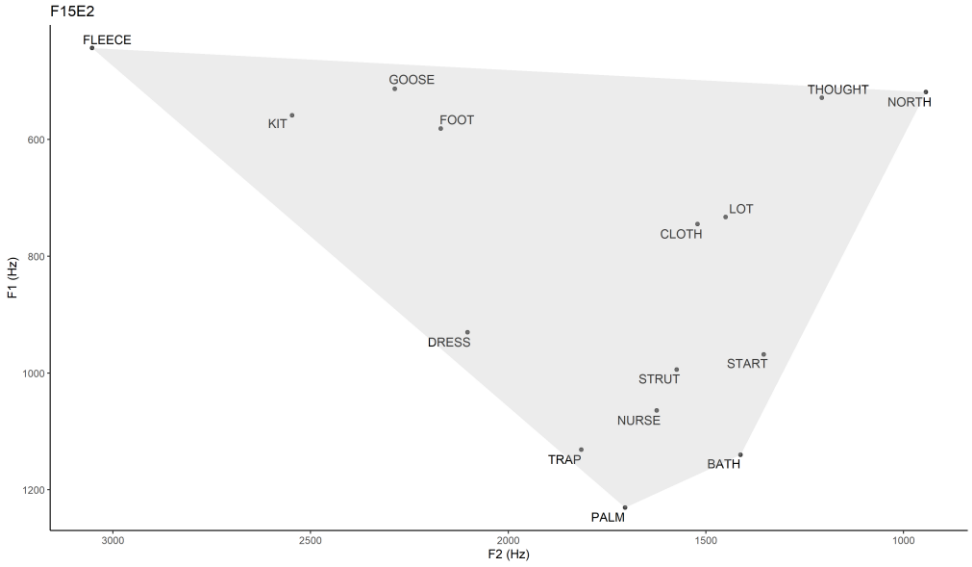
Appendix 1

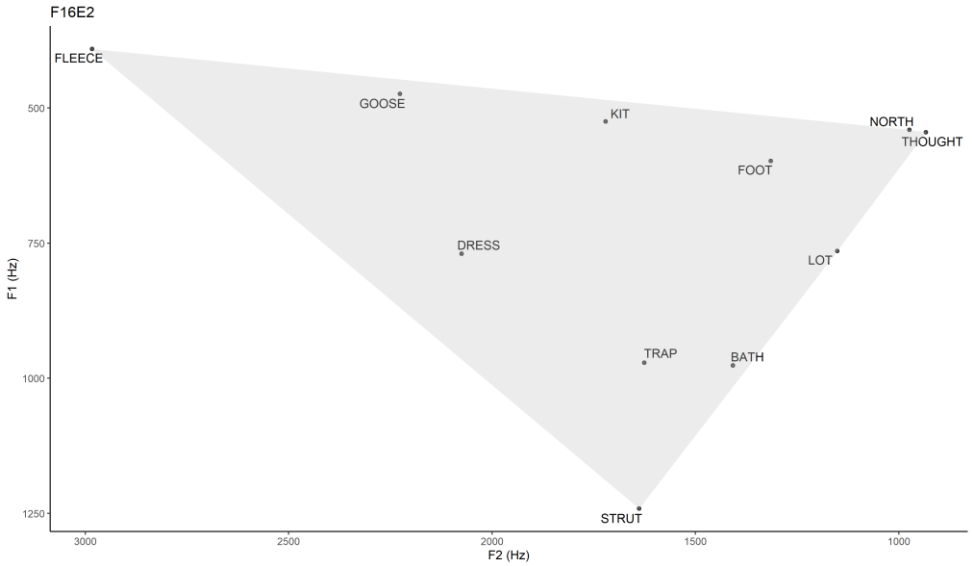
Vowel plots

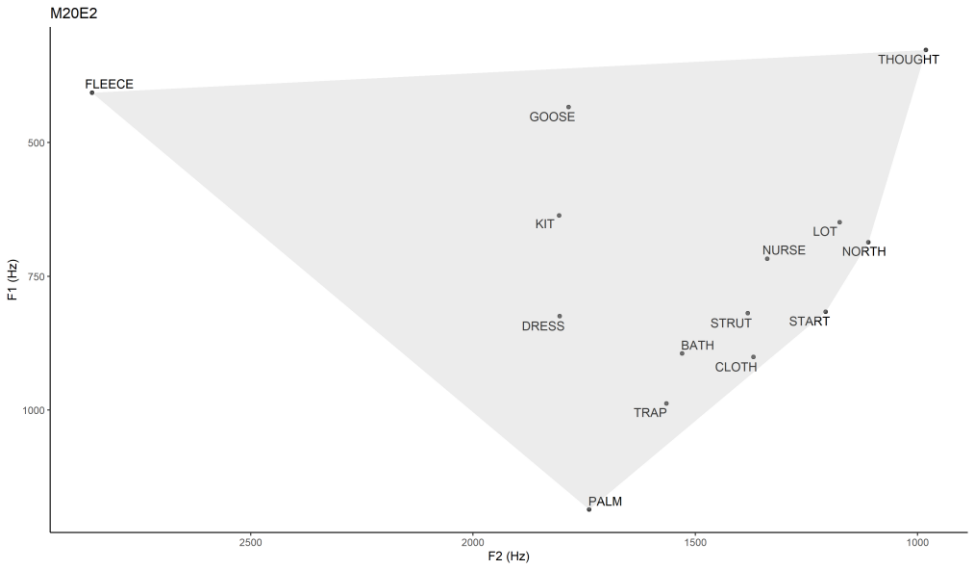
This appendix includes the vowel plots of the individual speakers based on their individual mean formant values of each vowel. The speaker is indicated in the top left corner of each vowel plot.

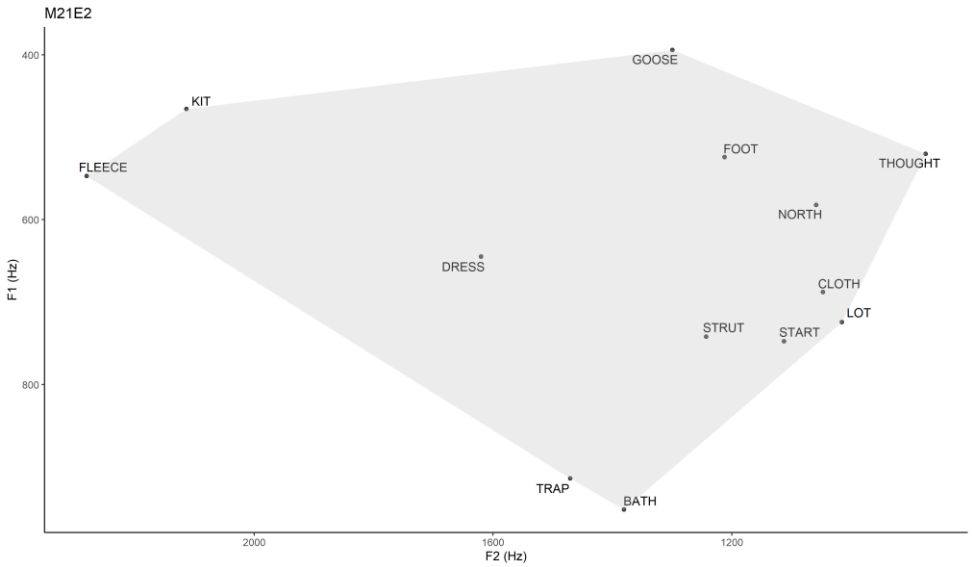
Received Pronunciation sub-corpus

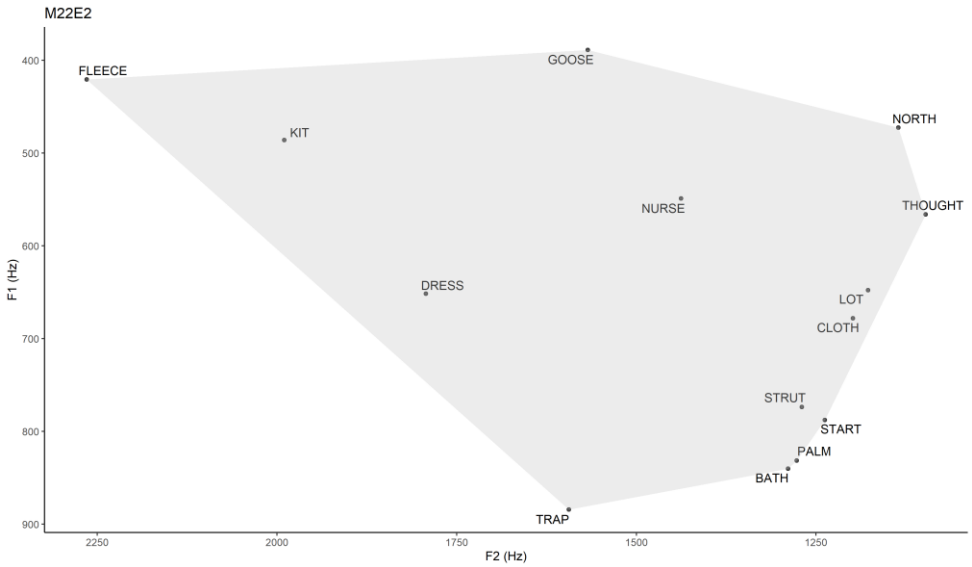
English G Access 2

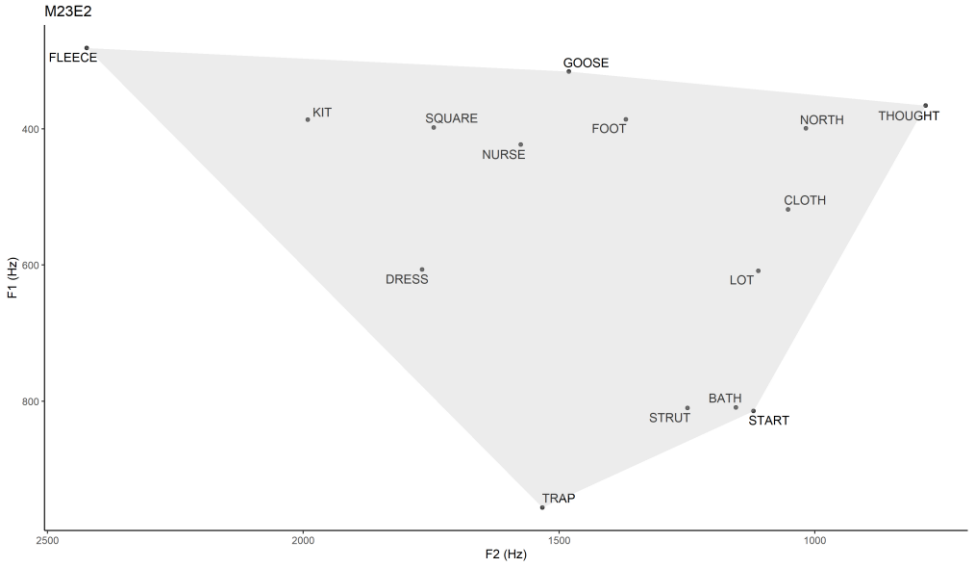


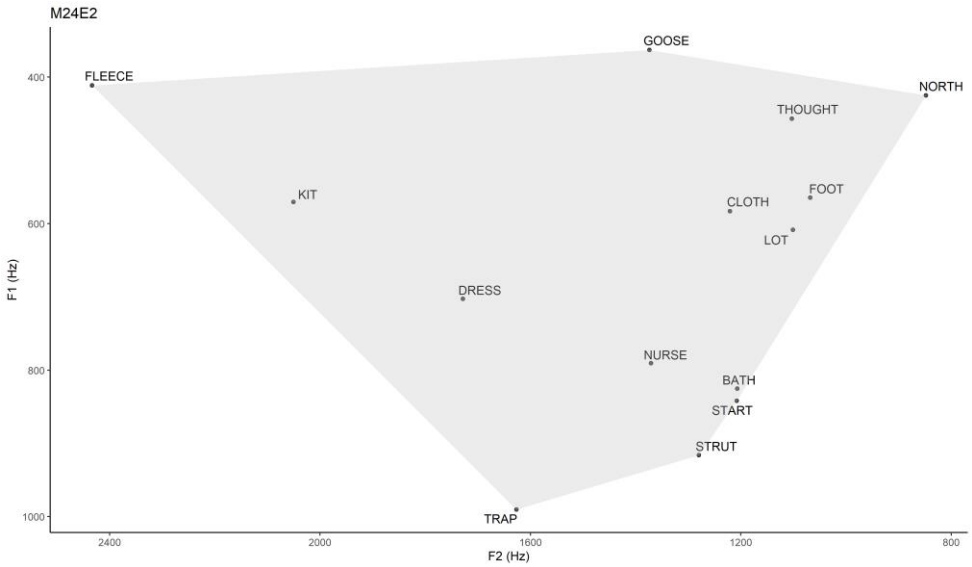



M20E2
THOUGHT
FLEECE
GOOSE
KIT
LOT
NURSE
NORTH
DRESS
STRUT
START
BATH
CLOTH
TRAP
PALM
F1 (Hz)
F2 (Hz)
500
750
1000
2500
2000
1500
1000

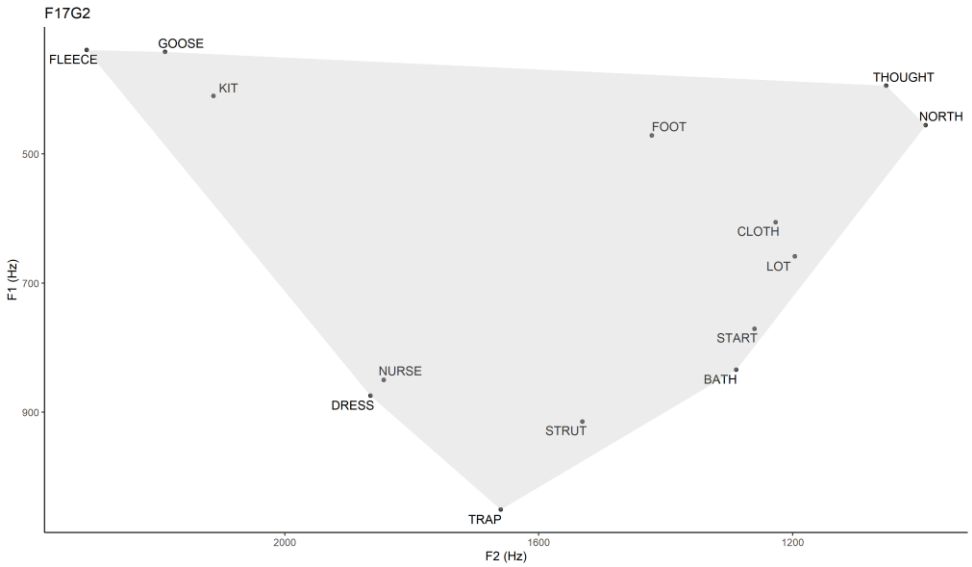

M21E2
GOOSE
KIT
FOOT
FLEECE
THOUGHT
NORTH
DRESS
CLOTH
LOT
STRUT
START
TRAP
BATH
F1 (Hz)
F2 (Hz)
400
600
800
2000
1600
1200

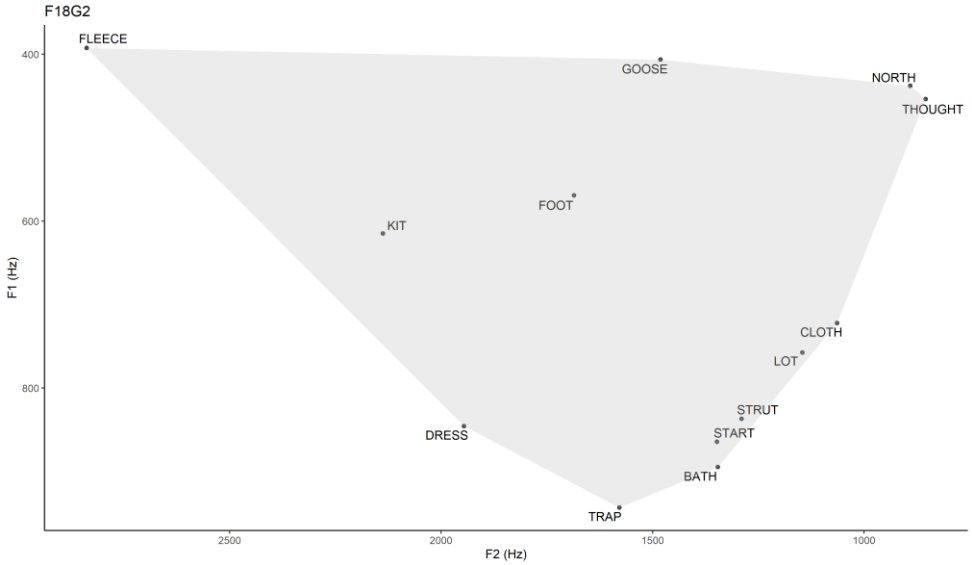

M22E2
GOOSE
FLEECE
NORTH
KIT
NURSE
THOUGHT
DRESS
LOT
CLOTH
STRUT
START
PALM
BATH
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2250
2000
1750
1500
1250

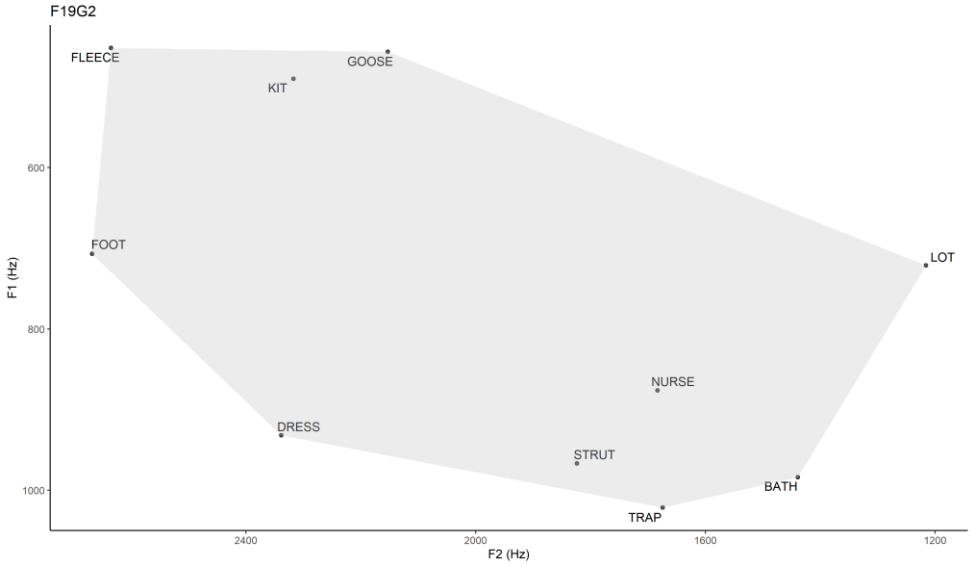

M23E2
FLEECE
GOOSE
KIT
SQUARE
FOOT
NORTH
THOUGHT
NURSE
CLOTH
DRESS
LOT
BATH
STRUT
START
TRAP
F1 (Hz)
400
600
800
2500
2000
1500
1000
F2 (Hz)

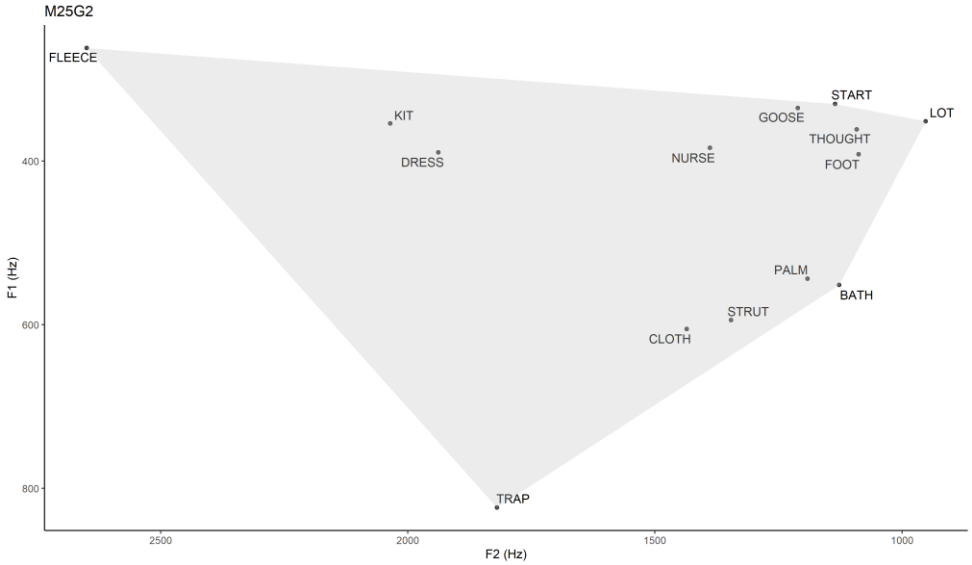

M24E2
GOOSE
FLEECE
NORTH
THOUGHT
KIT
FOOT
CLOTH
LOT
DRESS
NURSE
BATH
START
STRUT
TRAP
F1 (Hz)
400
600
800
1000
2400
2000
1600
1200
800
F2 (Hz)

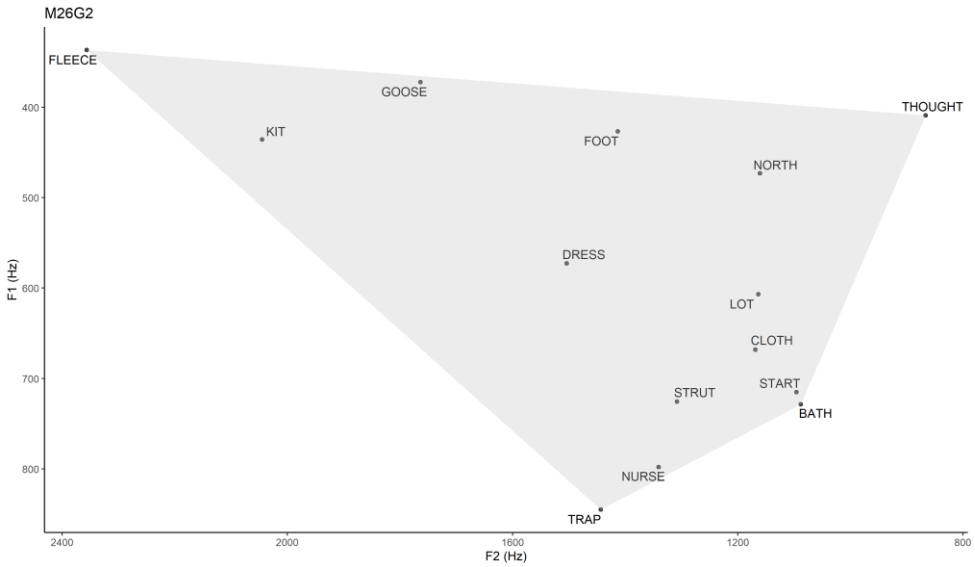
Green Line 2

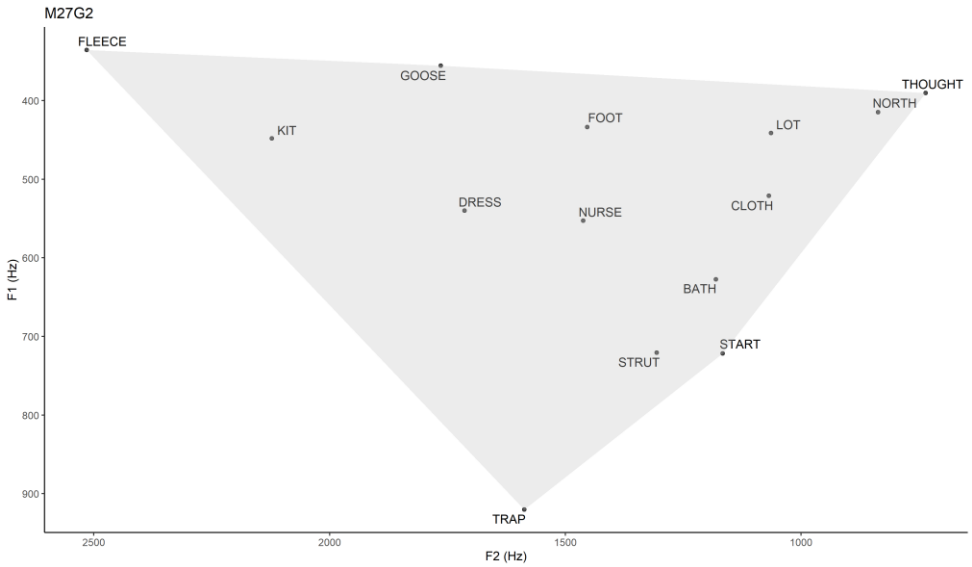

F17G2
GOOSE
FLEECE
KIT
THOUGHT
NORTH
FOOT
CLOTH
LOT
START
NURSE
BATH
DRESS
STRUT
TRAP
F1 (Hz)
500
700
900
2000
1600
1200
F2 (Hz)

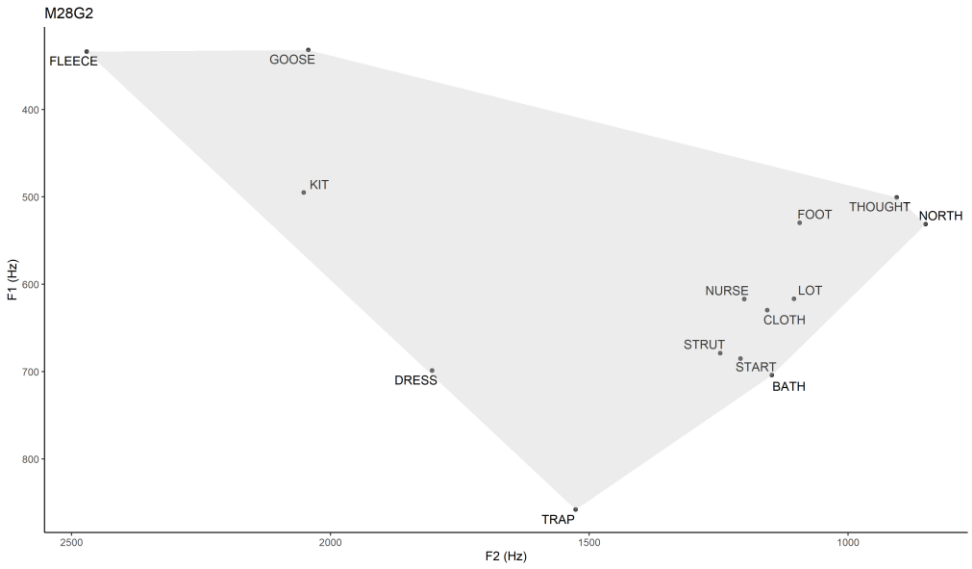

F18G2
FLEECE
GOOSE
NORTH
THOUGHT
FOOT
KIT
CLOTH
LOT
STRUT
DRESS
START
BATH
TRAP
F1 (Hz)
400
600
800
2500
2000
1500
1000
F2 (Hz)


F19G2
FLEECE
GOOSE
KIT
FOOT
LOT
NURSE
DRESS
STRUT
BATH
TRAP
F1 (Hz)
600
800
1000
2400
2000
1600
1200
F2 (Hz)

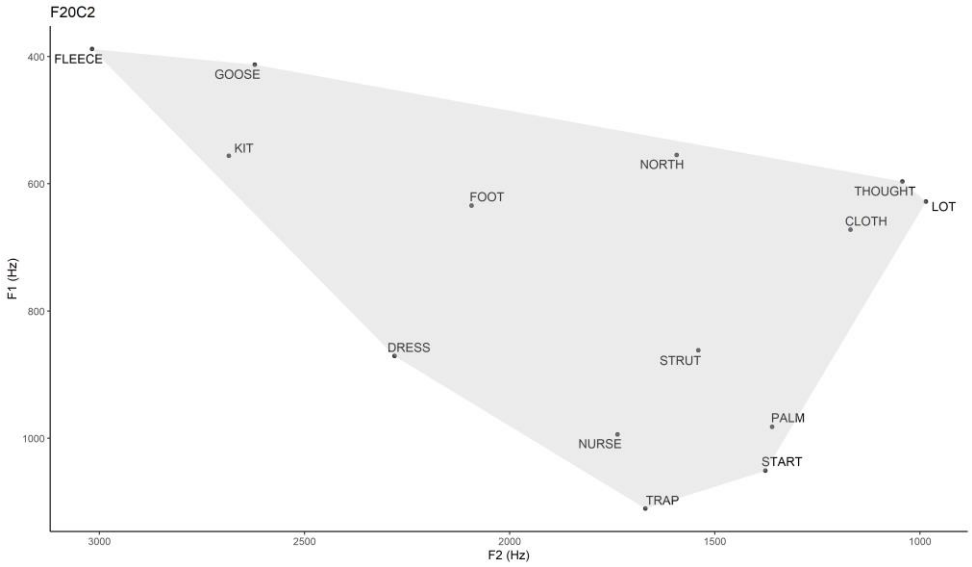

M25G2
FLEECE
KIT
DRESS
START
GOOSE
LOT
THOUGHT
NURSE
FOOT
PALM
BATH
STRUT
CLOTH
TRAP
F1 (Hz)
F2 (Hz)
400
600
800
2500
2000
1500
1000

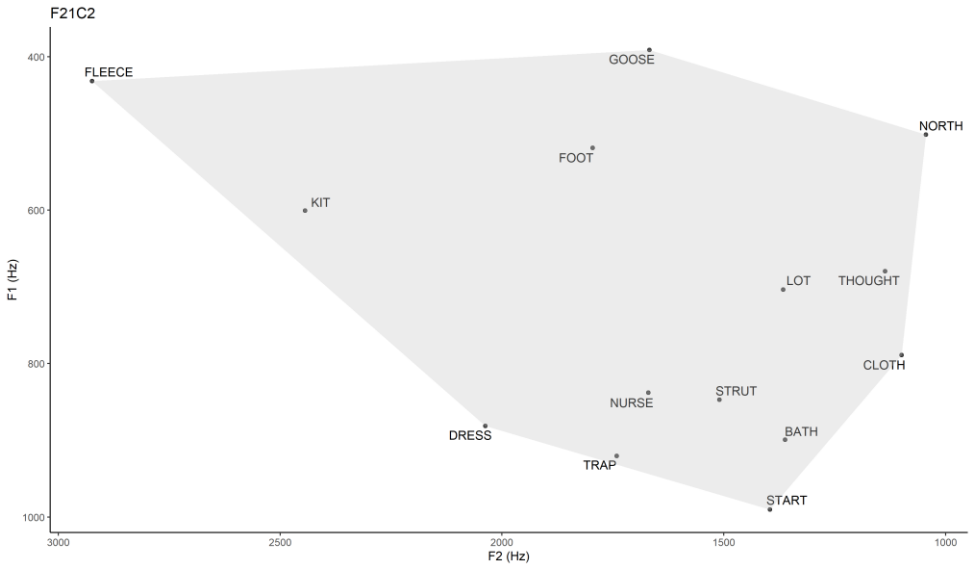

M26G2
FLEECE
GOOSE
THOUGHT
KIT
FOOT
NORTH
DRESS
LOT
CLOTH
START
STRUT
BATH
NURSE
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
2400
2000
1600
1200
800

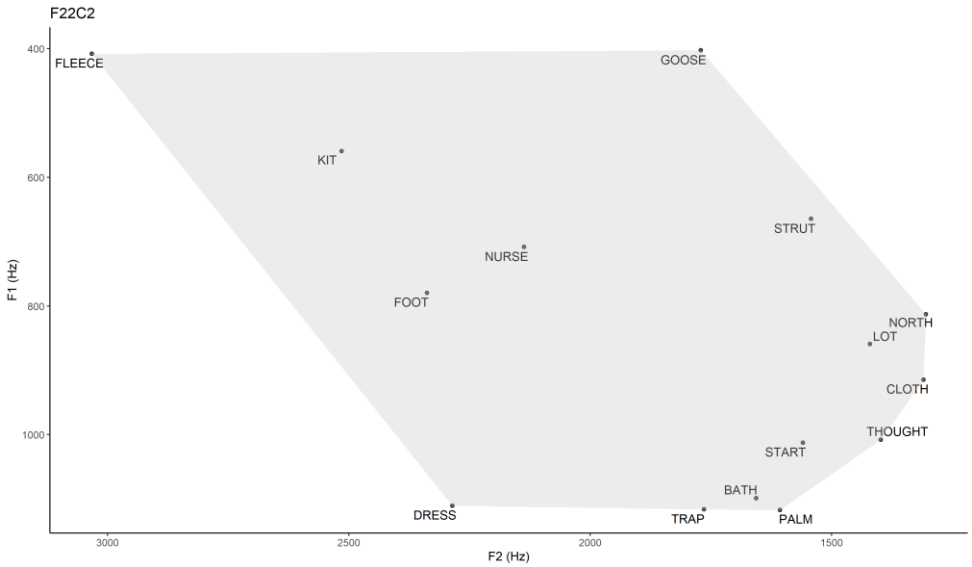

M27G2
FLEECE
GOOSE
THOUGHT
NORTH
FOOT
KIT
LOT
DRESS
NURSE
CLOTH
BATH
START
STRUT
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2500
2000
1500
1000

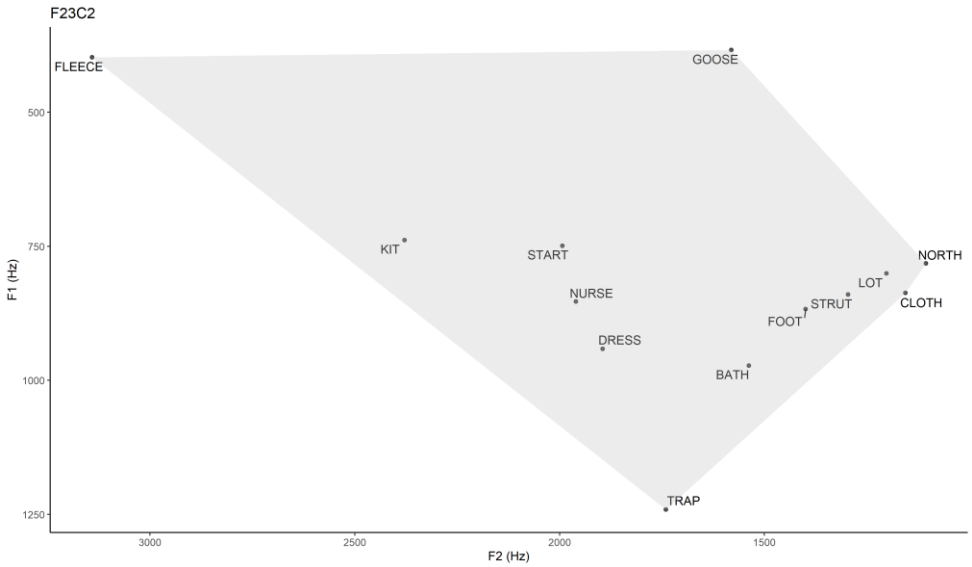


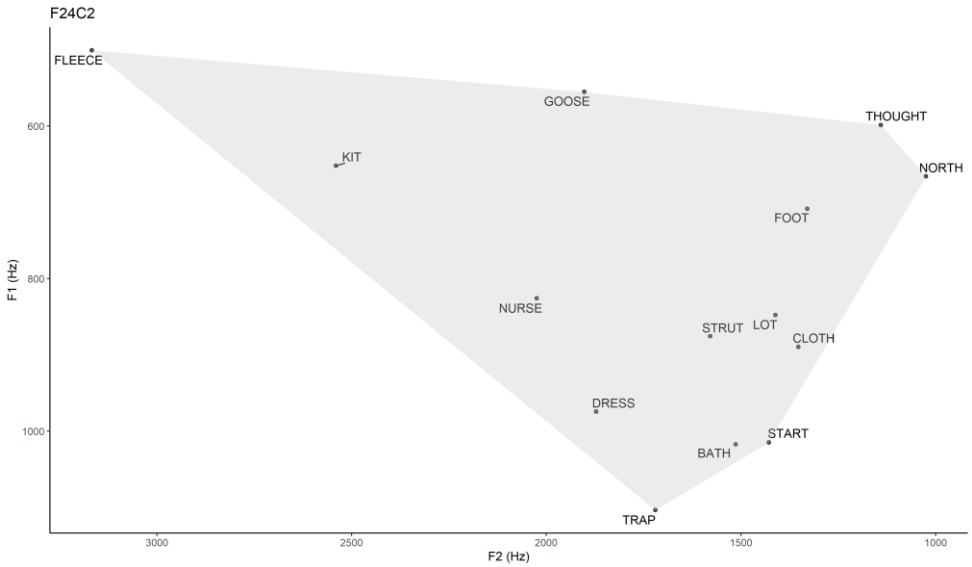
Camden Town 2

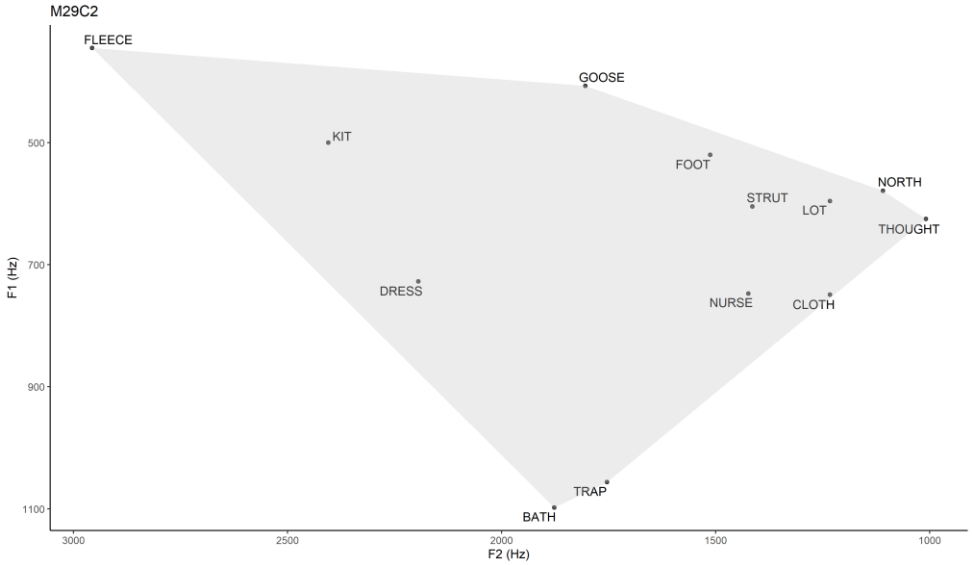


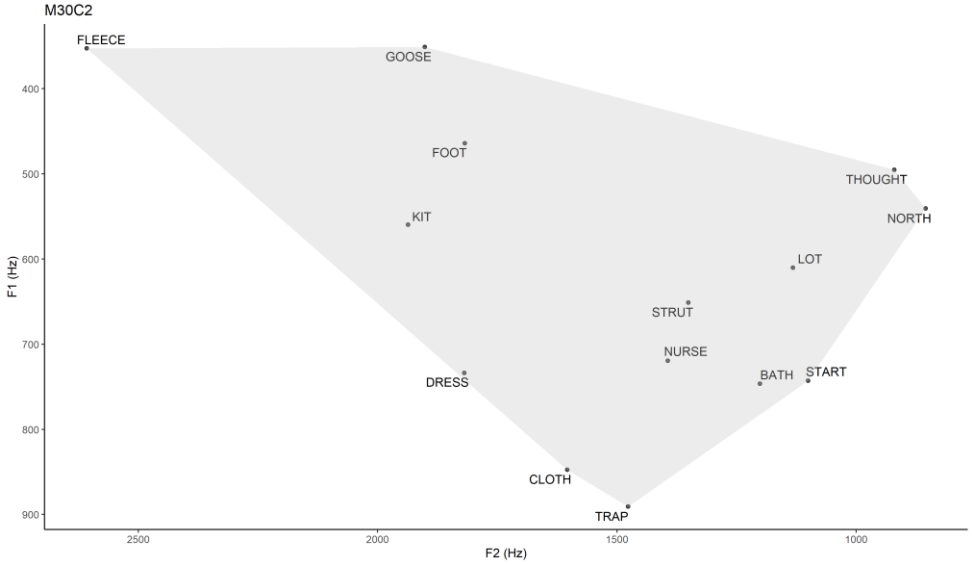



F22C2
FLEECE
GOOSE
KIT
STRUT
NURSE
FOOT
NORTH
LOT
CLOTH
THOUGHT
START
BATH
DRESS
TRAP
PALM
F1 (Hz)
F2 (Hz)
400
600
800
1000
3000
2500
2000
1500


F23C2
FLEECE
GOOSE
KIT
START
NORTH
LOT
NURSE
STRUT
CLOTH
FOOT
DRESS
BATH
TRAP
F1 (Hz)
F2 (Hz)
500
750
1000
1250
3000
2500
2000
1500

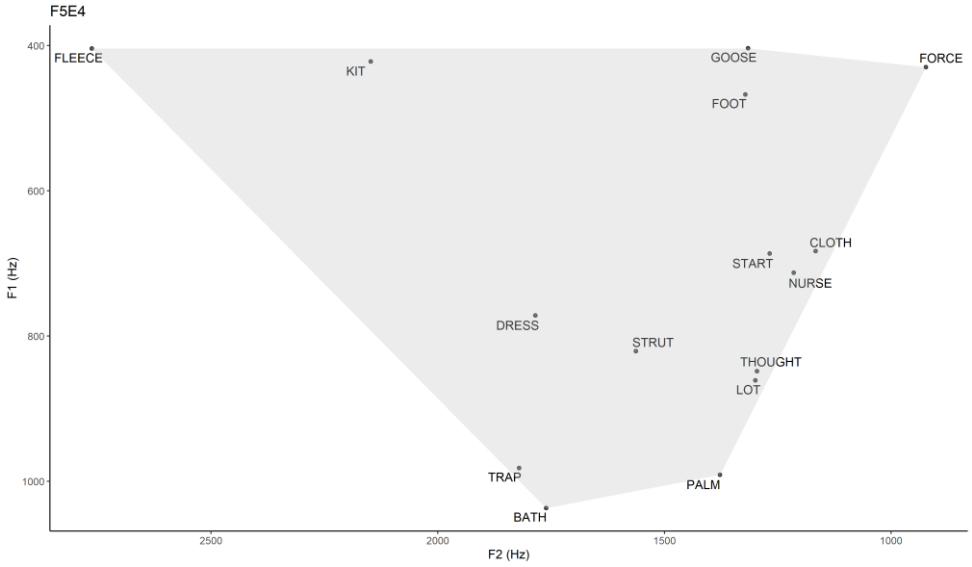

F24C2
FLEECE
GOOSE
THOUGHT
KIT
NORTH
FOOT
NURSE
STRUT
LOT
CLOTH
DRESS
START
BATH
TRAP
F1 (Hz)
F2 (Hz)
600
800
1000
3000
2500
2000
1500
1000

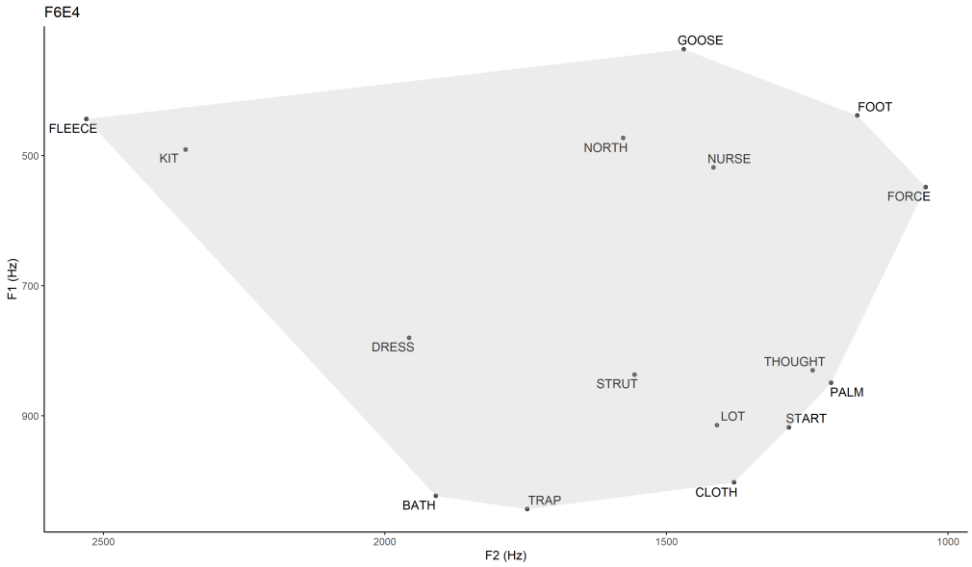

M29C2
FLEECE
GOOSE
KIT
FOOT
NORTH
STRUT
LOT
THOUGHT
DRESS
NURSE
CLOTH
TRAP
BATH
F1 (Hz)
F2 (Hz)
500
700
900
1100
3000
2500
2000
1500
1000

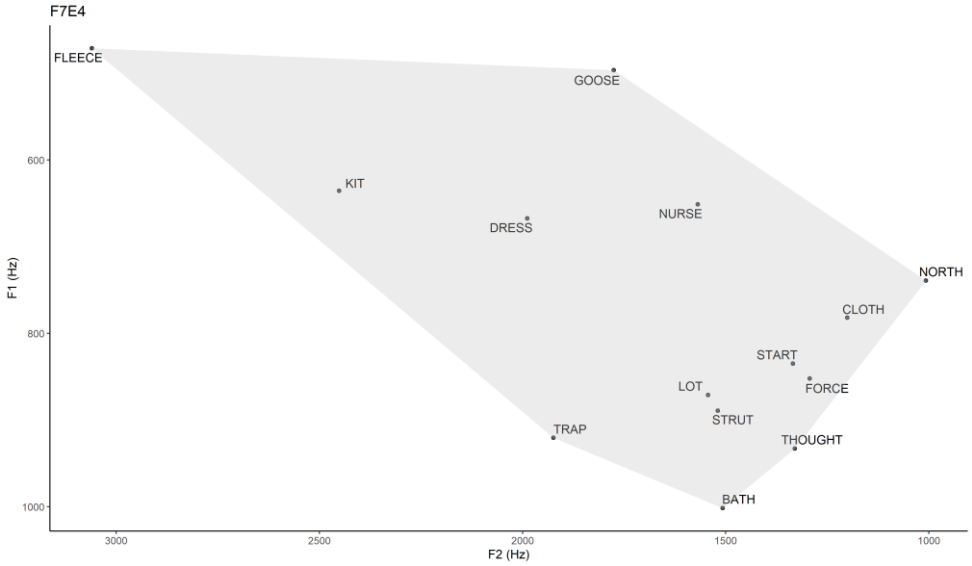

M30C2
FLEECE
GOOSE
FOOT
THOUGHT
KIT
NORTH
LOT
STRUT
NURSE
DRESS
BATH
START
CLOTH
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2500
2000
1500
1000

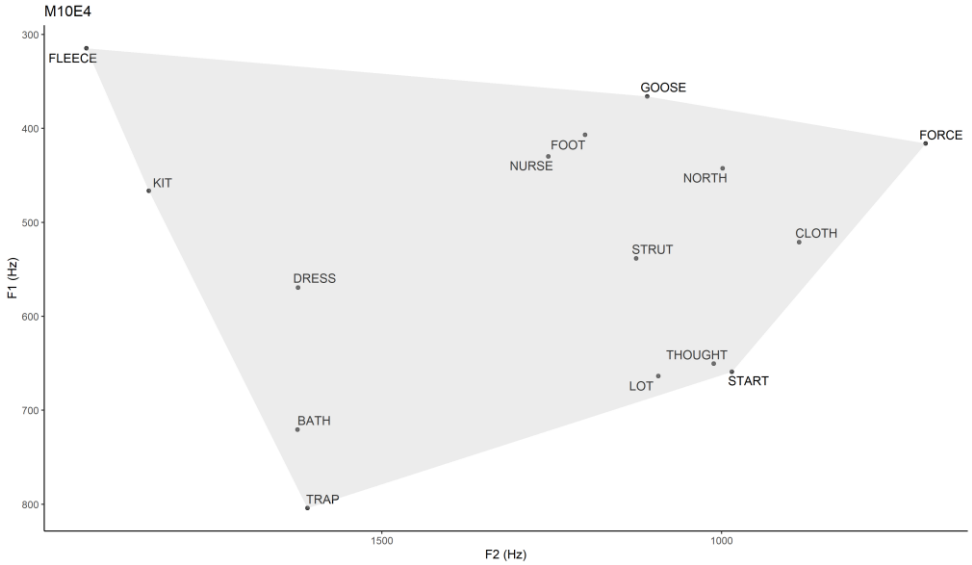
General American sub-corpus

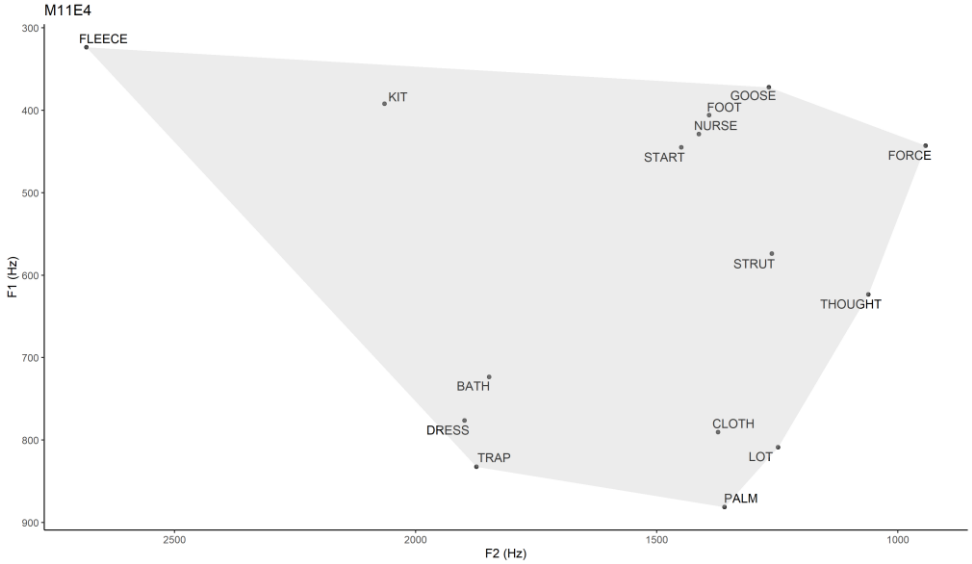
English G Access 4

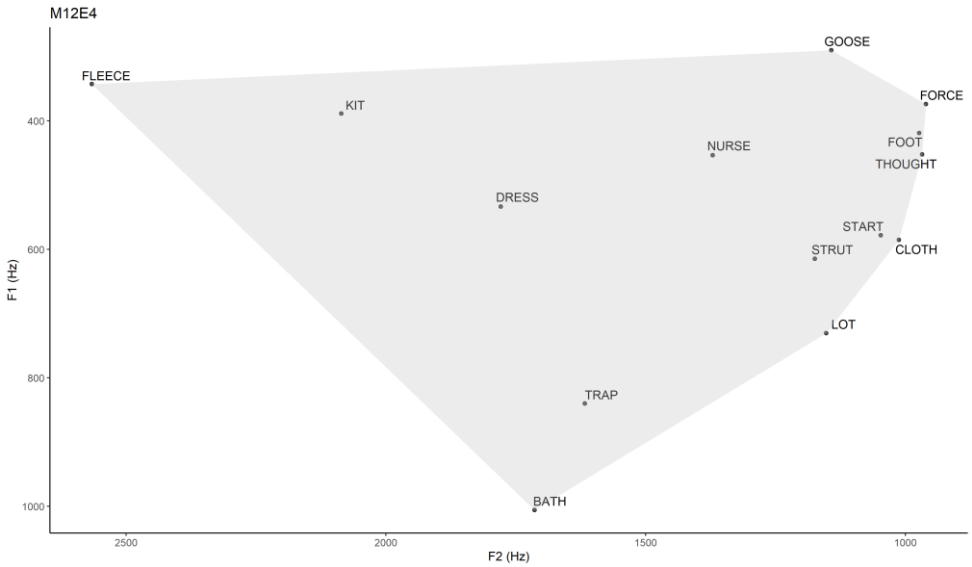


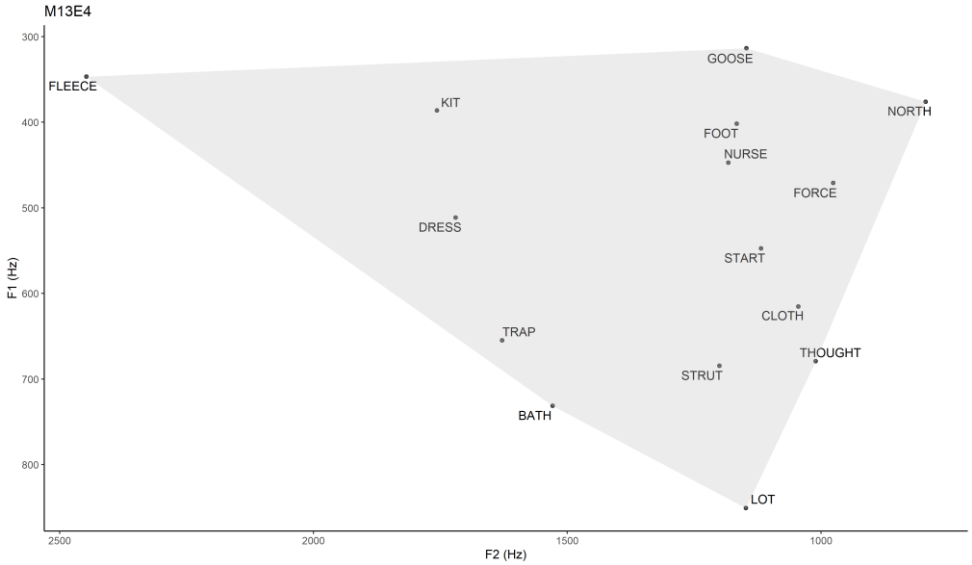




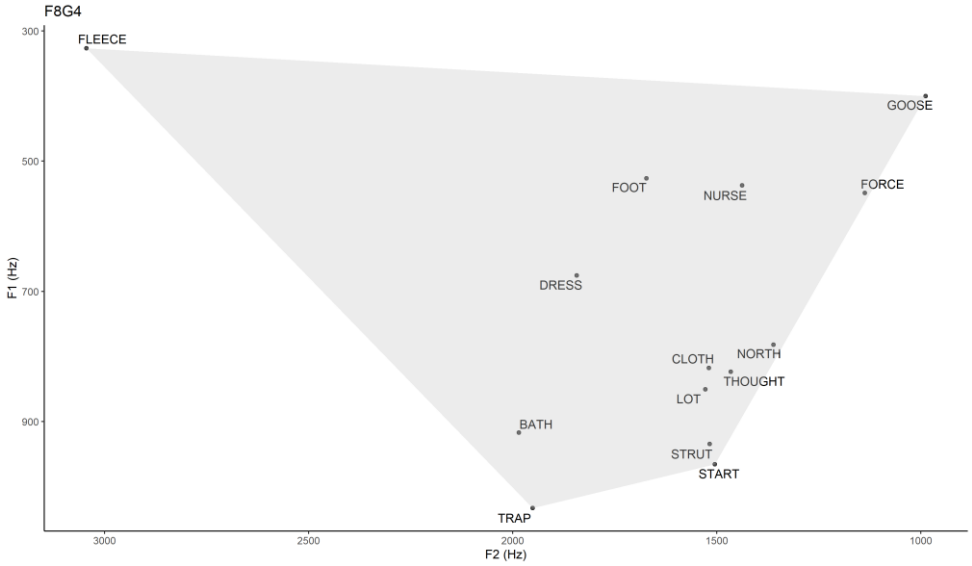

M10E4
FLEECE
GOOSE
FORCE
FOOT
NURSE
NORTH
KIT
CLOTH
STRUT
DRESS
THOUGHT
START
LOT
BATH
TRAP
F1 (Hz)
F2 (Hz)

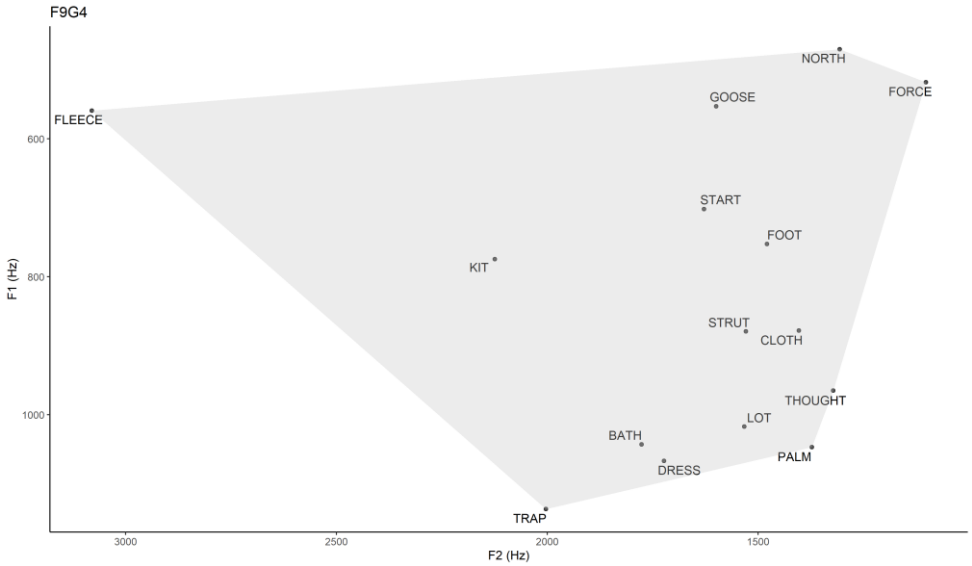

M11E4
FLEECE
KIT
GOOSE
FOOT
NURSE
START
FORCE
STRUT
THOUGHT
BATH
DRESS
CLOTH
TRAP
LOT
PALM
F1 (Hz)
F2 (Hz)

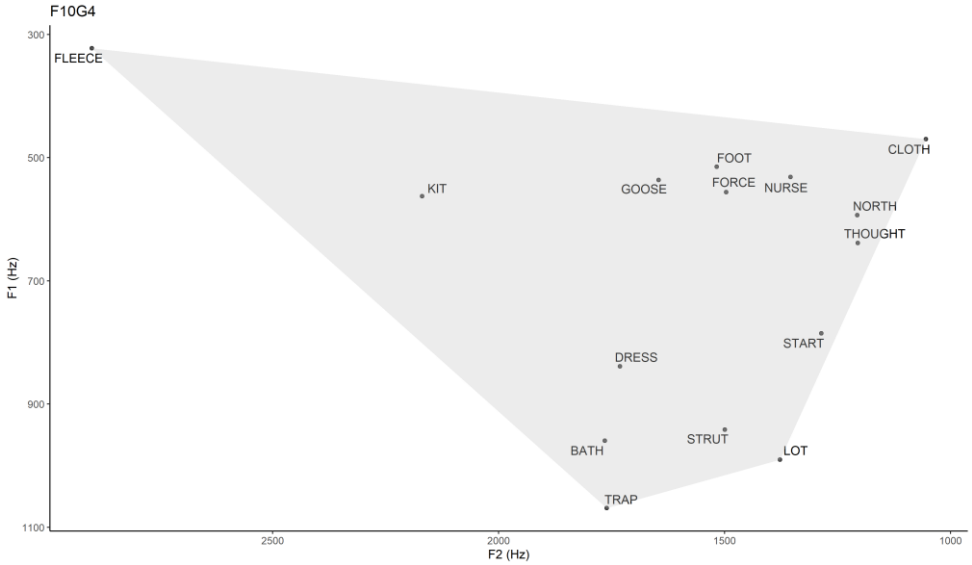

M12E4
GOOSE
FLEECE
FORCE
KIT
NURSE
FOOT
THOUGHT
DRESS
START
STRUT
CLOTH
LOT
TRAP
BATH
F1 (Hz)
F2 (Hz)

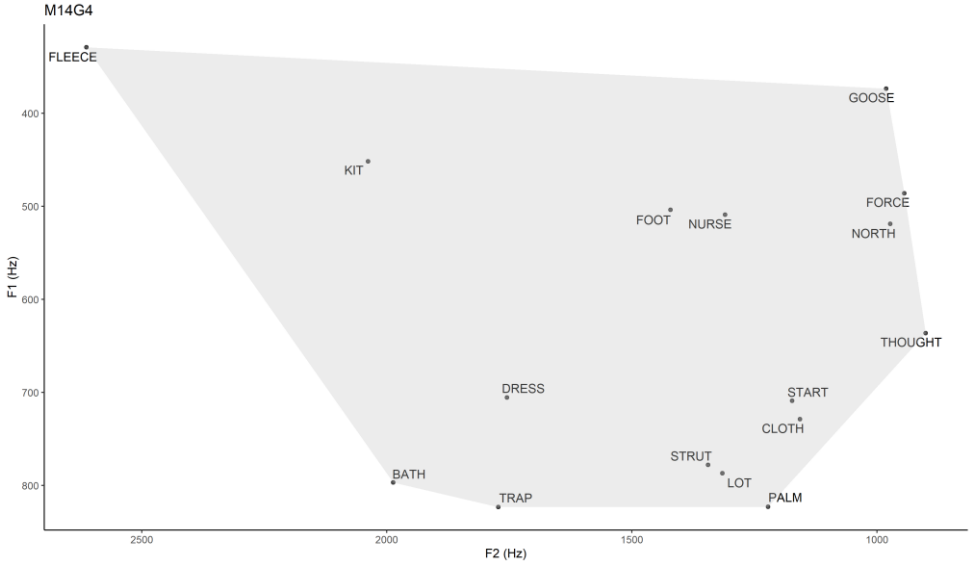

M13E4
GOOSE
FLEECE
KIT
NORTH
FOOT
NURSE
FORCE
DRESS
START
CLOTH
TRAP
THOUGHT
STRUT
BATH
LOT
F1 (Hz)
F2 (Hz)

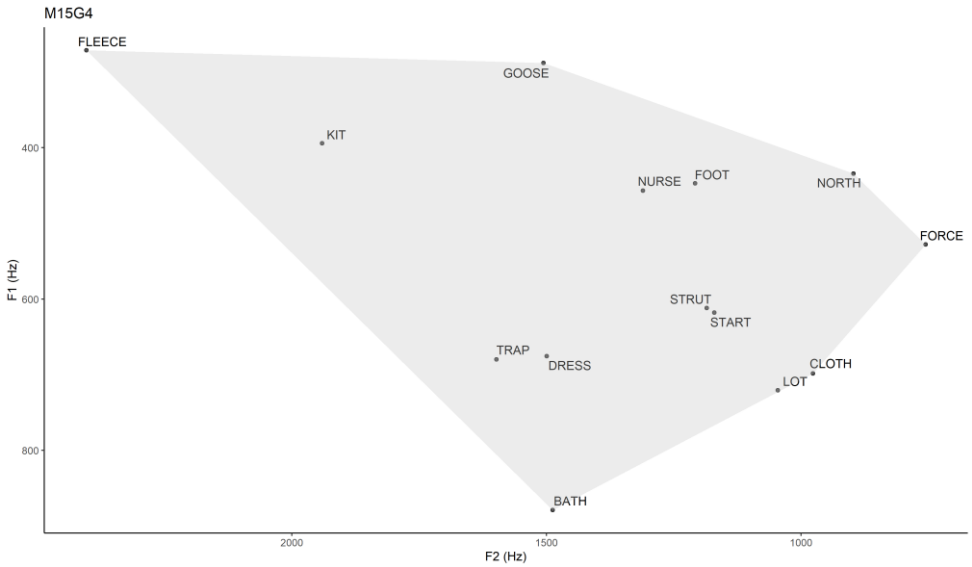
Green Line 4

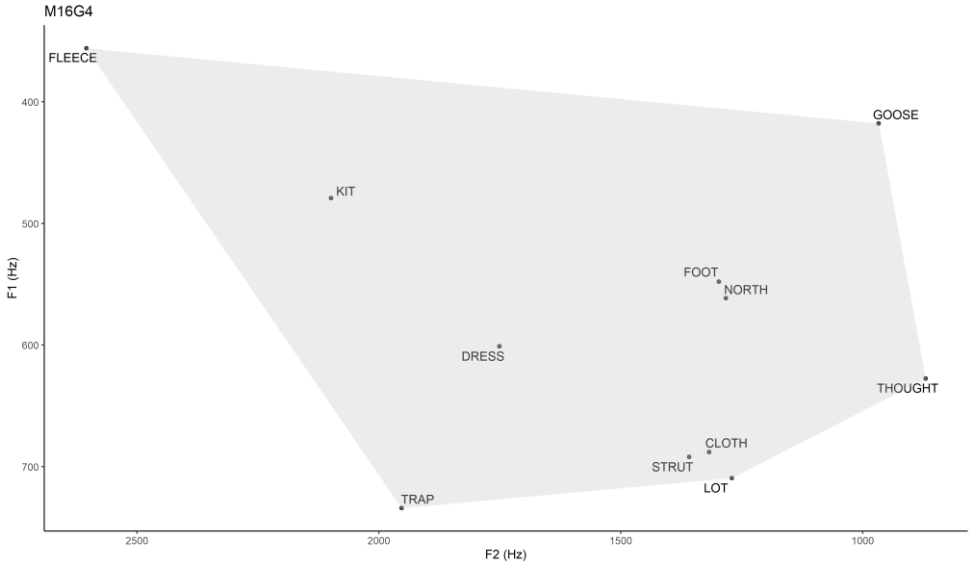

F8G4
FLEECE
GOOSE
FOOT
NURSE
FORCE
DRESS
CLOTH
NORTH
THOUGHT
LOT
BATH
STRUT
START
TRAP
F1 (Hz)
F2 (Hz)

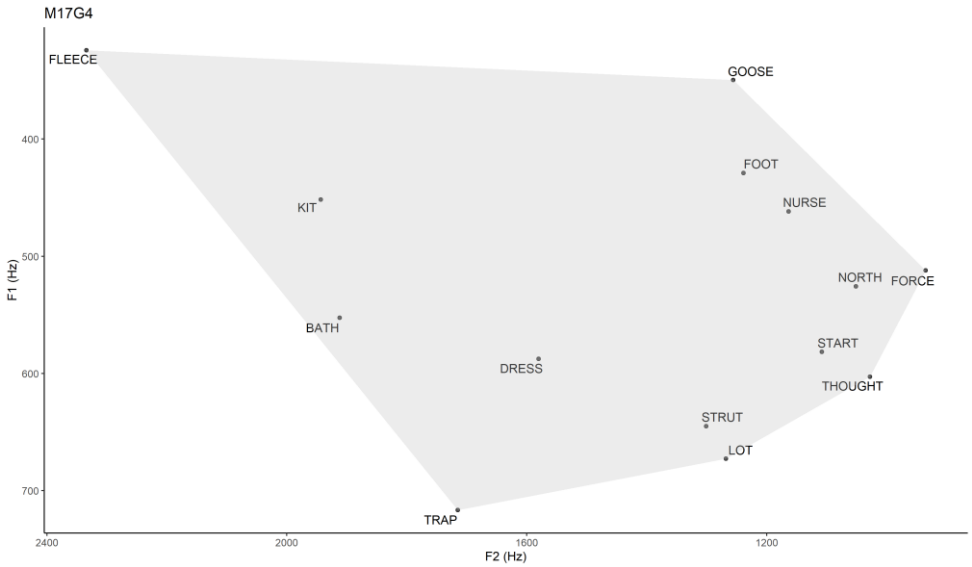

F9G4
NORTH
FORCE
GOOSE
FLEECE
START
FOOT
KIT
STRUT
CLOTH
THOUGHT
LOT
BATH
PALM
DRESS
TRAP
F1 (Hz)
F2 (Hz)


F10G4
FLEECE
CLOTH
FOOT
KIT
GOOSE
FORCE
NURSE
NORTH
THOUGHT
START
DRESS
STRUT
BATH
LOT
TRAP
F1 (Hz)
F2 (Hz)
300
500
700
900
1100
2500
2000
1500
1000

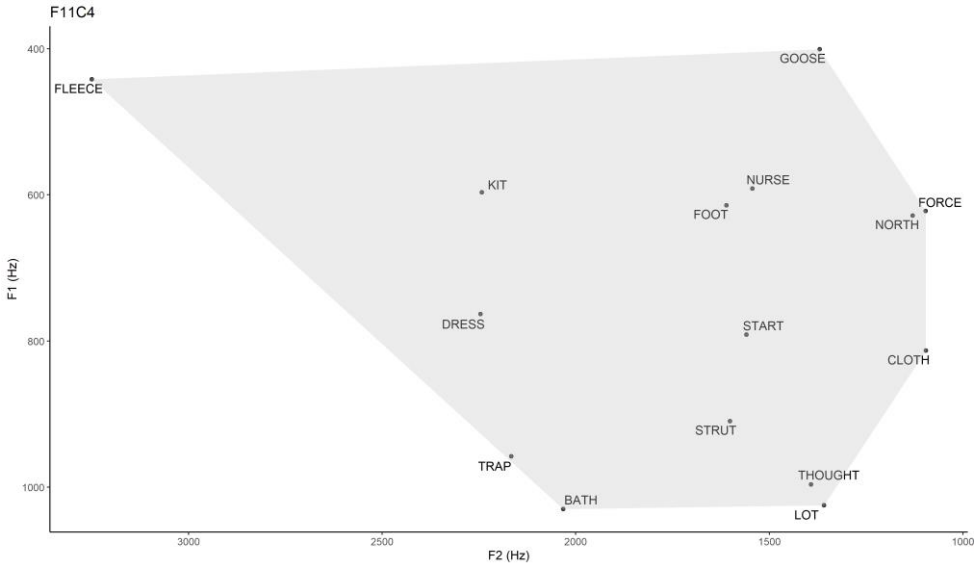

M14G4
FLEECE
GOOSE
KIT
FORCE
FOOT
NURSE
NORTH
THOUGHT
DRESS
START
CLOTH
STRUT
BATH
LOT
TRAP
PALM
F1 (Hz)
F2 (Hz)
400
500
600
700
800
2500
2000
1500
1000

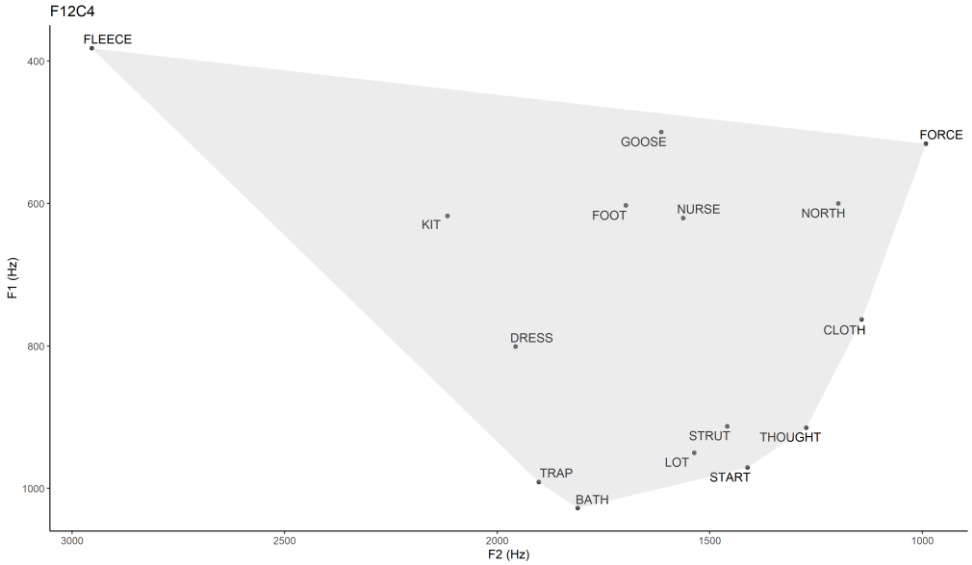

M15G4
FLEECE
GOOSE
KIT
NURSE
FOOT
NORTH
FORCE
STRUT
START
TRAP
DRESS
CLOTH
LOT
BATH
F1 (Hz)
F2 (Hz)
400
600
800
2000
1500
1000

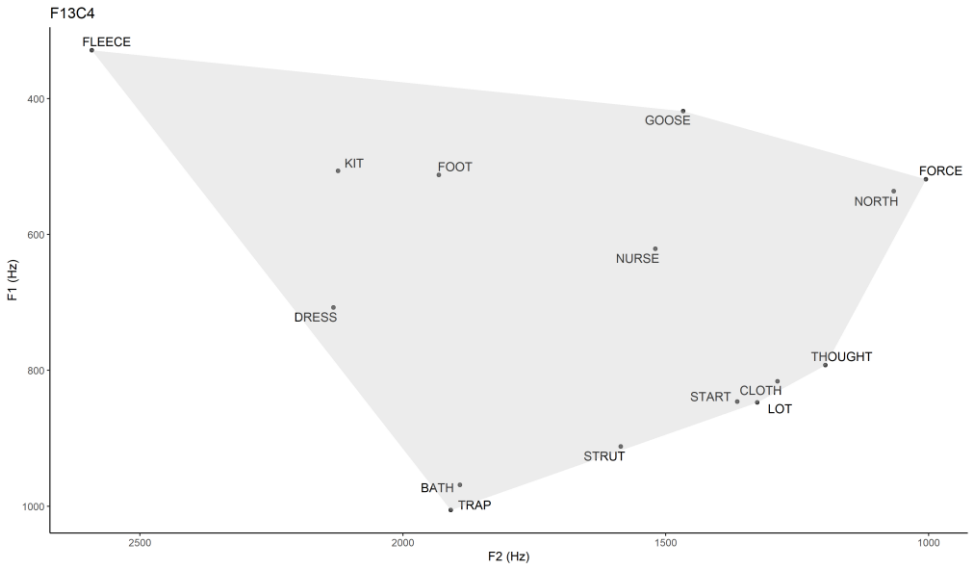


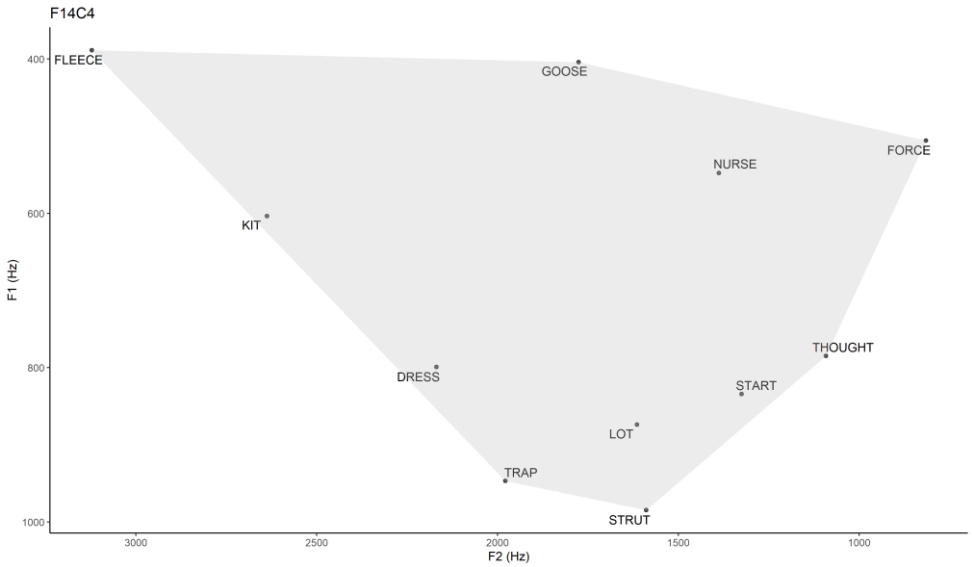


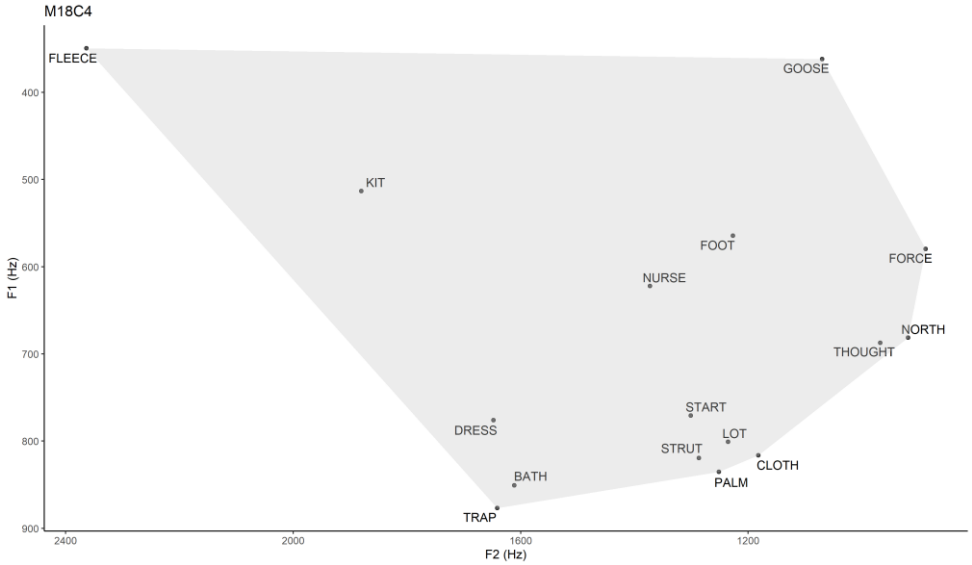
Camden Town 4

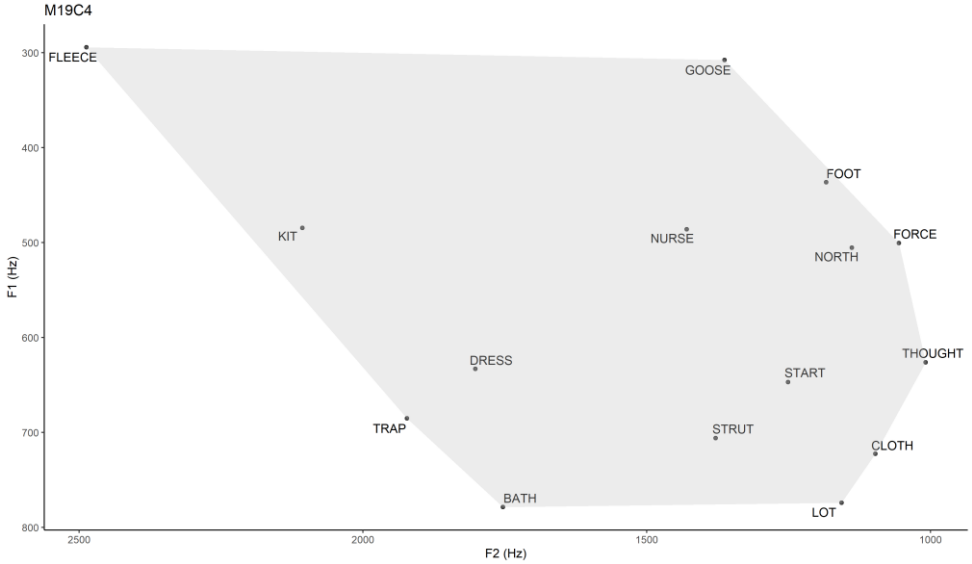



F12C4
FLEECE
GOOSE
FORCE
KIT
FOOT
NURSE
NORTH
DRESS
CLOTH
STRUT
THOUGHT
LOT
TRAP
START
BATH
F1 (Hz)
F2 (Hz)
400
600
800
1000
3000
2500
2000
1500
1000


F13C4
FLEECE
GOOSE
KIT
FOOT
FORCE
NORTH
NURSE
DRESS
THOUGHT
START
CLOTH
LOT
STRUT
BATH
TRAP
F1 (Hz)
F2 (Hz)
400
600
800
1000
2500
2000
1500
1000


F14C4
FLEECE
GOOSE
FORCE
NURSE
KIT
THOUGHT
DRESS
START
LOT
TRAP
STRUT
F1 (Hz)
F2 (Hz)
400
600
800
1000
3000
2500
2000
1500
1000

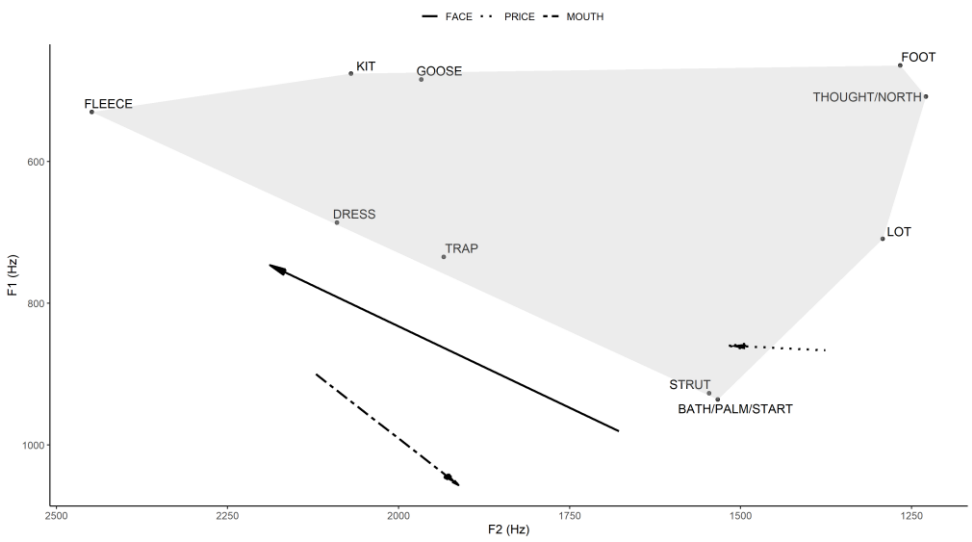

M18C4
FLEECE
GOOSE
KIT
FOOT
FORCE
NURSE
NORTH
THOUGHT
START
DRESS
LOT
STRUT
CLOTH
BATH
PALM
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2400
2000
1600
1200


M19C4
FLEECE
GOOSE
FOOT
KIT
NURSE
FORCE
NORTH
THOUGHT
DRESS
START
TRAP
STRUT
CLOTH
BATH
LOT
F1 (Hz)
F2 (Hz)
300
400
500
600
700
800
2500
2000
1500
1000

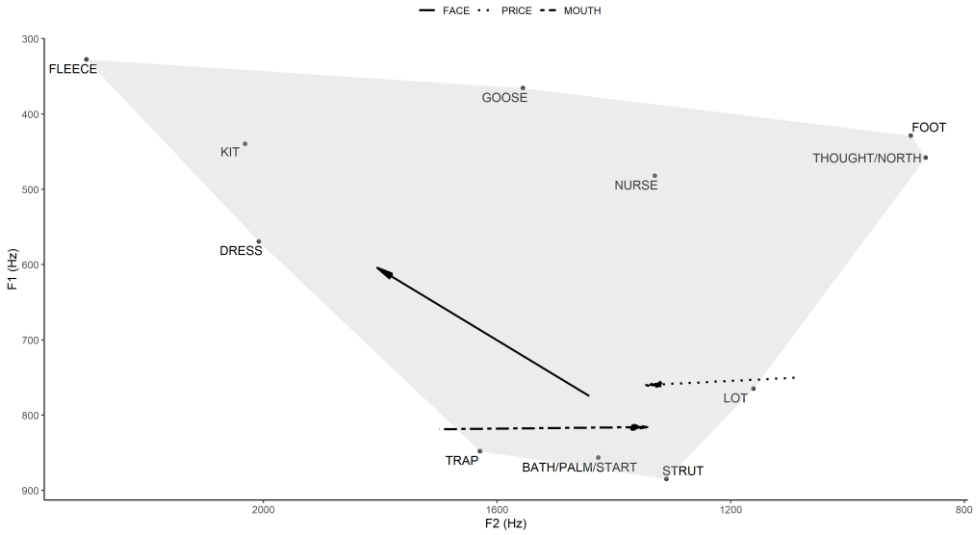
Australian English sub-corpus

English G Access 5

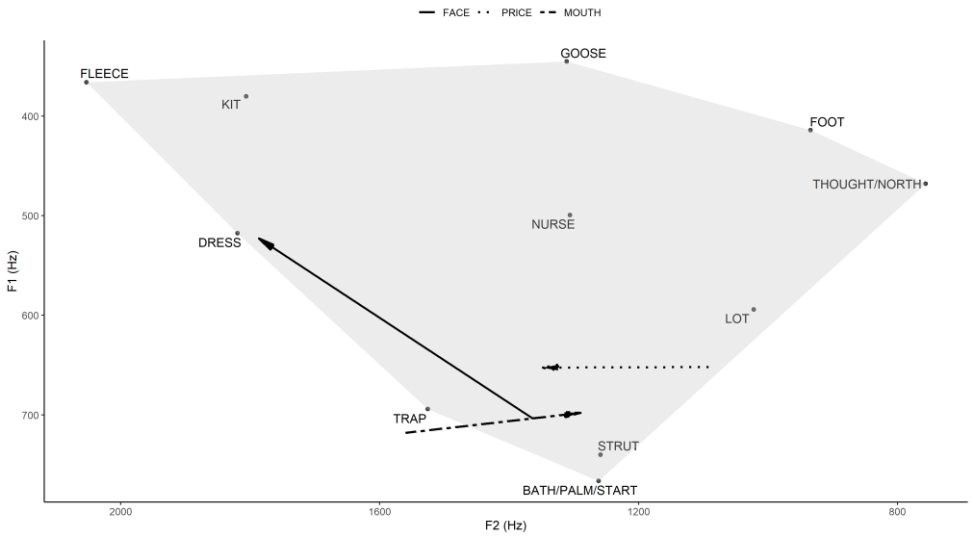
F1E5



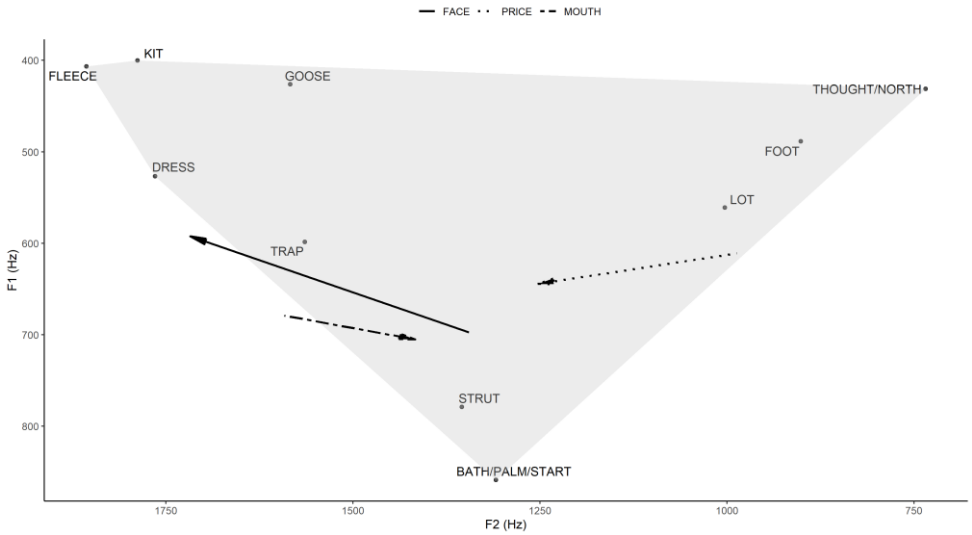
M1E5



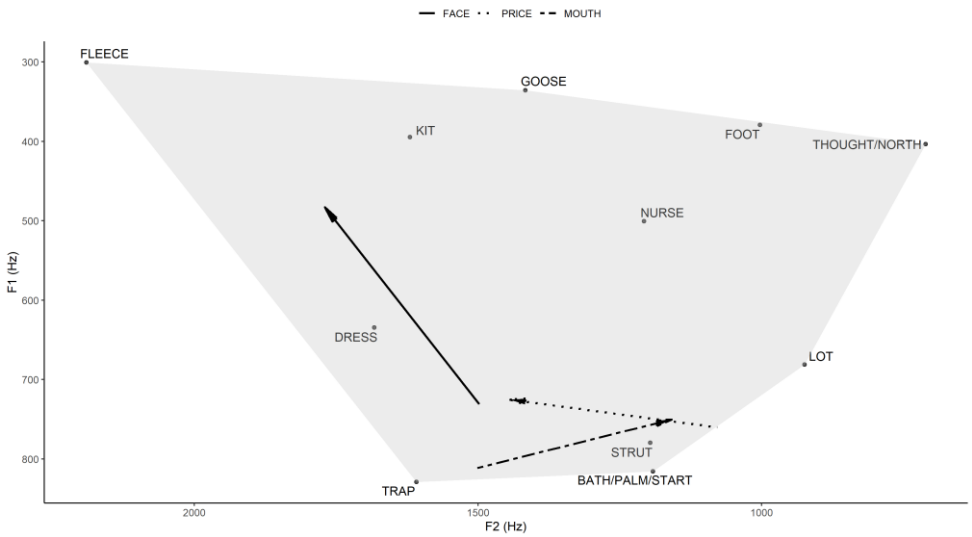
M2E5



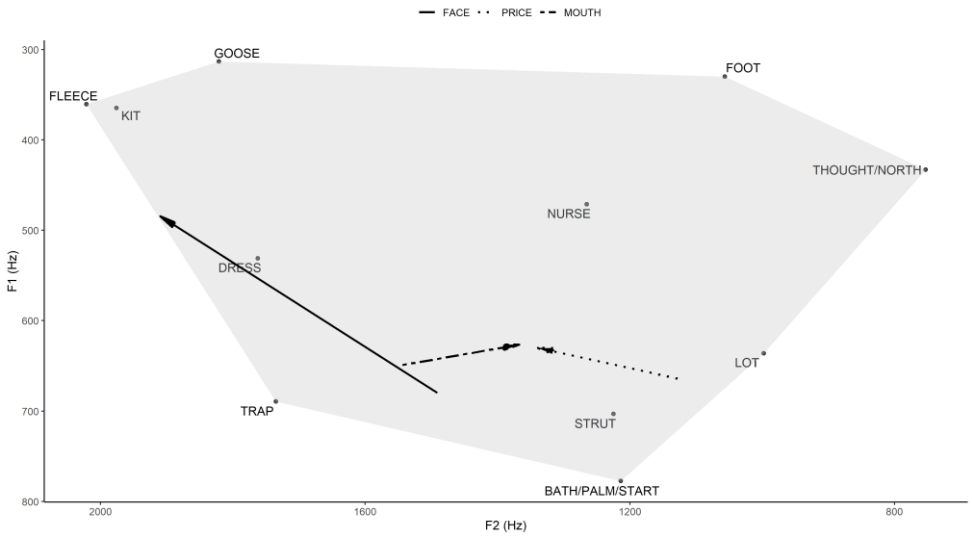
M3E5



M4E5

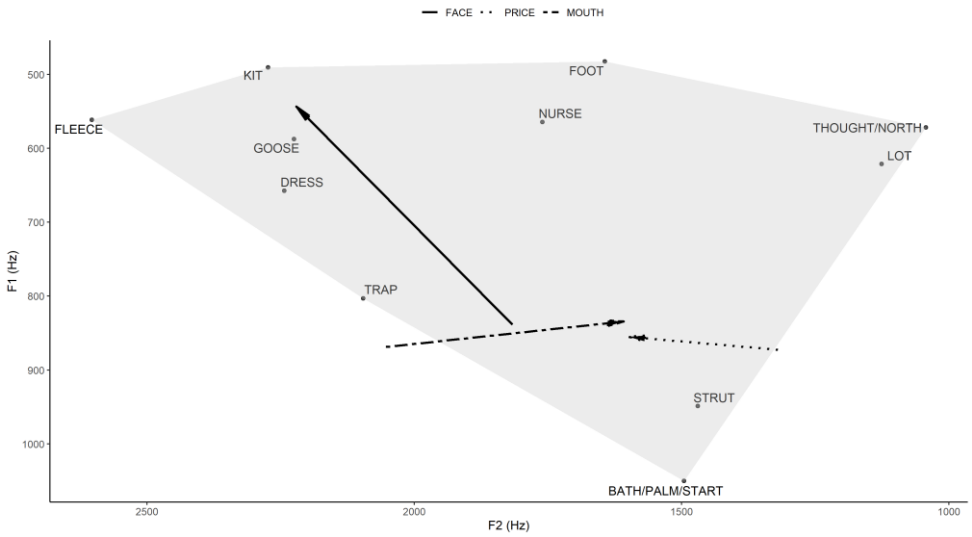


M5E5

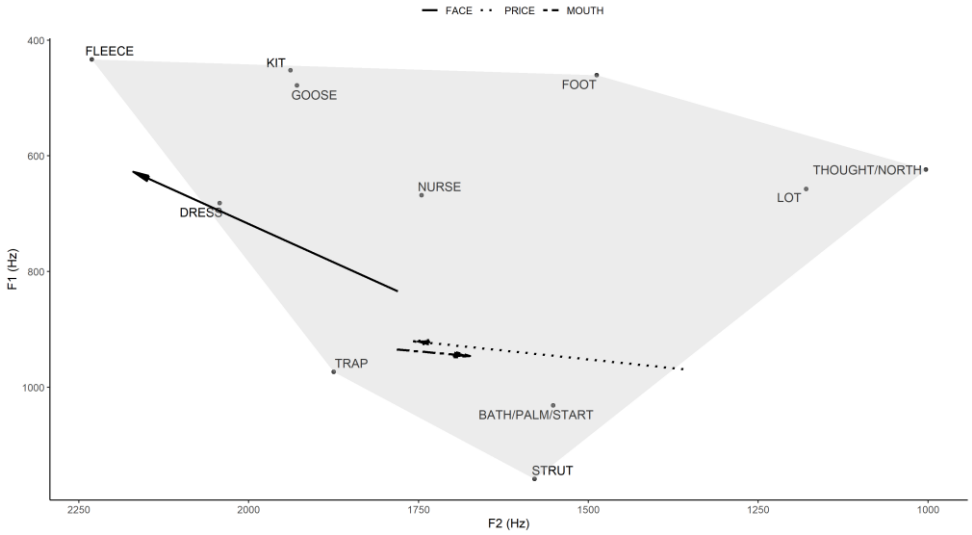


Green Line 5

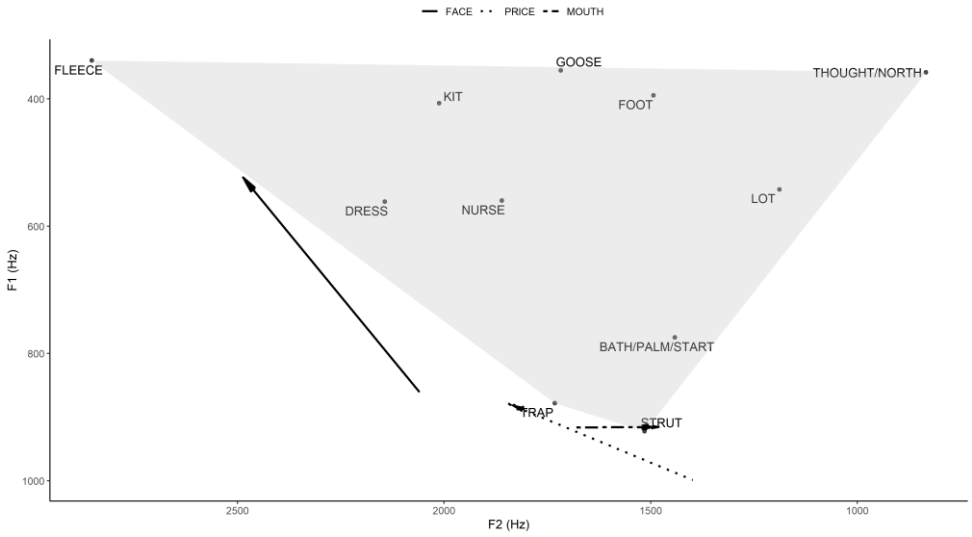
F2G5



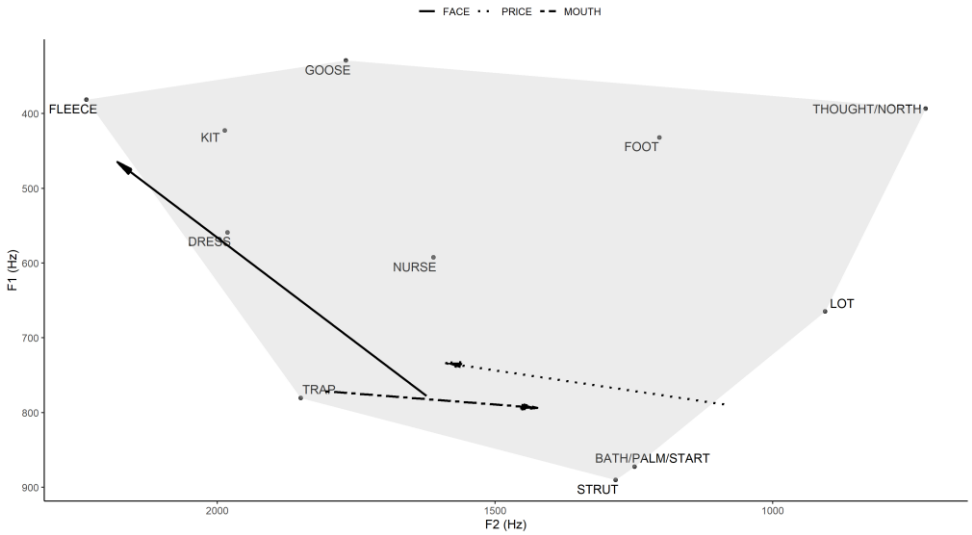
F3G5



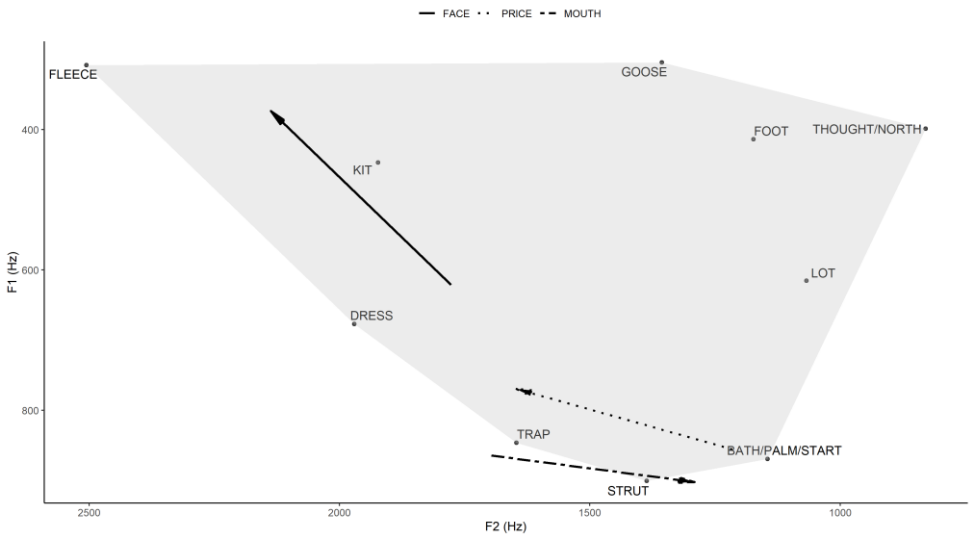
F4G5



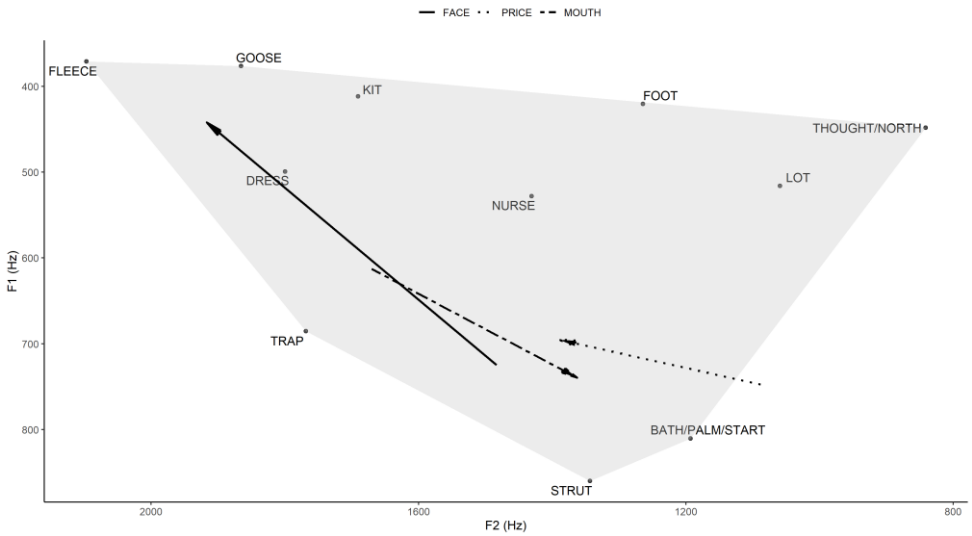
M6G5

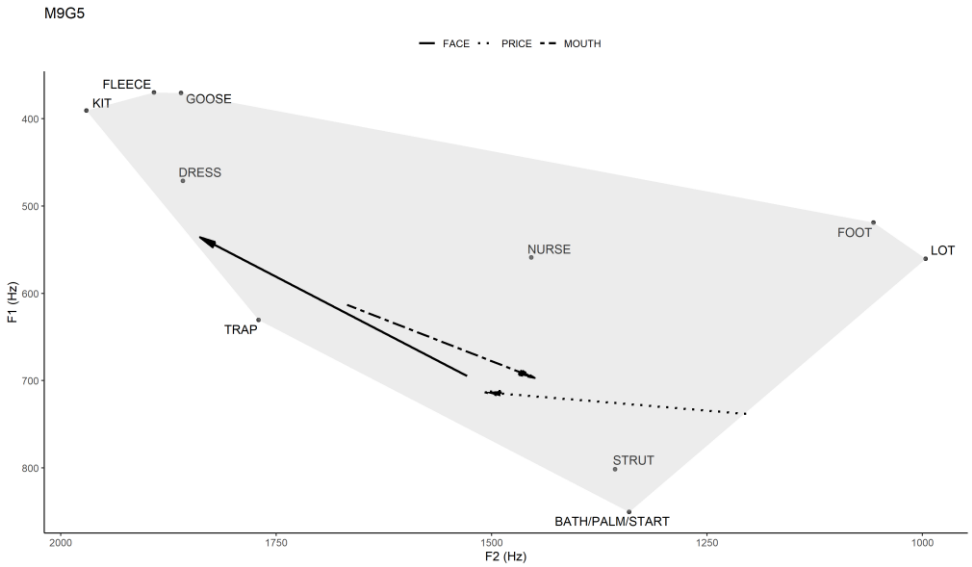


M7G5



M8G5



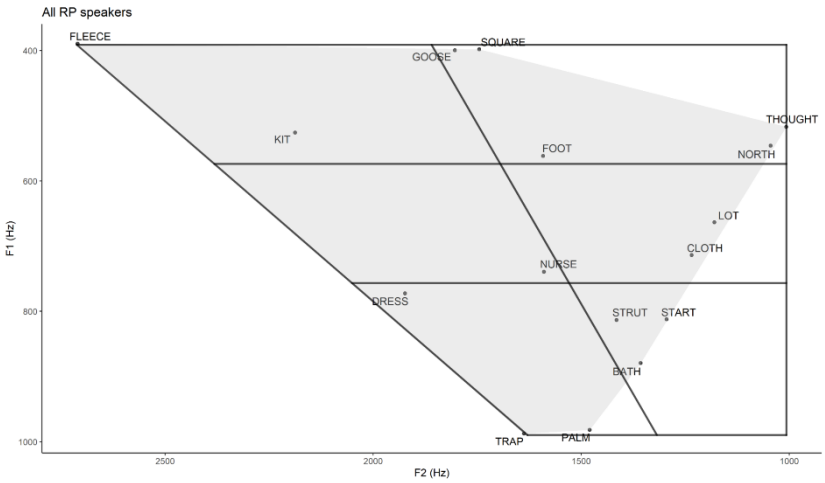

M9G5
FACE
PRICE
MOUTH
FLEECE
KIT
GOOSE
DRESS
NURSE
FOOT
LOT
TRAP
STRUT
BATH/PALM/START
F1 (Hz)
F2 (Hz)
400
500
600
700
800
2000
1750
1500
1250
1000

Appendix 2

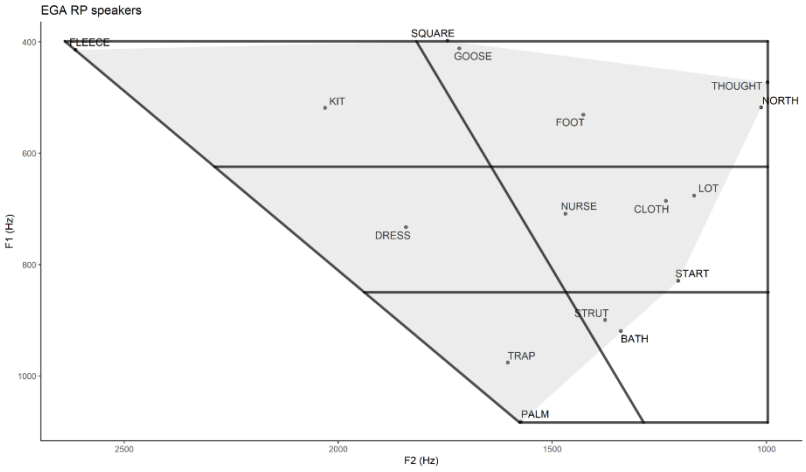
Vowel plots with articulatory categorisation

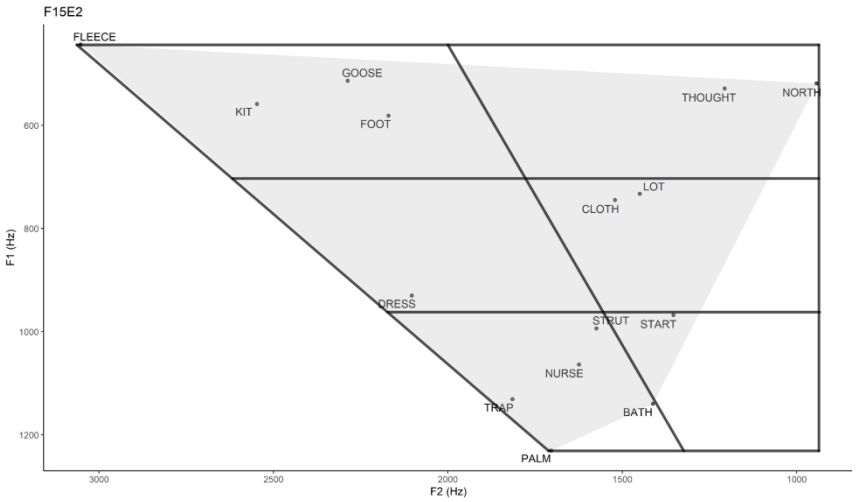
This appendix includes the grouped and speaker-individual vowel plots with the articulatory categorisation The method of categorisation has been presented in Section 3.3.1. In each vowel plot, the speaker or speaker group is indicated in the top left corner of each vowel plot.

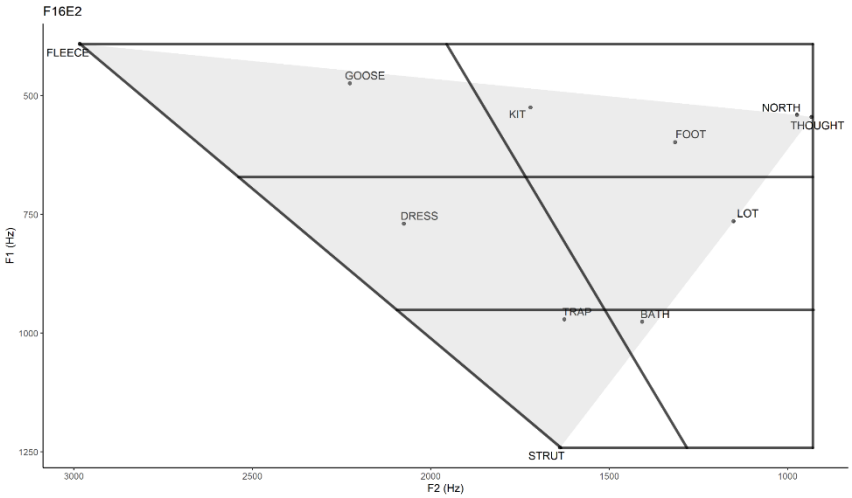
Received Pronunciation sub-corpus

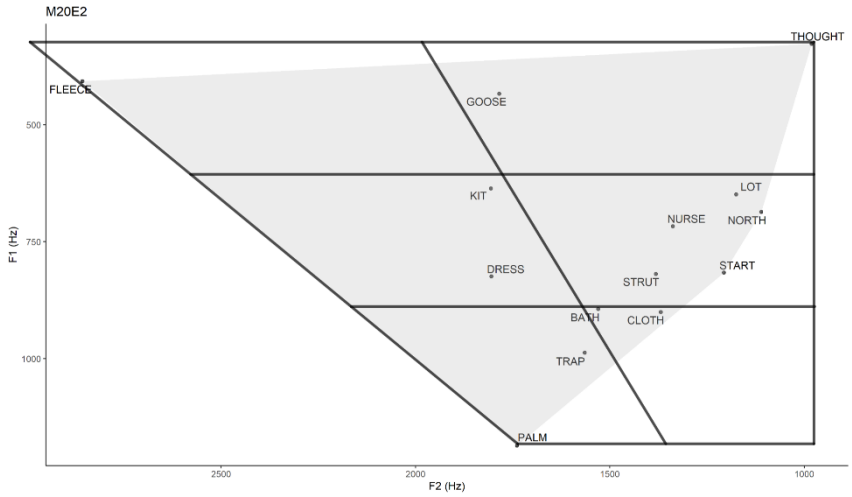


English G Access 2

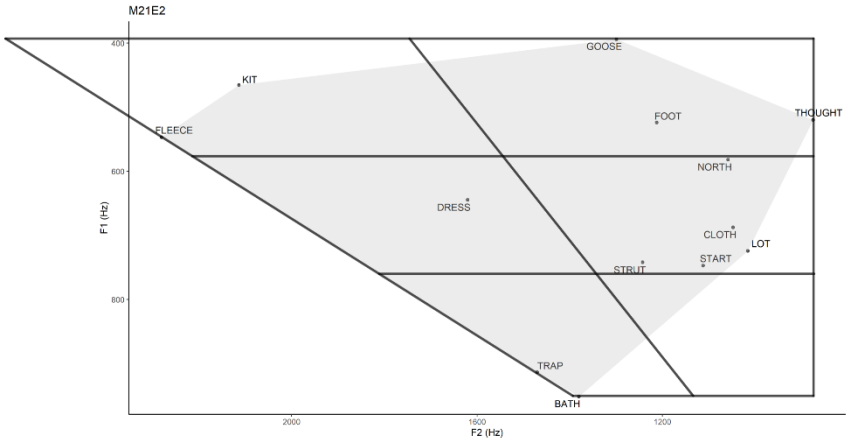


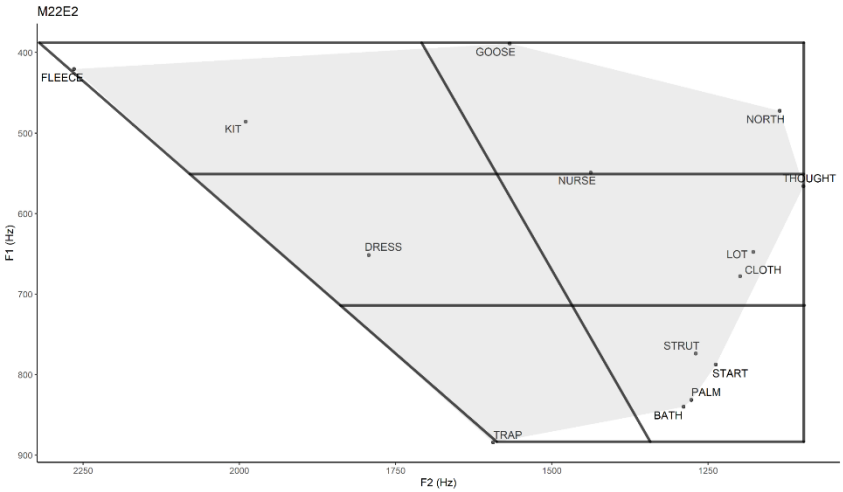

F15E2
FLEECE
GOOSE
KIT
FOOT
THOUGHT
NORTH
LOT
CLOTH
DRESS
STRUT
START
NURSE
TRAP
BATH
PALM
F1 (Hz)
F2 (Hz)
600
800
1000
1200
3000
2500
2000
1500
1000

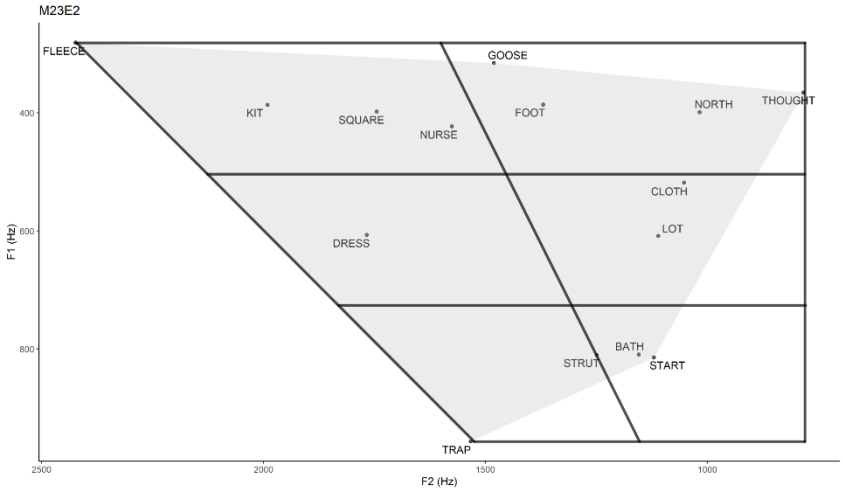

F16E2
FLEECE
GOOSE
KIT
NORTH
THOUGHT
FOOT
DRESS
LOT
TRAP
BATH
STRUT
F1 (Hz)
F2 (Hz)
500
750
1000
1250
3000
2500
2000
1500
1000

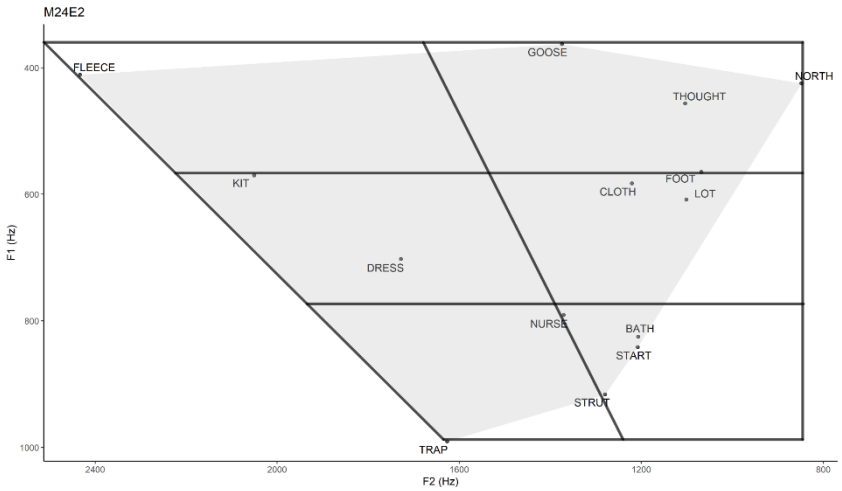

M20E2
THOUGHT
FLEECE
GOOSE
LOT
KIT
NURSE
NORTH
START
DRESS
STRUT
BATH
CLOTH
TRAP
PALM
F1 (Hz)
F2 (Hz)
500
750
1000
2500
2000
1500
1000

,5

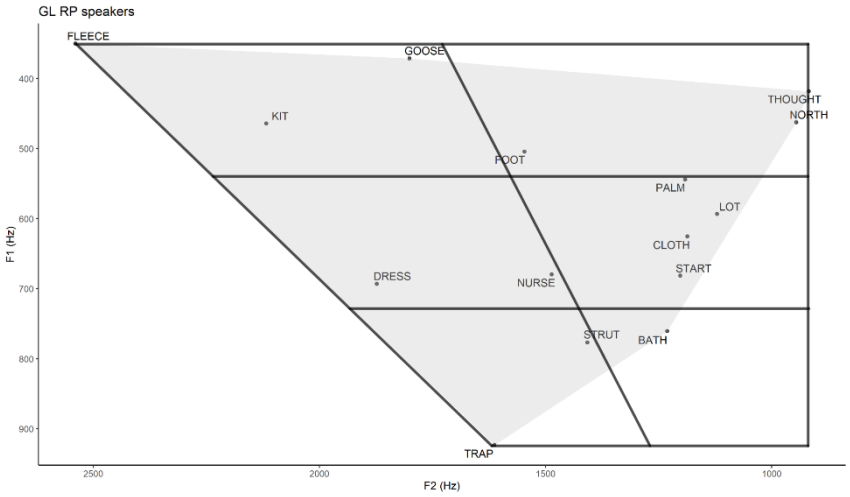

M21E2
GOOSE
KIT
FOOT
THOUGHT
FLEECE
NORTH
DRESS
CLOTH
LOT
START
STRUT
TRAP
BATH
F1 (Hz)
F2 (Hz)

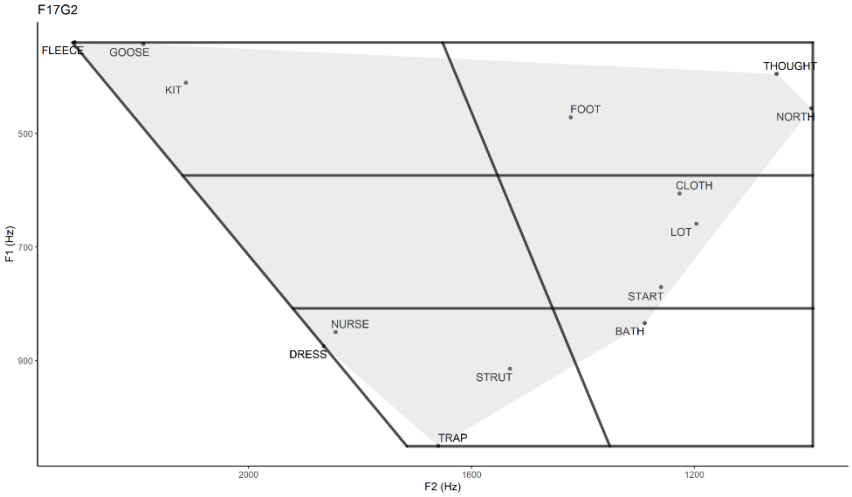

M22E2
GOOSE
FLEECE
NORTH
KIT
NURSE
THOUGHT
DRESS
LOT
CLOTH
STRUT
START
PALM
BATH
TRAP
F1 (Hz)
F2 (Hz)

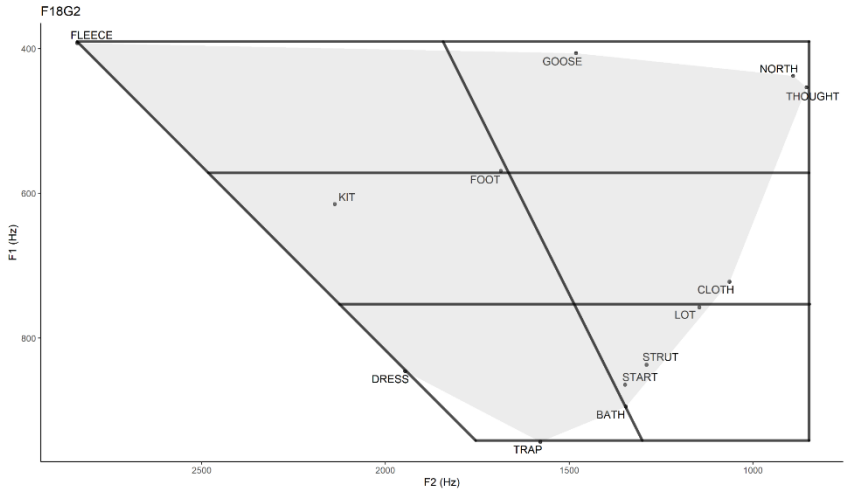

M23E2
FLEECE
GOOSE
KIT
SQUARE
FOOT
NORTH
THOUGHT
NURSE
CLOTH
LOT
DRESS
BATH
STRUT
START
TRAP
F1 (Hz)
F2 (Hz)

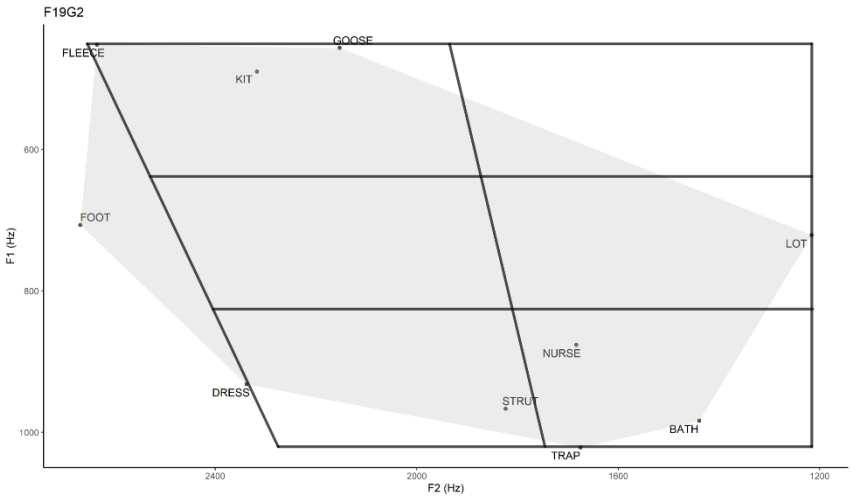

M24E2
FLEECE
GOOSE
NORTH
THOUGHT
KIT
FOOT
CLOTH
LOT
DRESS
NURSE
BATH
START
STRUT
TRAP
F1 (Hz)
F2 (Hz)
400
600
800
1000
2400
2000
1600
1200
800

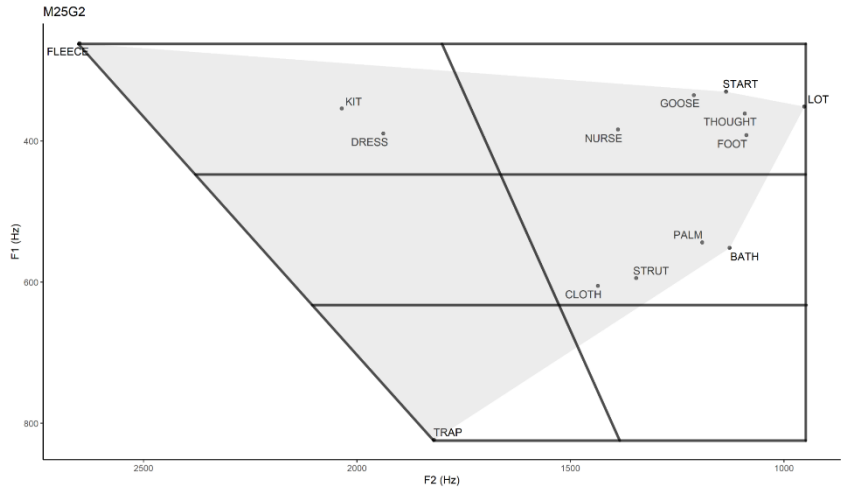
Green Line 2

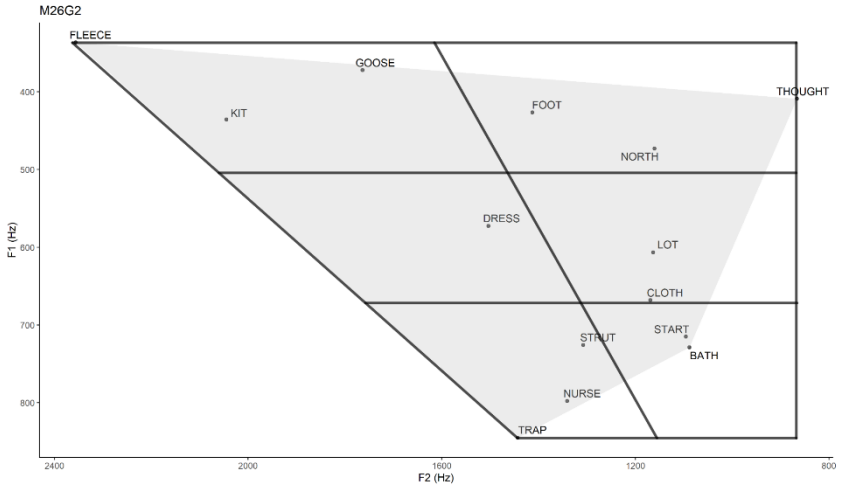

GL RP speakers
FLEECE
GOOSE
KIT
THOUGHT
NORTH
FOOT
PALM
LOT
CLOTH
START
DRESS
NURSE
STRUT
BATH
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2500
2000
1500
1000

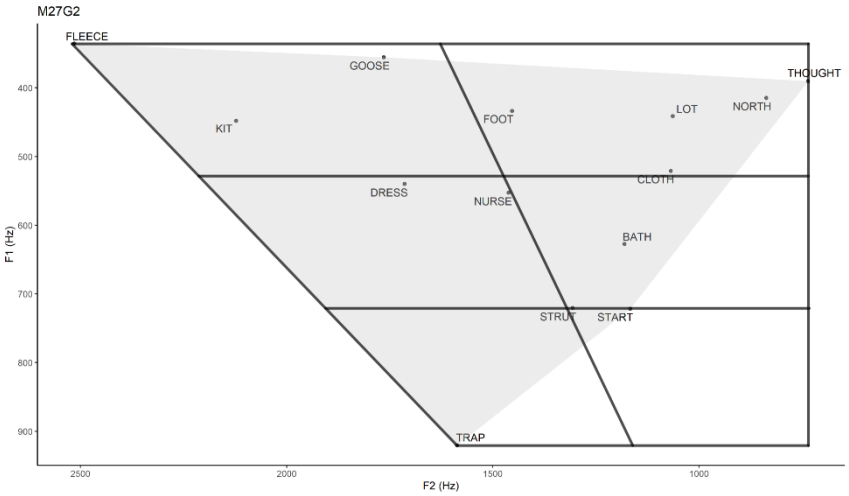

F17G2
FLEECE
GOOSE
KIT
THOUGHT
FOOT
NORTH
CLOTH
LOT
START
NURSE
BATH
DRESS
STRUT
TRAP
F1 (Hz)
F2 (Hz)
500
700
900
2000
1600
1200

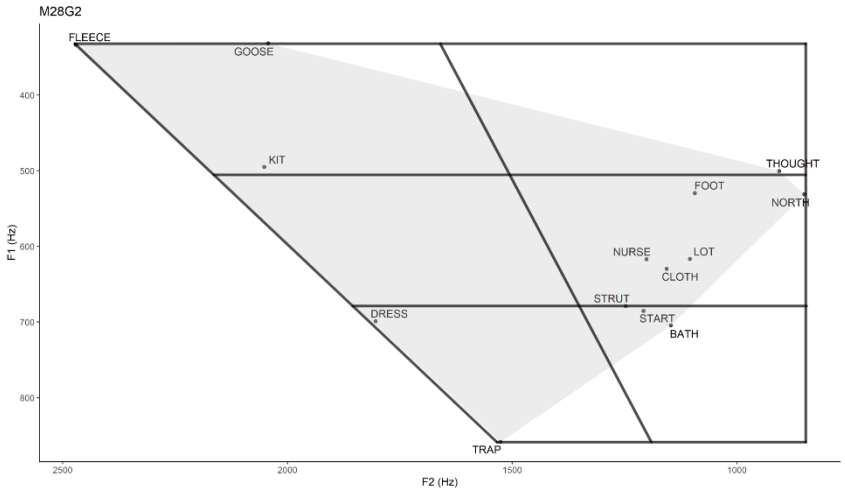

F18G2
FLEECE
GOOSE
NORTH
THOUGHT
FOOT
KIT
CLOTH
LOT
STRUT
START
DRESS
BATH
TRAP
400
600
800
F1 (Hz)
2500
2000
1500
1000
F2 (Hz)


F19G2
GOOSE
FLEECE
KIT
FOOT
LOT
NURSE
DRESS
STRUT
BATH
TRAP
600
800
1000
F1 (Hz)
2400
2000
1600
1200
F2 (Hz)

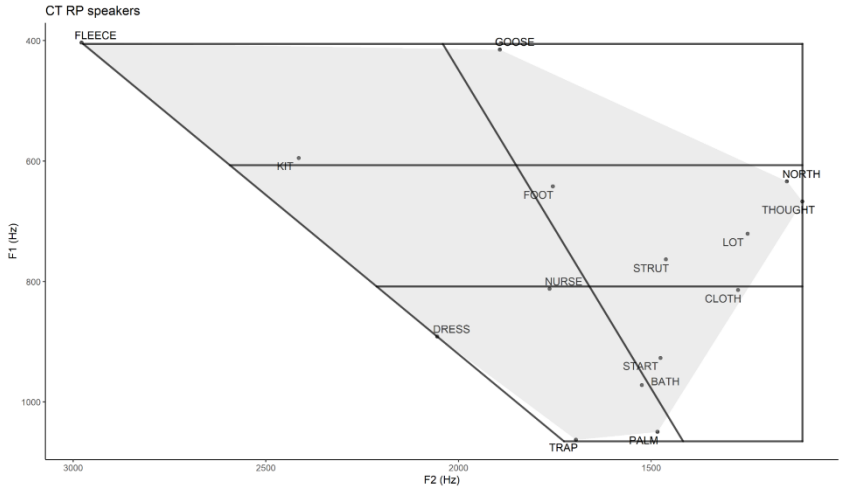

M25G2
FLEECE
START
KIT
GOOSE
LOT
THOUGHT
DRESS
NURSE
FOOT
PALM
BATH
STRUT
CLOTH
TRAP
400
600
800
F1 (Hz)
2500
2000
1500
1000
F2 (Hz)

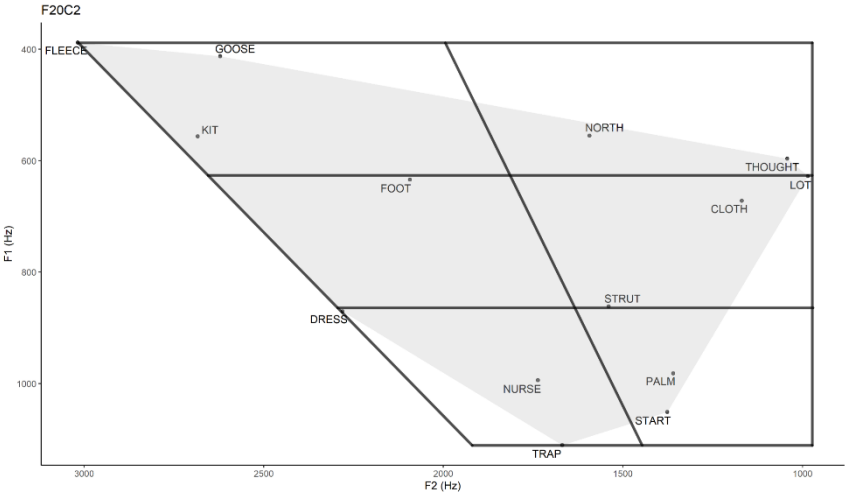

M26G2
FLEECE
GOOSE
THOUGHT
KIT
FOOT
NORTH
DRESS
LOT
CLOTH
START
STRUT
BATH
NURSE
TRAP
F1 (Hz)
F2 (Hz)

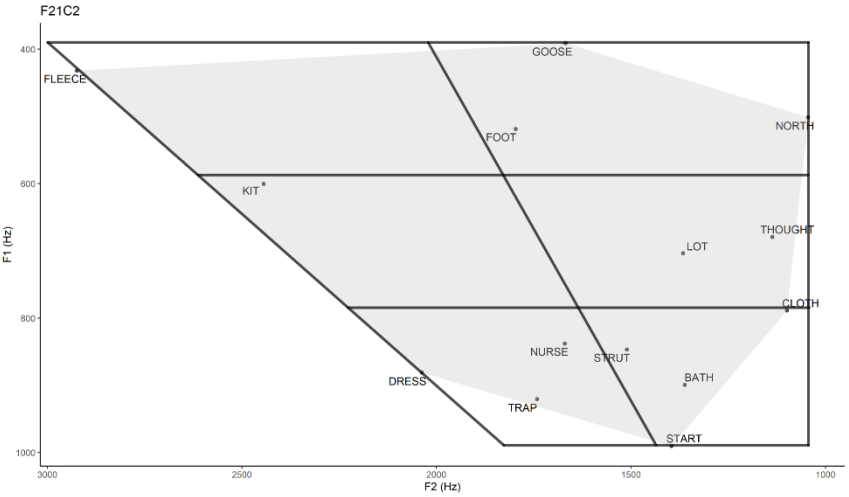

M27G2
FLEECE
GOOSE
THOUGHT
FOOT
LOT
NORTH
KIT
CLOTH
DRESS
NURSE
BATH
STRUT
START
TRAP
F1 (Hz)
F2 (Hz)

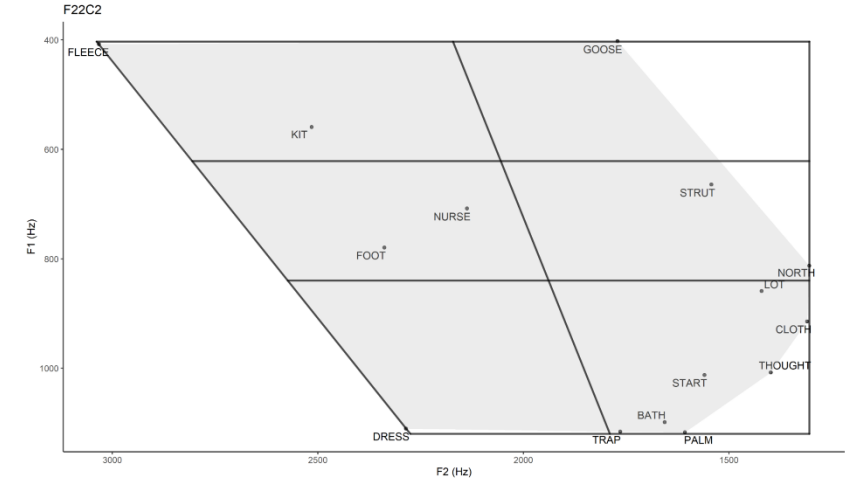

M28G2
FLEECE
GOOSE
KIT
THOUGHT
FOOT
NORTH
NURSE
LOT
CLOTH
STRUT
DRESS
START
BATH
TRAP
F1 (Hz)
F2 (Hz)

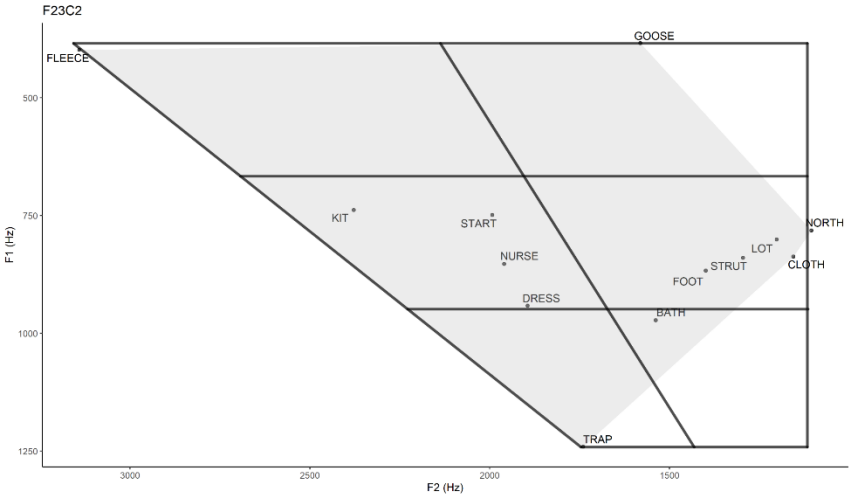
Camden Town 2

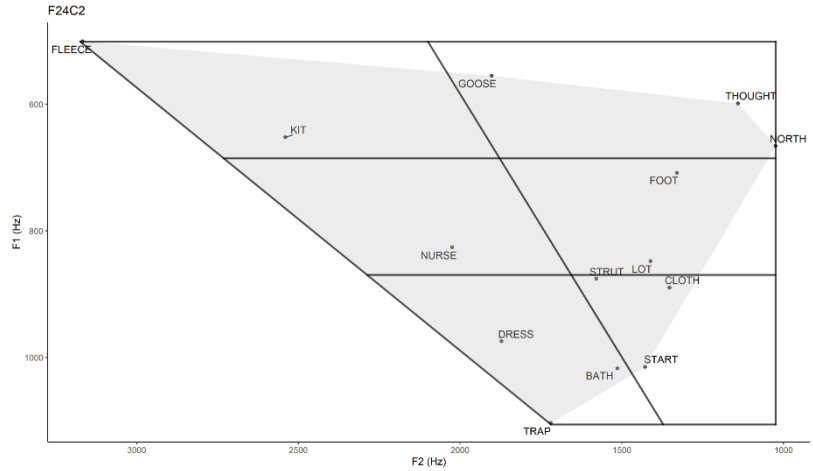


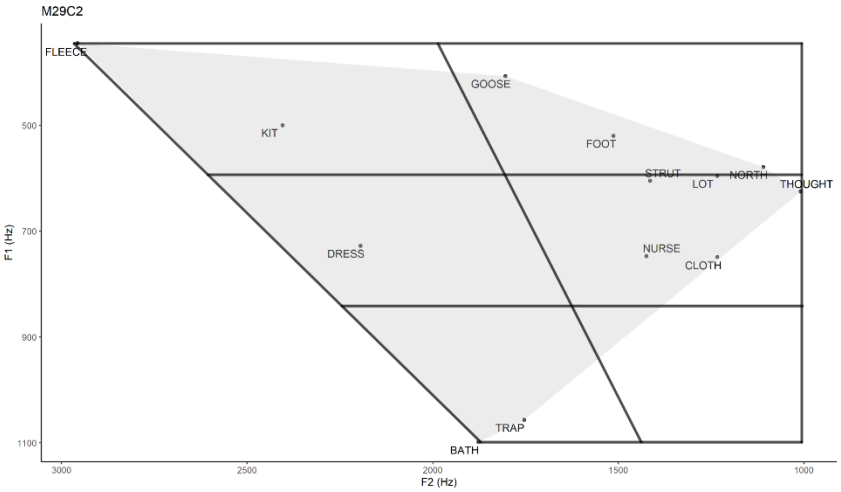


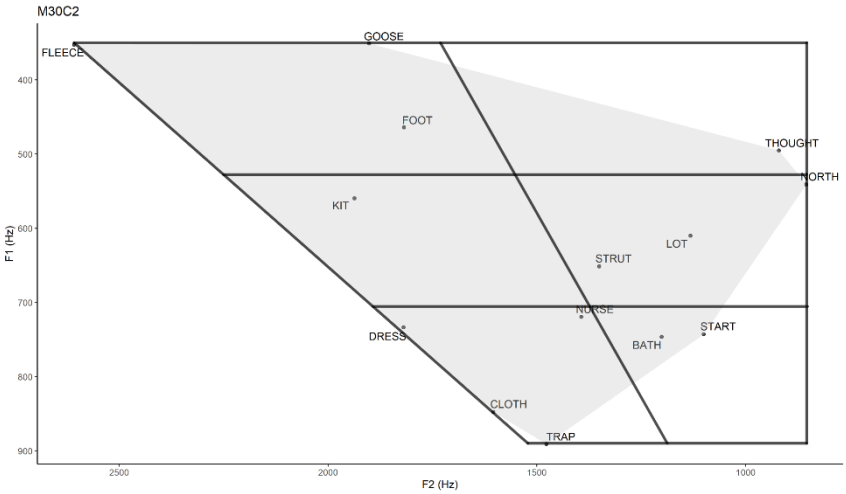



F22C2
FLEECE
GOOSE
KIT
STRUT
NURSE
FOOT
NORTH
LOT
CLOTH
THOUGHT
START
BATH
DRESS
TRAP
PALM
F1 (Hz)
F2 (Hz)
400
600
800
1000
3000
2500
2000
1500

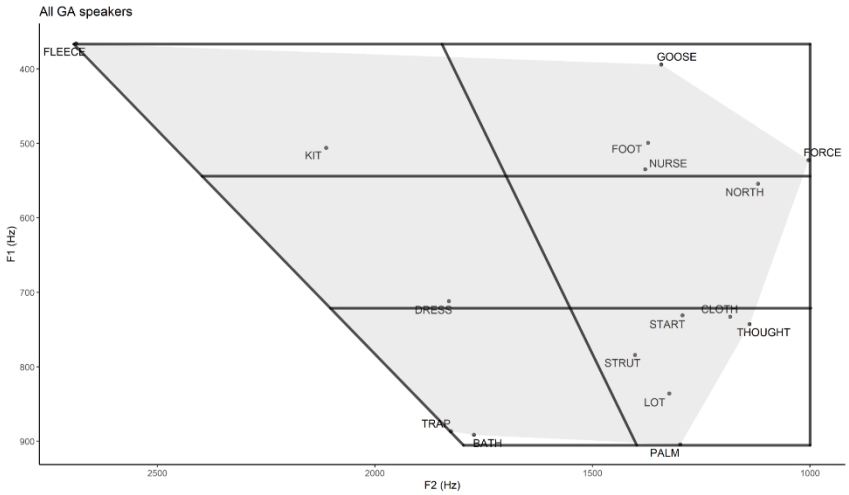

F23C2
GOOSE
FLEECE
KIT
START
NORTH
LOT
NURSE
STRUT
CLOTH
FOOT
DRESS
BATH
TRAP
F1 (Hz)
F2 (Hz)
500
750
1000
1250
3000
2500
2000
1500


F24C2
FLEECE
GOOSE
THOUGHT
KIT
NORTH
FOOT
NURSE
STRUT
LOT
CLOTH
DRESS
START
BATH
TRAP
F1 (Hz)
F2 (Hz)
600
800
1000
3000
2500
2000
1500
1000

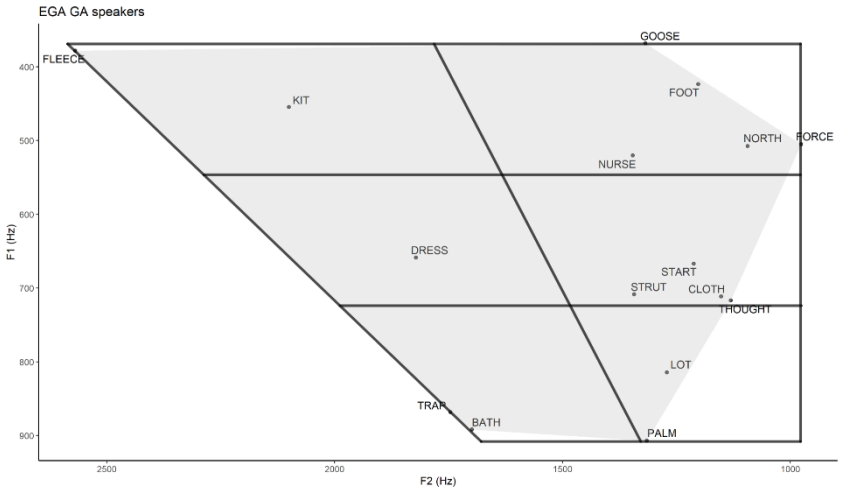

M29C2
FLEECE
GOOSE
KIT
FOOT
STRUT
LOT
NORTH
THOUGHT
DRESS
NURSE
CLOTH
TRAP
BATH
F1 (Hz)
F2 (Hz)
500
700
900
1100
3000
2500
2000
1500
1000

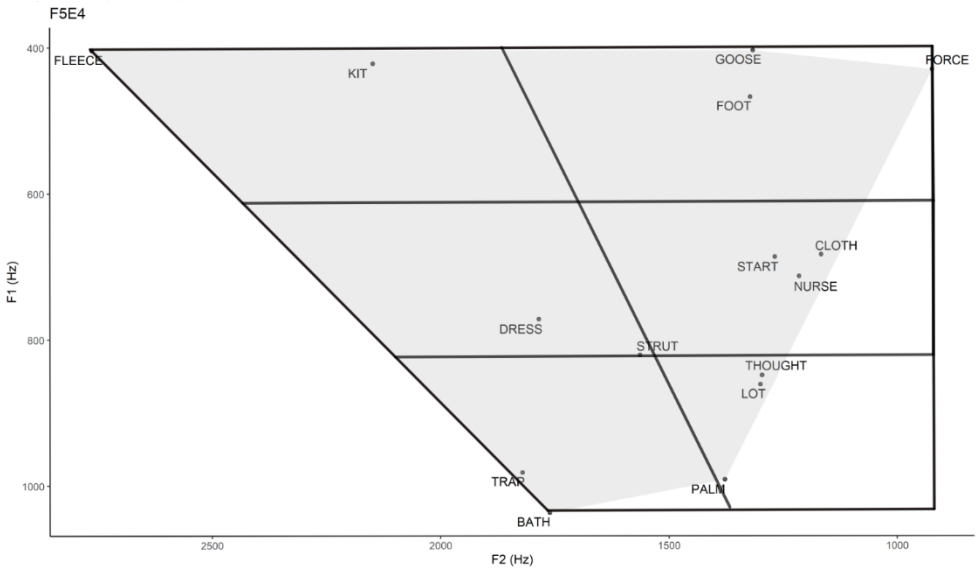

M30C2
GOOSE
FLEECE
FOOT
THOUGHT
NORTH
KIT
LOT
STRUT
NURSE
START
DRESS
BATH
CLOTH
TRAP
F1 (Hz)
F2 (Hz)
400
500
600
700
800
900
2500
2000
1500
1000

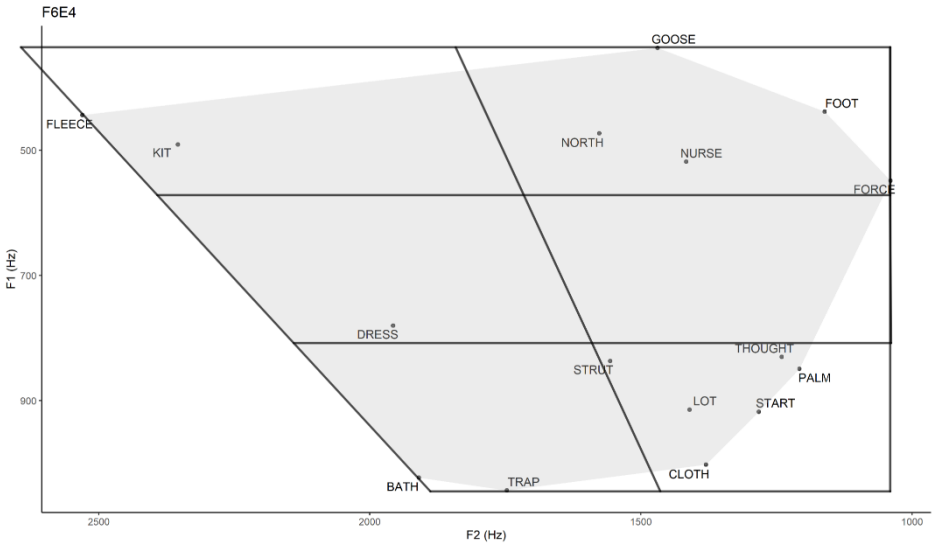
General American sub-corpus

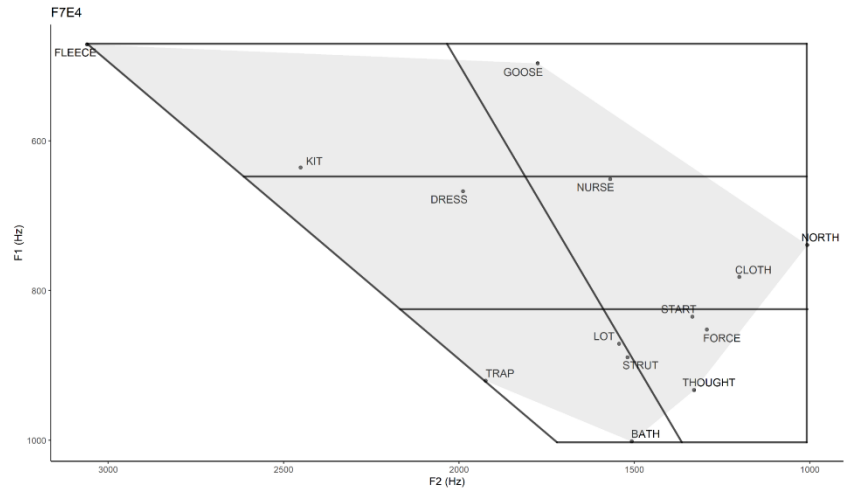


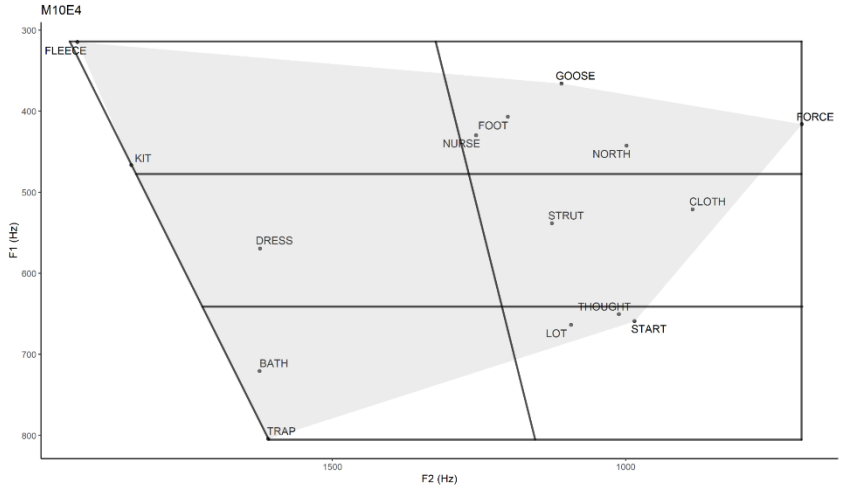
English G Access 4

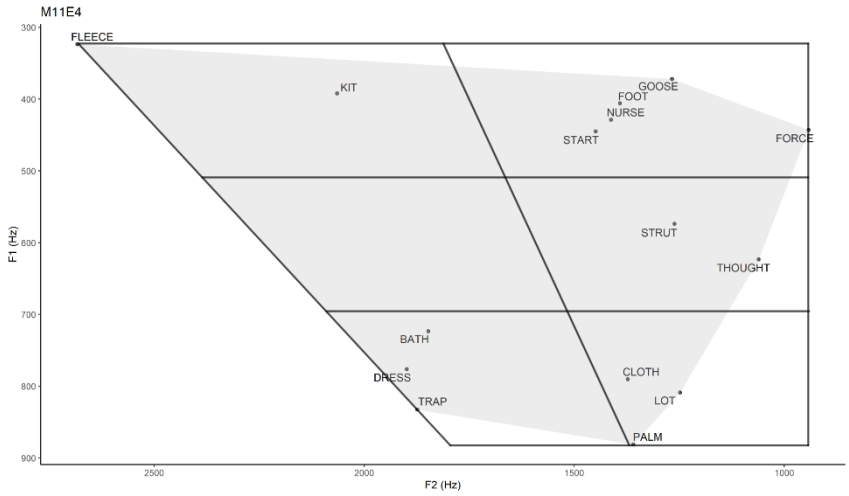


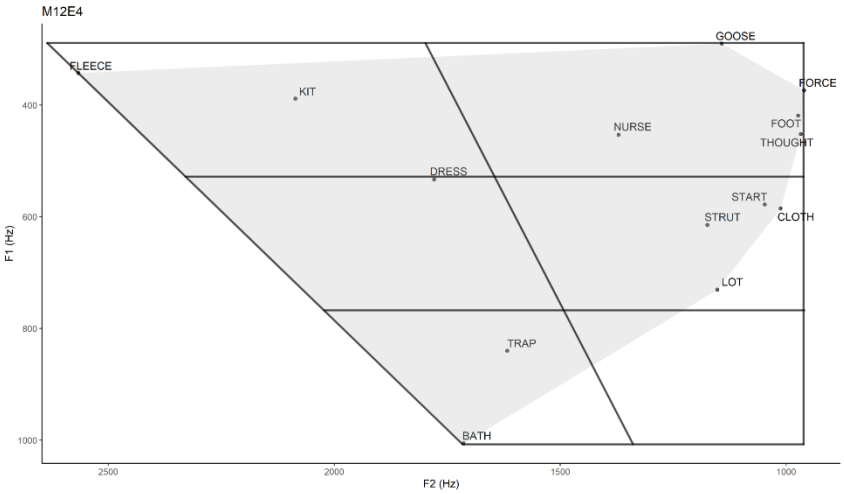

F5E4
FLEECE
KIT
GOOSE
FORCE
FOOT
CLOTH
START
NURSE
DRESS
STRUT
THOUGHT
LOT
TRAP
PALM
BATH
F1 (Hz)
F2 (Hz)
400
600
800
1000
2500
2000
1500
1000

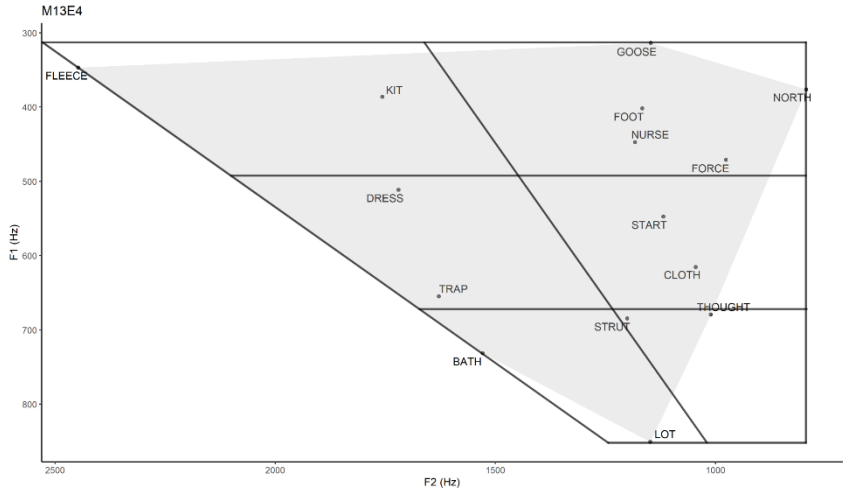

F6E4
GOOSE
FOOT
FLEECE
KIT
NORTH
NURSE
FORCE
DRESS
THOUGHT
STRUT
PALM
LOT
START
BATH
TRAP
CLOTH
F1 (Hz)
F2 (Hz)
500
700
900
2500
2000
1500
1000


F7E4
FLEECE
GOOSE
KIT
NURSE
DRESS
NORTH
CLOTH
START
LOT
FORCE
STRUT
TRAP
THOUGHT
BATH
F1 (Hz)
F2 (Hz)
600
800
1000
3000
2500
2000
1500
1000

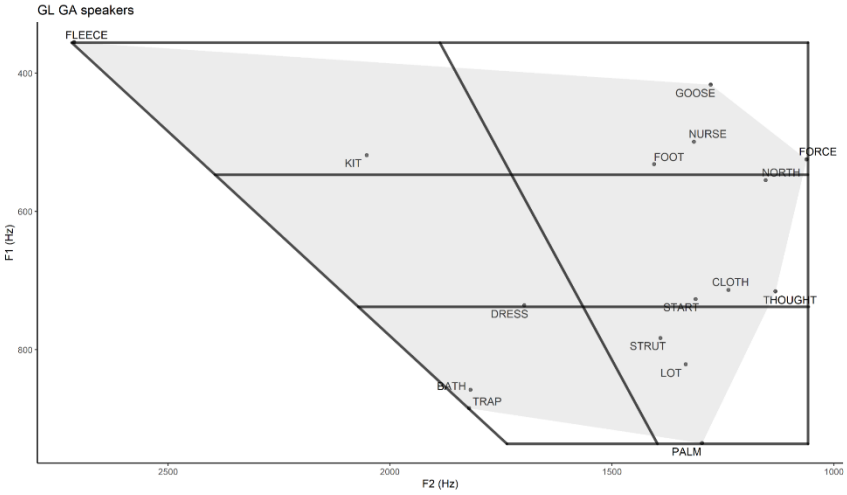

M10E4
FLEECE
GOOSE
FORCE
FOOT
NURSE
NORTH
KIT
CLOTH
STRUT
DRESS
THOUGHT
LOT
START
BATH
TRAP
F1 (Hz)
F2 (Hz)

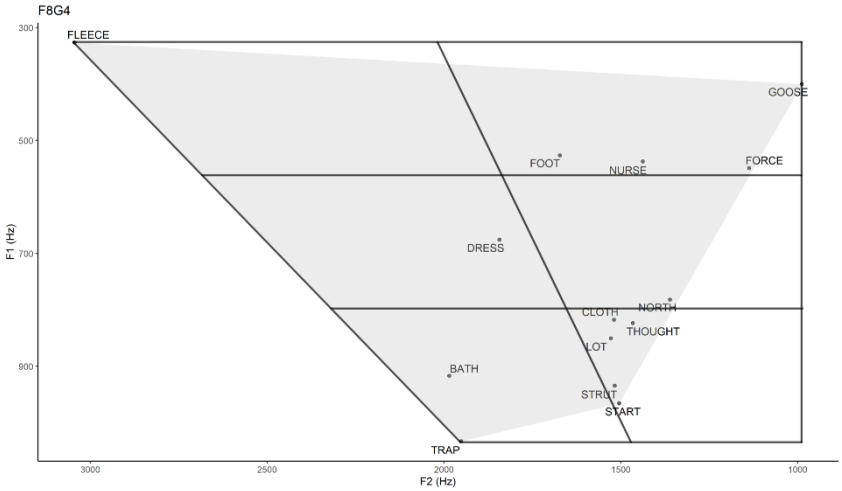

M11E4
FLEECE
KIT
GOOSE
FOOT
NURSE
START
FORCE
STRUT
THOUGHT
BATH
DRESS
CLOTH
TRAP
LOT
PALM
F1 (Hz)
F2 (Hz)

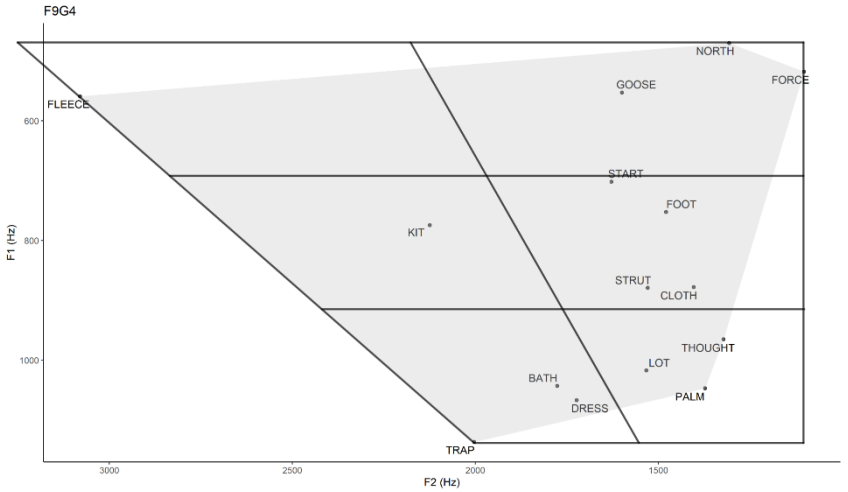

M12E4
GOOSE
FLEECE
KIT
FORCE
NURSE
FOOT
THOUGHT
DRESS
START
STRUT
CLOTH
LOT
TRAP
BATH
F1 (Hz)
F2 (Hz)

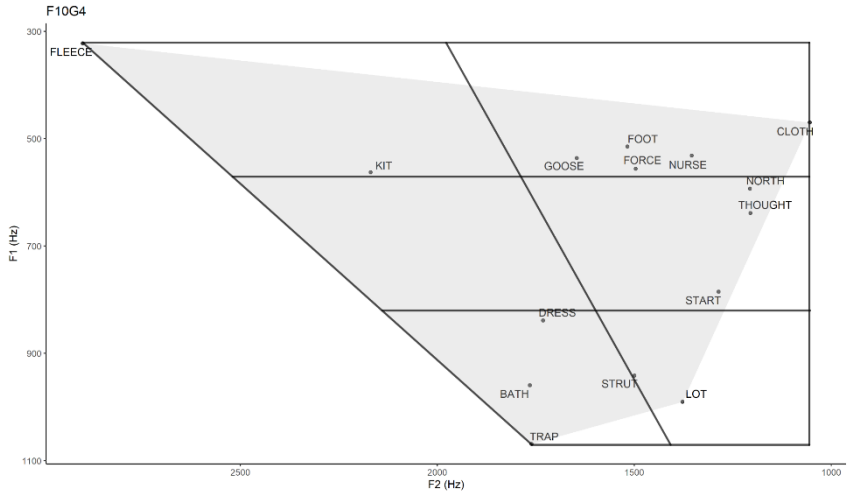

M13E4
FLEECE
KIT
GOOSE
NORTH
FOOT
NURSE
FORCE
DRESS
START
CLOTH
TRAP
THOUGHT
STRUT
BATH
LOT
F1 (Hz)
F2 (Hz)

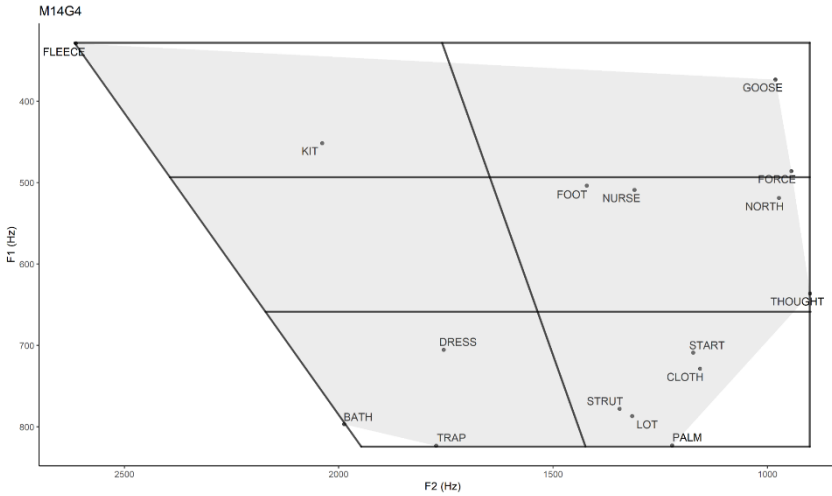
Green Line 4

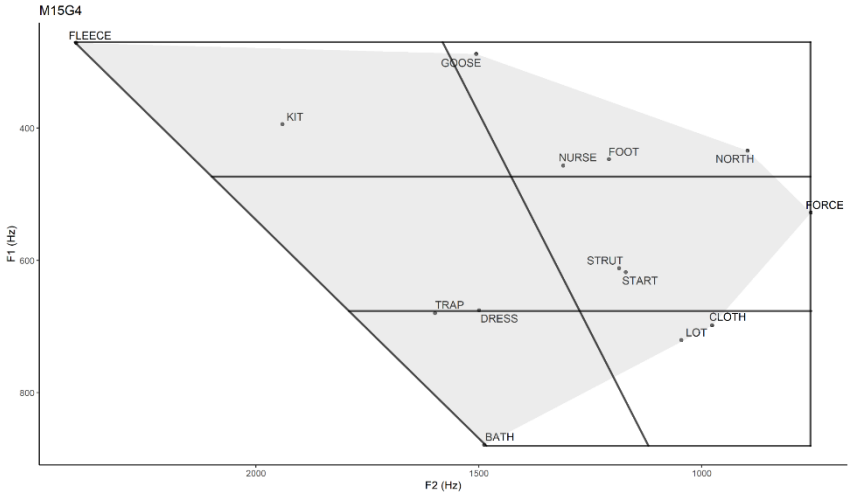

GL GA speakers
FLEECE
GOOSE
NURSE
KIT
FOOT
FORCE
NORTH
CLOTH
THOUGHT
DRESS
START
STRUT
LOT
BATH
TRAP
PALM
F1 (Hz)
F2 (Hz)

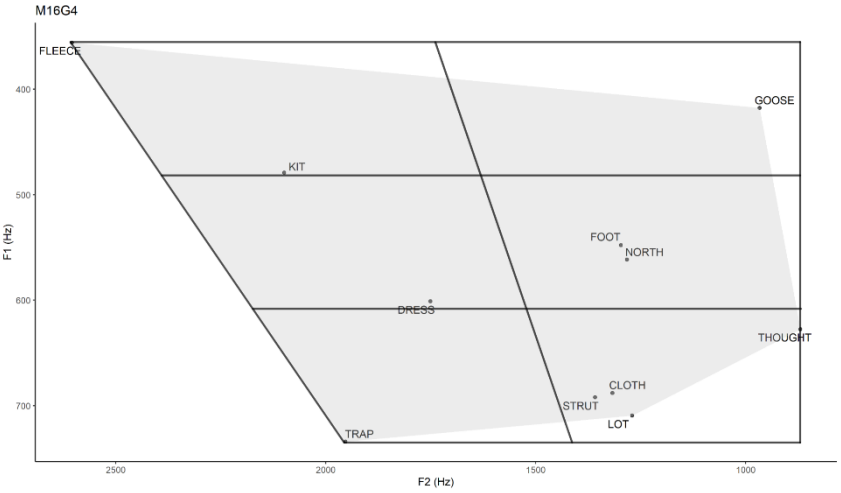

F8G4
FLEECE
GOOSE
FOOT
NURSE
FORCE
DRESS
CLOTH
NORTH
THOUGHT
LOT
BATH
STRUT
START
TRAP
F1 (Hz)
F2 (Hz)

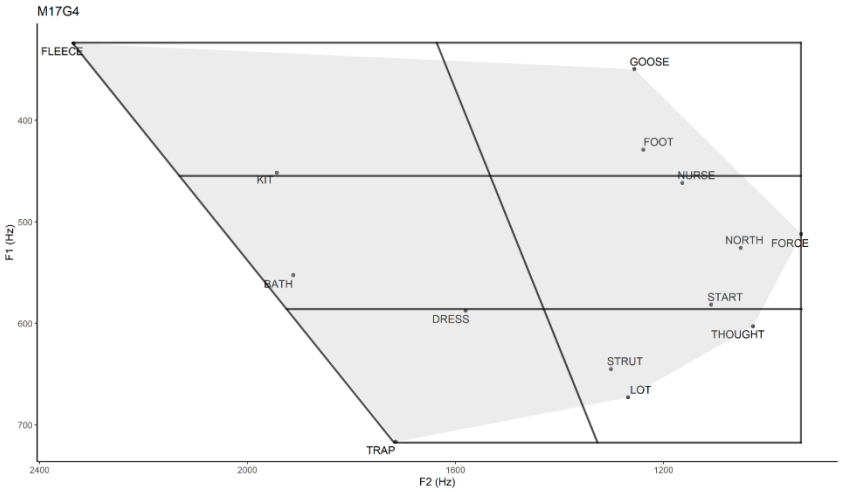

F9G4
NORTH
GOOSE
FORCE
FLEECE
START
FOOT
KIT
STRUT
CLOTH
THOUGHT
LOT
BATH
PALM
DRESS
TRAP
F1 (Hz)
F2 (Hz)


F10G4
FLEECE
CLOTH
FOOT
KIT
GOOSE
FORCE
NURSE
NORTH
THOUGHT
START
DRESS
STRUT
BATH
LOT
TRAP
F1 (Hz)
F2 (Hz)

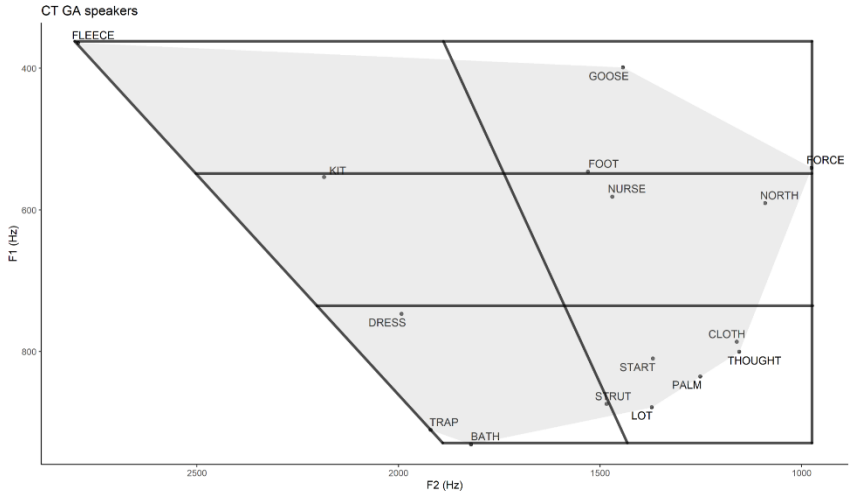

M14G4
FLEECE
GOOSE
KIT
FORCE
FOOT
NURSE
NORTH
THOUGHT
DRESS
START
CLOTH
STRUT
BATH
LOT
TRAP
PALM
F1 (Hz)
F2 (Hz)

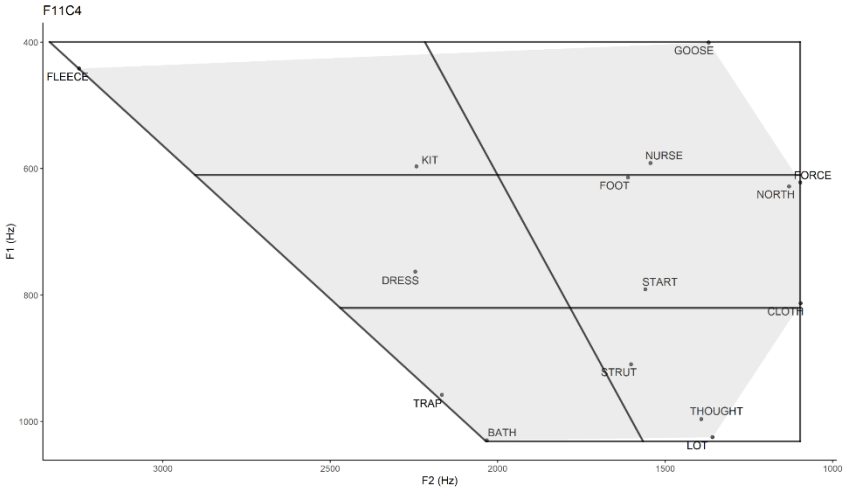

M15G4
FLEECE
GOOSE
KIT
NURSE
FOOT
NORTH
FORCE
STRUT
START
TRAP
DRESS
CLOTH
LOT
BATH
F1 (Hz)
F2 (Hz)

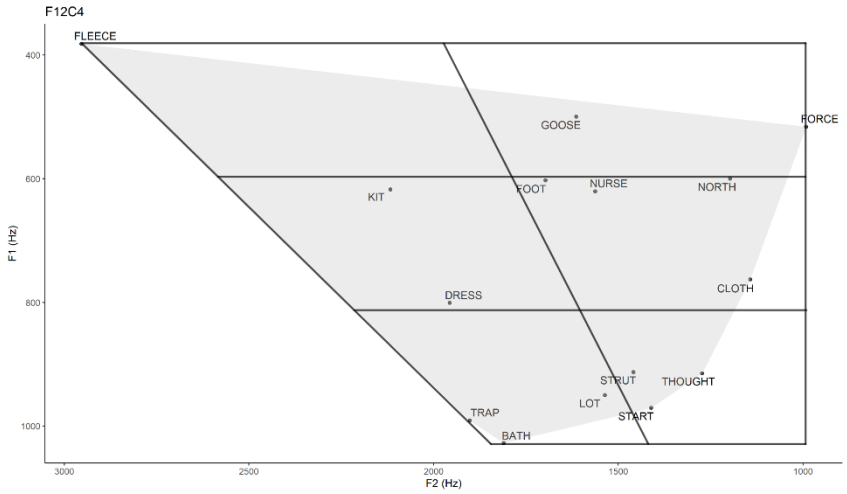

M16G4
FLEECE
GOOSE
KIT
FOOT
NORTH
DRESS
THOUGHT
CLOTH
STRUT
LOT
TRAP
F1 (Hz)
F2 (Hz)

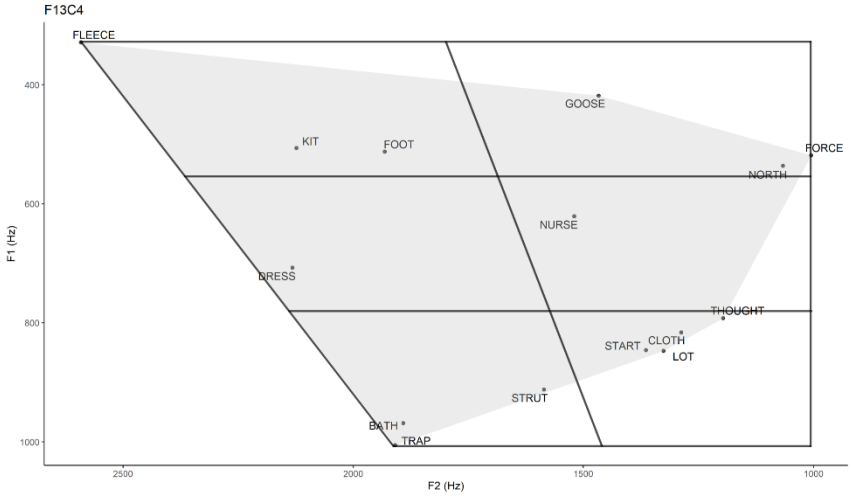

M17G4
FLEECE
GOOSE
FOOT
KIT
NURSE
NORTH
FORCE
BATH
START
DRESS
THOUGHT
STRUT
LOT
TRAP
F1 (Hz)
F2 (Hz)

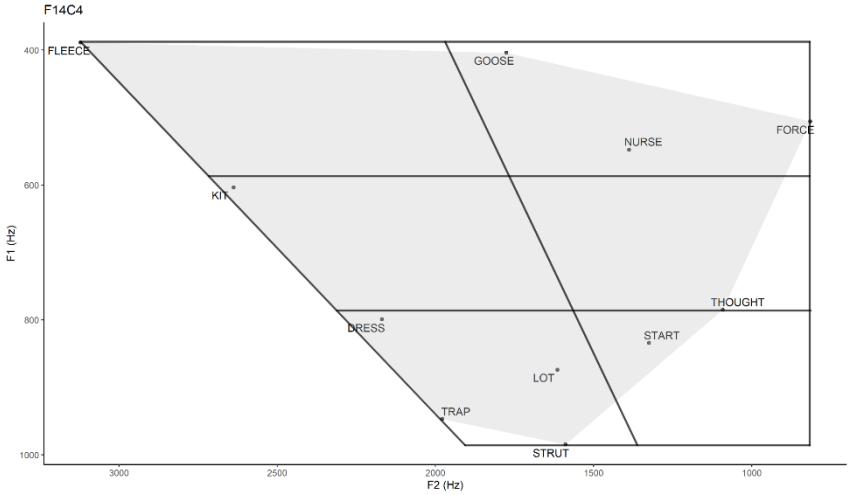
Camden Town 4

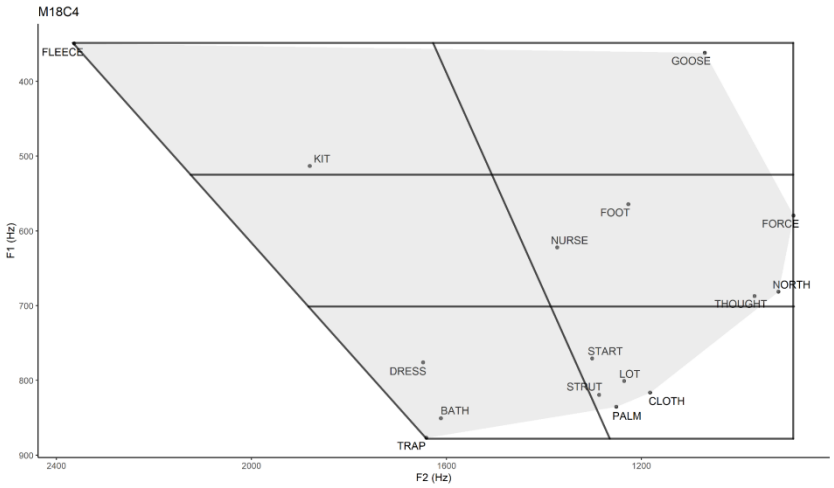


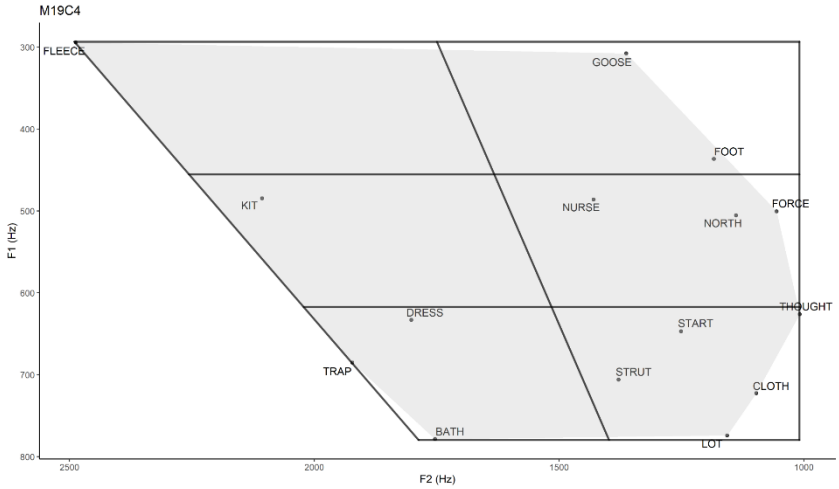




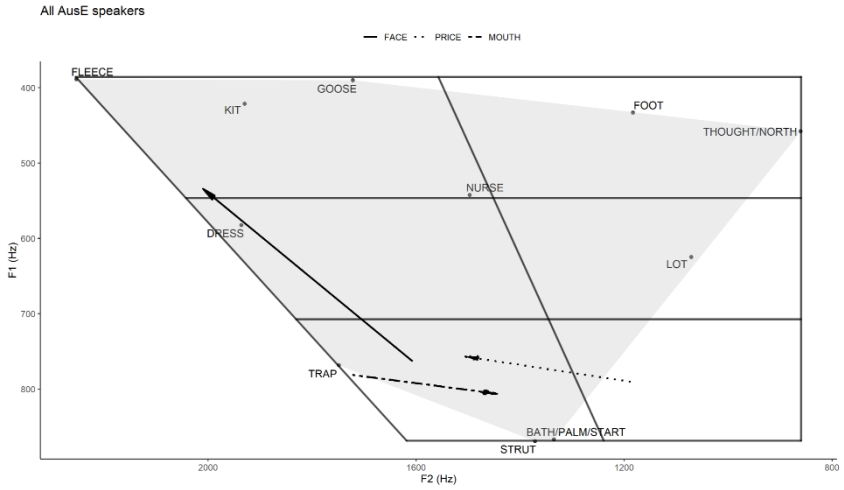

F13C4
FLEECE
GOOSE
KIT
FOOT
FORCE
NORTH
NURSE
DRESS
THOUGHT
START
CLOTH
LOT
STRUT
BATH
TRAP
F1 (Hz)
F2 (Hz)


F14C4
FLEECE
GOOSE
FORCE
NURSE
KIT
THOUGHT
DRESS
START
LOT
TRAP
STRUT
F1 (Hz)
F2 (Hz)

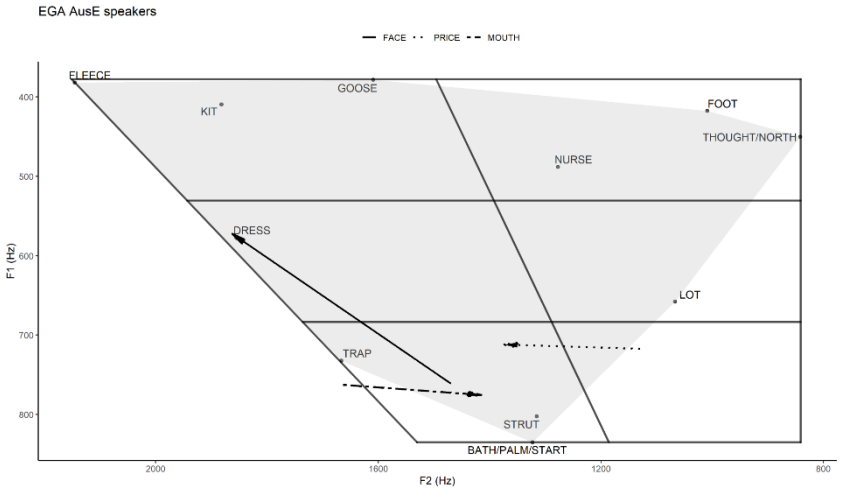

M18C4
FLEECE
GOOSE
KIT
FOOT
FORCE
NURSE
NORTH
THOUGHT
START
DRESS
LOT
STRUT
CLOTH
BATH
PALM
TRAP
F1 (Hz)
F2 (Hz)


M19C4
FLEECE
GOOSE
FOOT
KIT
NURSE
FORCE
NORTH
DRESS
THOUGHT
START
TRAP
STRUT
CLOTH
BATH
LOT
300
400
500
600
700
800
2500
2000
1500
1000
F1 (Hz)
F2 (Hz)

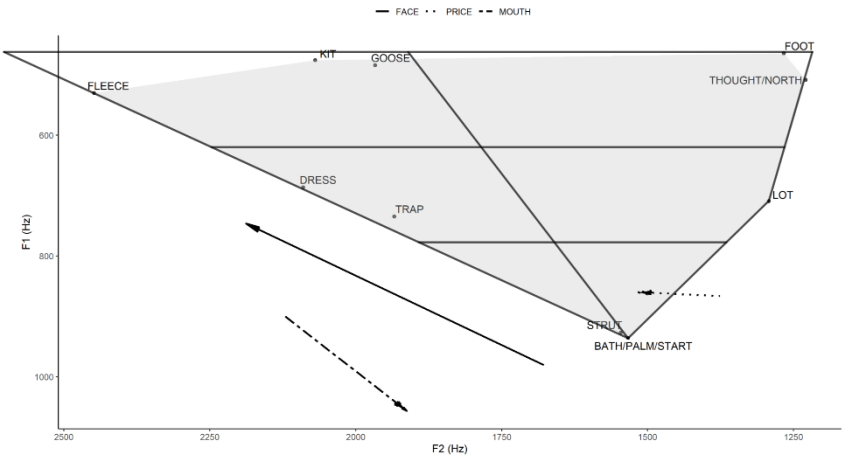
Australian English sub-corpus



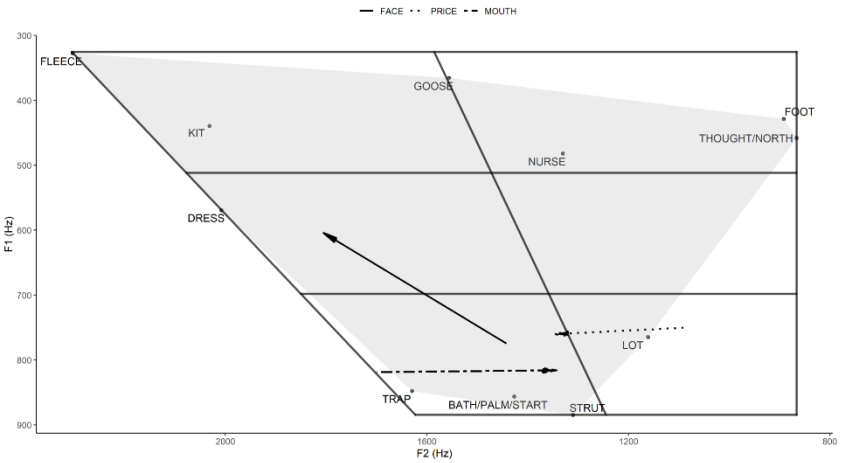
English G Access 5



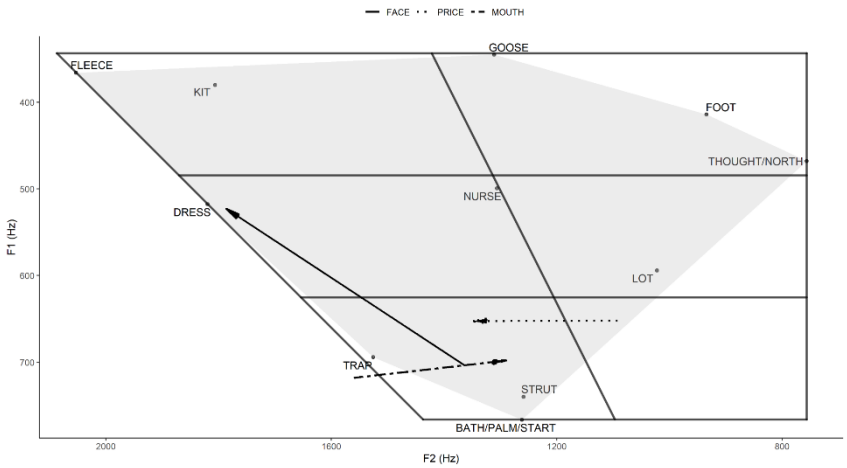
F1E5



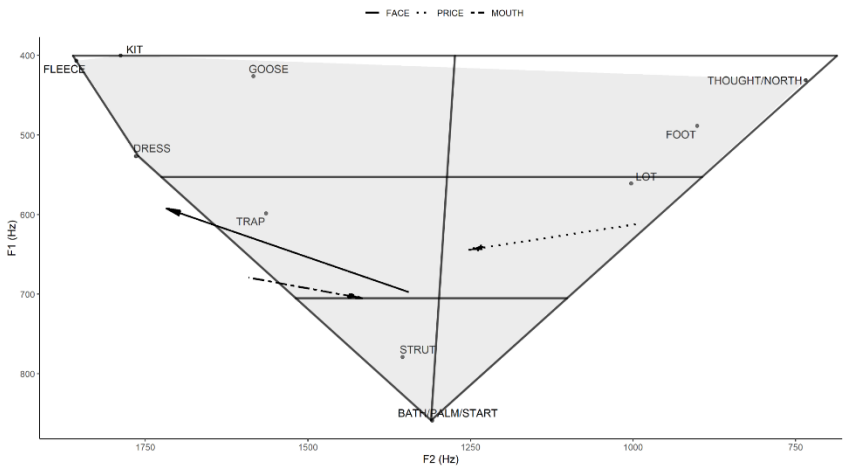
M1E5



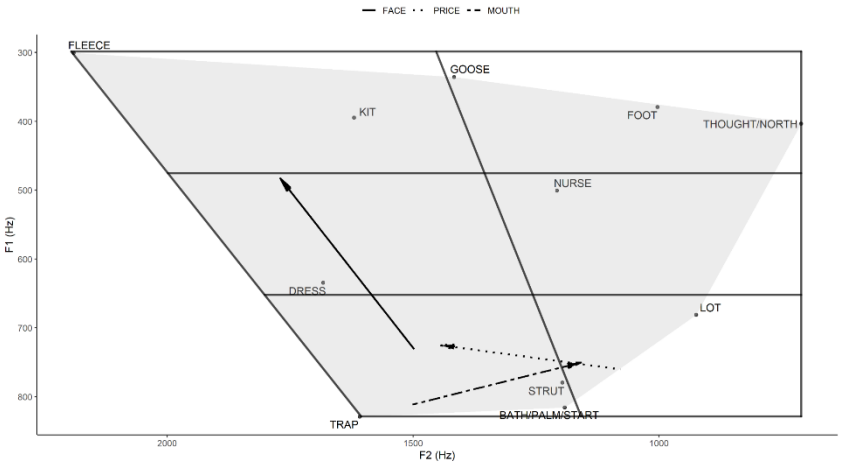
M2E5



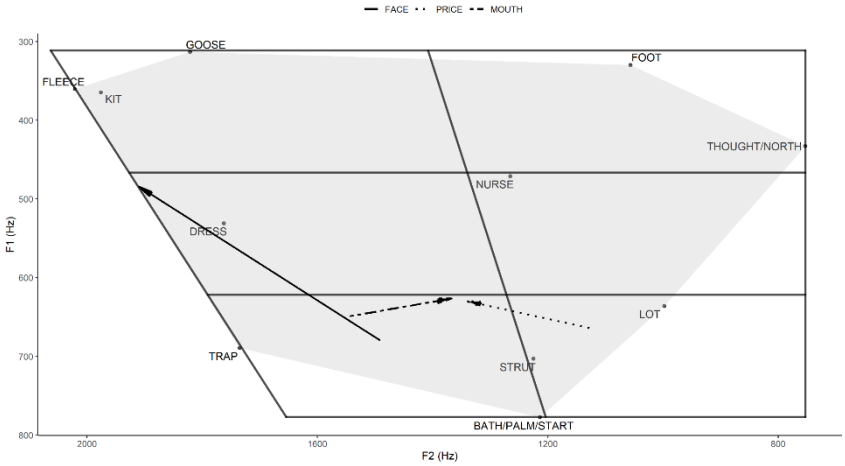
M3E5



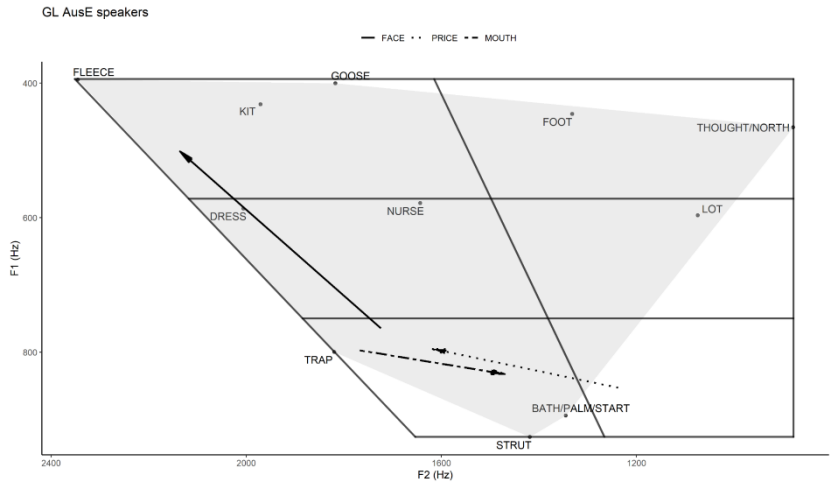
M4E5

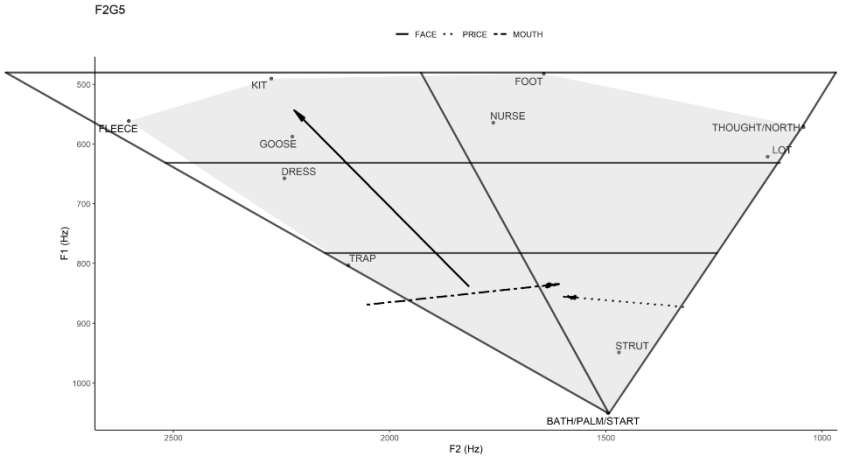


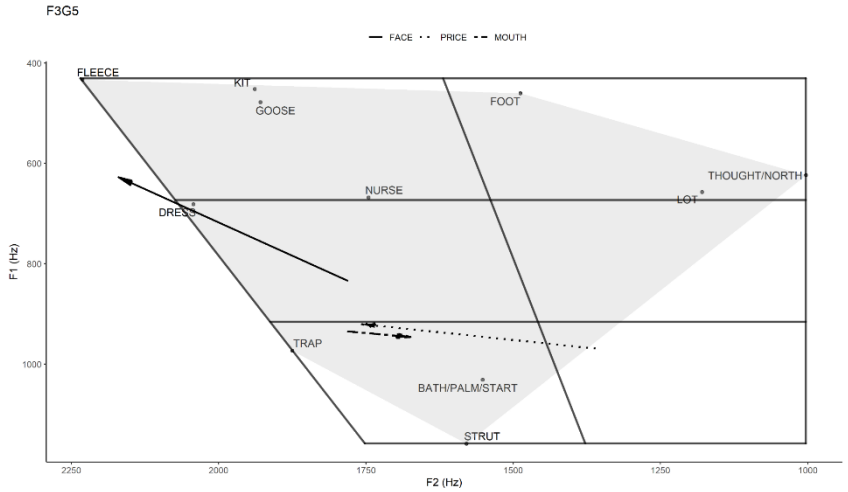
M5E5



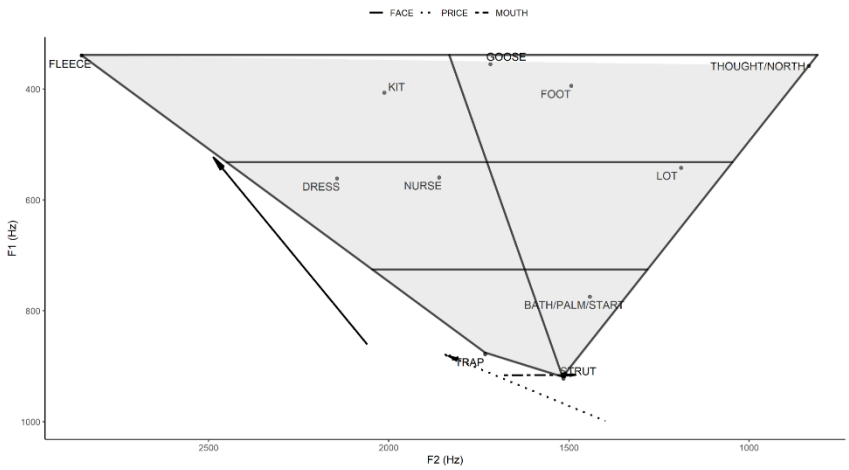
Green Line 5


GL AusE speakers
FACE PRICE MOUTH
FLEECE
GOOSE
KIT
FOOT
THOUGHT/NORTH
DRESS
NURSE
LOT
TRAP
BATH/PALM/START
STRUT
F1 (Hz)
F2 (Hz)

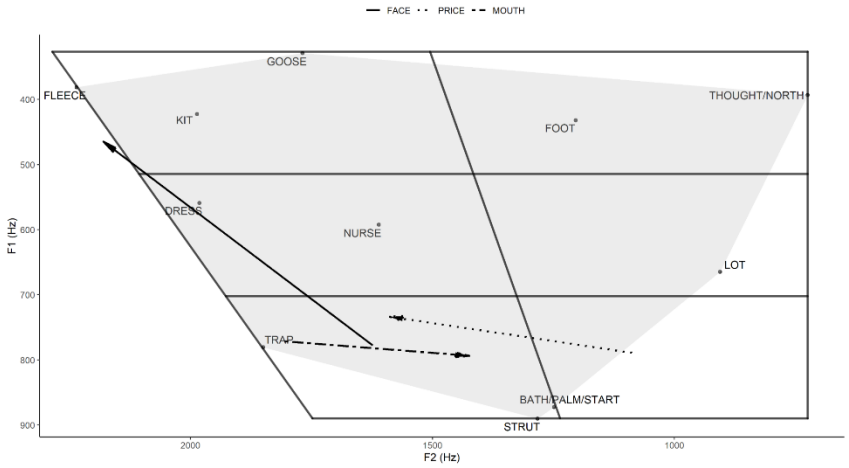

F2G5
FACE PRICE MOUTH
KIT
FOOT
NURSE
FLEECE
GOOSE
THOUGHT/NORTH
LOT
DRESS
TRAP
STRUT
BATH/PALM/START
F1 (Hz)
F2 (Hz)


F3G5
FACE PRICE MOUTH
FLEECE
KIT
GOOSE
FOOT
THOUGHT/NORTH
NURSE
LOT
DRESS
TRAP
BATH/PALM/START
STRUT
F1 (Hz)
F2 (Hz)

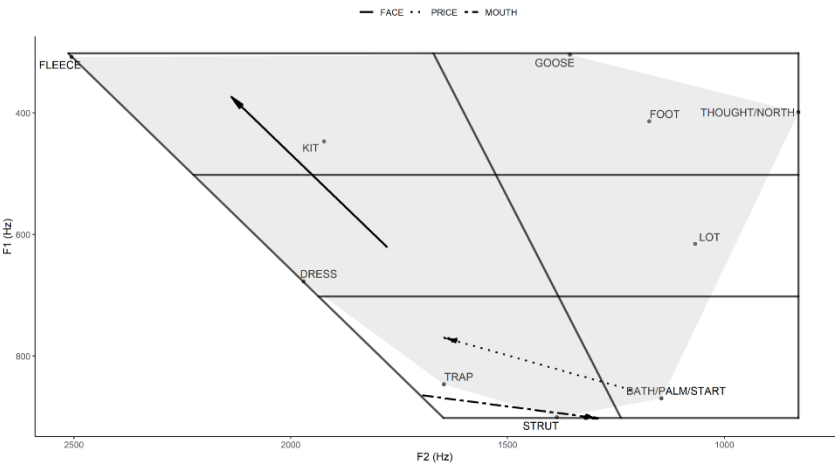
F4G5



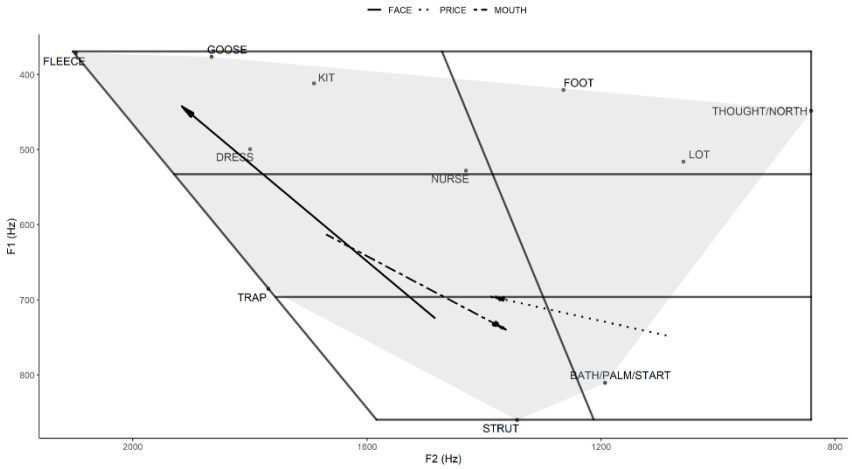
M6G5



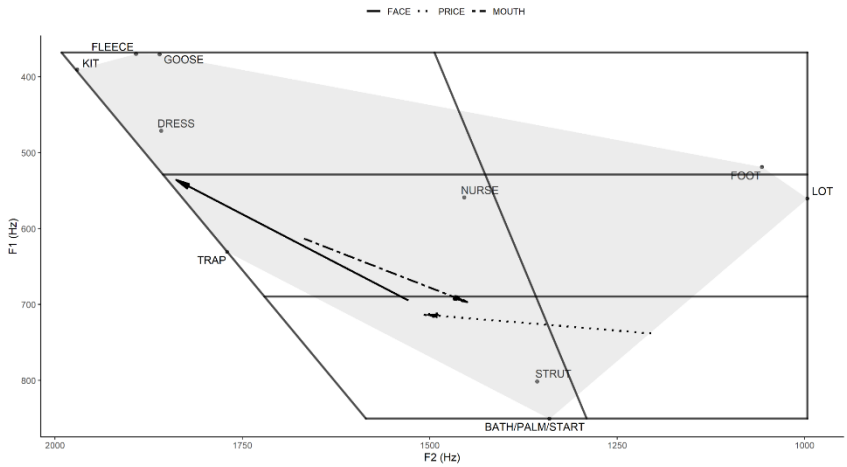
M7G5



M8G5


FACE
PRICE
MOUTH
FLEECE
GOOSE
KIT
FOOT
THOUGHT/NORTH
DRESS
LOT
NURSE
TRAP
BATH/PALM/START
STRUT
F1 (Hz)
F2 (Hz)
400
500
600
700
800
2000
1600
1200
800

M9G5


FACE
PRICE
MOUTH
FLEECE
KIT
GOOSE
DRESS
FOOT
LOT
NURSE
TRAP
STRUT
BATH/PALM/START
F1 (Hz)
F2 (Hz)
400
500
600
700
800
2000
1750
1500
1250
1000

Appendix 3

Praat Scripts

This appendix includes the two Praat scripts that were used to extract the formant values. Praat Script 1 was used for the Australian English sub-corpus. This script was subsequently adapted to simplify and expedite the analysis process, with leaving the core mechanics intact.[1] This Praat Script 2 was used for the analysis of the General American and Received Pronunciation corpus data. The '#' symbol indicates comments in the script. These lines are not functional and are only used to describe or explain the following lines of code.

[1] A huge thanks to Noa Ibrahim, our student assistant, who took my original script and adapted it to streamline the analysis process.

Praat Script 1

```
#This script was created 24th January 2019 by Lisa Scheiwe.
#adapt file path
#open sound file and text grid in Praat. Select both, open script, run script.

thisSound$ = selected$ ("Sound")
thisTextGrid$ = selected$ ("TextGrid")

#set up default values for formant analysis. Do not change these!
	timeStepDefault = 0
	numFormantsDefault = 5
	maxFormantDefault = 5500
	windowLengthDefault = 0.025
	preEmphasisDefault = 50

#Set up formant variables; change values if necessary, for male speakers, maxFormant=5000
		timeStep = timeStepDefault
		numFormants = numFormantsDefault
		maxFormant = maxFormantDefault
		windowLength = windowLengthDefault
		preEmphasis = preEmphasisDefault

	#create a formant object
	select Sound 'thisSound$'
	formantObj = To Formant (burg): timeStep, numFormants, maxFormant, windowLength, preEmphasis
	thisFormant$ = selected$ ("Formant")

select TextGrid 'thisTextGrid$'


	numberOfTiers = Get number of tiers

	#This only works if tier 1 = speaker, tier 2 = word, tier3 = lexical set
		for tiernumber from 1 to numberOfTiers
			numberofintervals = Get number of intervals: tiernumber
			for intervalnumber from 1 to numberofintervals
```

```
select TextGrid 'thisTextGrid$'

	if tiernumber = 1
		speaker$ = Get label of interval: tiernumber, intervalnumber
		if speaker$ = ""
		ignore = 1
		else
		fileappend "C:\Users\Lisa Scheiwe\sciebo\PhD\Schoolbooks\Australian English\English G Access\Analyse Transcription\EGA5-5\speakerresults.txt" 'thisTextGrid$' 'tab$' 'speaker$' 'newline$'
		endif

	elsif tiernumber = 2
		word$ = Get label of interval: tiernumber, intervalnumber
		if word$ = ""
		ignore = 1
		else
			start_time_word = Get start point: tiernumber, intervalnumber
			end_time_word = Get end point: tiernumber, intervalnumber
		fileappend "C:\Users\Lisa Scheiwe\sciebo\PhD\Schoolbooks\Australian English\English G Access\Analyse Transcription\EGA5-5\wordresults.txt" 'word$' 'tab$' 'start_time_word:3' 'tab$'  'end_time_word:3' 'newline$'
		endif

	else tiernumber = 3
		begin = Get start point: tiernumber, intervalnumber
		end = Get end point: tiernumber, intervalnumber
		dur = (end-begin)
		mid = begin+(dur/2)
		onset = begin+(dur*0.3)
		glide = end-(dur*0.3)

		vowel$ = Get label of interval: tiernumber, intervalnumber
		if vowel$ = ""
		ignore = 1
		else
			select Formant 'thisFormant$'

				of1 = Get value at time: 1, 'onset', "Hertz", "Linear"
				of2 = Get value at time: 2, 'onset', "Hertz", "Linear"
				of3 = Get value at time: 3, 'onset', "Hertz", "Linear"
```

```
mf1 = Get value at time: 1, 'mid', "Hertz", "Linear"
mf2 = Get value at time: 2, 'mid', "Hertz", "Linear"
mf3 = Get value at time: 3, 'mid', "Hertz", "Linear"


gf1 = Get value at time: 1, 'glide', "Hertz", "Linear"
gf2 = Get value at time: 2, 'glide', "Hertz", "Linear"
gf3 = Get value at time: 3, 'glide', "Hertz", "Linear"
fileappend "C:\Users\Lisa Scheiwe\sciebo\PhD\Schoolbooks\Australian English\English G Access\Analyse Transcription\EGA5-5\vowelresults.txt" 'vowel$' 'tab$' 'begin:3' 'tab$' 'end:3' 'tab$' 'dur:3' 'tab$' 'onset:3' 'tab$' 'of1:0' 'tab$' 'of2:0' 'tab$' 'of3:0' 'tab$' 'mid:3' 'tab$' 'mf1:0' 'tab$' 'mf2:0' 'tab$' 'mf3:0' 'tab$' 'glide:3' 'tab$' 'gf1:0' 'tab$' 'gf2:0' 'tab$' 'gf3:0' 'newline$'
                endif
            endif
        endfor
    endfor
```

Praat Script 2

#This script was created and adapted from Script 1 by Noa Ibrahim 30th January 2024.

Enter file path and audio track name here

path$ = "PUT-PATH-HERE"

track$ = "PUT-TRACK-HERE"

Enter speaker codes of male and female speakers here. Remove # if more speakers are needed.

```
maleSpeaker1$ = "PUT-SPEAKER-CODE-HERE"
maleSpeaker2$ = "PUT-SPEAKER-CODE-HERE"
maleSpeaker3$ = "PUT-SPEAKER-CODE-HERE"
#maleSpeaker4$ = "PUT-SPEAKER-CODE-HERE"
#maleSpeaker5$ = "PUT-SPEAKER-CODE-HERE"
#maleSpeaker6$ = "PUT-SPEAKER-CODE-HERE"
#maleSpeaker7$ = "PUT-SPEAKER-CODE-HERE"

femaleSpeaker1$ = "PUT-SPEAKER-CODE-HERE"
femaleSpeaker2$ = "PUT-SPEAKER-CODE-HERE"
femaleSpeaker3$ = "PUT-SPEAKER-CODE-HERE"
#femaleSpeaker4$ = "PUT-SPEAKER-CODE-HERE"
#femaleSpeaker5$ = "PUT-SPEAKER-CODE-HERE"
#femaleSpeaker6$ = "PUT-SPEAKER-CODE-HERE"
#femaleSpeaker7$ = "PUT-SPEAKER-CODE-HERE"

#-------------------------------------------ANALYSIS-----------------------------------------
```

```
filename$ = "Results_"+track$+""
outputPath$ = ""+path$+"\"+filename$+".csv"
writeFileLine: "'outputPath$'", "track;speaker;name; gender; word; start_time_word; end_time_word; lexical_set; min_lex_set; type; start_time_vowel; end_time_vowel; dur_vowel; time_onset; o_F1;o_F2;o_F3;time_mid;m_F1;m_F2;m_F3;time_glide;g_F1;g_F2;g_F3"

thisSound$ = selected$("Sound")
thisTextGrid$ = selected$("TextGrid")

select TextGrid 'thisTextGrid$'

#-------------------------------------------MALE ANALYSIS-----------------------------------
```

```
appendInfoLine: "working on the males..."

numberOfVowels = Get number of intervals: 3

select Sound 'thisSound$'
To Formant (burg)... 0 5 5000 0.025 50
select TextGrid 'thisTextGrid$'

for currentInterval from 1 to numberOfVowels
        vowel$ = Get label of interval: 3, currentInterval
                if vowel$ <> ""
                        vowel_start = Get start time of interval: 3, currentInterval
                        vowel_end = Get end time of interval: 3, currentInterval
                        duration = vowel_end - vowel_start
                        midpoint = vowel_start + duration/2
                        onset = vowel_start+(duration*0.3)
                        glide = vowel_end-(duration*0.3)

                        word = Get interval at time: 2, vowel_start
                        word_start = Get start time of interval: 2,word
                        word_end = Get end time of interval: 2,word
                        word$ = Get label of interval: 2, word

                        nameSpeaker = Get interval at time: 1, vowel_start
                        nameSpeaker$ = Get label of interval: 1, nameSpeaker

#####enter all male speakers in the parentheses (Format: "or nameSpeaker = maleSpeaker[NUMMER]$)
                if (nameSpeaker$ = maleSpeaker1$ or nameSpeaker$ = maleSpeaker2$)

                        select Formant 'thisSound$'

                        o_F1 = Get value at time... 1 onset Hertz Linear
                        o_F2 = Get value at time... 2 onset Hertz Linear
                        o_F3 = Get value at time... 3 onset Hertz Linear

                        m_F1 = Get value at time... 1 midpoint Hertz Linear
                        m_F2 = Get value at time... 2 midpoint Hertz Linear
```

```
                    m_F3 = Get value at time... 3 midpoint Hertz Linear

                    g_F1 = Get value at time... 1 glide Hertz Linear
                    g_F2 = Get value at time... 2 glide Hertz Linear
                    g_F3 = Get value at time... 3 glide Hertz Linear



                    appendFileLine:"'outputPath$'",
                                ...track$,";",
                                ...nameSpeaker$,";",
                                ...'"',";",
                                ..."m",";",
                                ...word$,";",
                                ...fixed$(word_start,3),";",
                                ...fixed$(word_end,3),";",
                                ...vowel$,";",
                                ...'"',";",
                                ...'"',";",
                                ...fixed$(vowel_start,3),";",
                                ...fixed$(vowel_end,3),";",
                                ...fixed$(duration,3),";",
                                ...fixed$(onset,3),";",
                                ...fixed$(o_F1,0),";",
                                ...fixed$(o_F2,0),";",
                                ...fixed$(o_F3,0),";",
                                ...fixed$(midpoint,3),";",
                                ...fixed$(m_F1,0),";",
                                ...fixed$(m_F2,0),";",
                                ...fixed$(m_F3,0),";",
                                ...fixed$(glide,3),";",
                                ...fixed$(g_F1,0),";",
                                ...fixed$(g_F2,0),";",
                                ...fixed$(g_F3,0)
                    endif
            endif
select TextGrid 'thisTextGrid$'
```

```
endfor

appendInfoLine: "ok, all done on the males..."

#------------------------------FEMALE ANALYSIS----------------------------------------

appendInfoLine: "working on the females now..."

select Sound 'thisSound$'
To Formant (burg)... 0 5 5500 0.025 50
select TextGrid 'thisTextGrid$'

for currentInterval from 1 to numberOfVowels
        vowel$ = Get label of interval: 3, currentInterval
                if vowel$ <> ""
                        vowel_start = Get start time of interval: 3, currentInterval
                        vowel_end = Get end time of interval: 3, currentInterval
                        duration = vowel_end - vowel_start
                        midpoint = vowel_start + duration/2
                        onset = vowel_start+(duration*0.3)
                        glide = vowel_end-(duration*0.3)

                        word = Get interval at time: 2, vowel_start
                        word_start = Get start time of interval: 2,word
                        word_end = Get end time of interval: 2,word
                        word$ = Get label of interval: 2, word

                        nameSpeaker = Get interval at time: 1, vowel_start
                        nameSpeaker$ = Get label of interval: 1, nameSpeaker

#####enter all female speakers in the parentheses (Format: "or nameSpeaker = femaleSpeaker[NUMMER]$)
                        if nameSpeaker$ = (femaleSpeaker1$)

                                select Formant 'thisSound$'

                                o_F1 = Get value at time... 1 onset Hertz Linear
                                o_F2 = Get value at time... 2 onset Hertz Linear
                                o_F3 = Get value at time... 3 onset Hertz Linear
```

```
                        m_F1 = Get value at time... 1 midpoint Hertz Linear
                        m_F2 = Get value at time... 2 midpoint Hertz Linear
                        m_F3 = Get value at time... 3 midpoint Hertz Linear

                        g_F1 = Get value at time... 1 glide Hertz Linear
                        g_F2 = Get value at time... 2 glide Hertz Linear
                        g_F3 = Get value at time... 3 glide Hertz Linear

                    appendFileLine:"'outputPath$'",
                            ...track$,";",
                            ...nameSpeaker$,";",
                            ...""",";",
                            ..."f",";",
                            ...word$,";",
                            ...fixed$(word_start,3),";",
                            ...fixed$(word_end,3),";",
                            ...vowel$,";",
                            ...""",";",
                            ...""",";",
                            ...fixed$(vowel_start,3),";",
                            ...fixed$(vowel_end,3),";",
                            ...fixed$(duration,3),";",
                            ...fixed$(onset,3),";",
                            ...fixed$(o_F1,0),";",
                            ...fixed$(o_F2,0),";",
                            ...fixed$(o_F3,0),";",
                            ...fixed$(midpoint,3),";",
                            ...fixed$(m_F1,0),";",
                            ...fixed$(m_F2,0),";",
                            ...fixed$(m_F3,0),";",
                            ...fixed$(glide,3),";",
                            ...fixed$(g_F1,0),";",
                            ...fixed$(g_F2,0),";",
                            ...fixed$(g_F3,0)
                        endif
                endif
select TextGrid 'thisTextGrid$'
endfor
appendInfoLine: "ok, all done!"
```

Appendix 4

R scripts

This appendix includes an overview of the R packages and the R scripts used to create the vowel plots. Four different types of scripts were used in this study: A script to plot the speaker-individual vowel plots, a script to plot the grouped textbook vowel plots, a script to plot the grouped vowel plot over all speakers of a sub-corpus, and a script to calculate the mean formant values and standard deviations per vowel per variety. For the vowel plot scripts, only the scripts for Received Pronunciation and Australian English are included, as the General American script is identical to the Received Pronunciation script. The script to calculate the mean formant values and SD is only included for Received Pronunciation, to exemplify the procedure.

R packages used in the scripts

These R packages are also cited in the References.

- dplyr (version 1.0.2; Wickham H et al., 2020)
- forcats (version 0.5.0; Wickham H, 2020)
- ggConvexHull (Version 0.1.0; Martin C, 2017)
- ggplot2 (version 3.3.2; Wickham H, 2016)
- ggrepel (version 0.8.2; Slowikowski K, 2020)
- purr (Version 0.3.4, Henry L. and Wickham H, 2020)
- readxl (version 1.3.1; Wickham H, Bryan J, 2019)
- stringr (version 1.4.0; Wickham H, 2019)
- tibble (Version 3.0.3; Müller K, Wickham H, 2020)
- tidyr (version 1.1.2; Wickham H, 2020)
- tidyverse (version 1.3.0; Wickham H et al., 2019)

Script for the speaker individual vowel plots

Received Pronunciation

Plotting Monophthongs for RP speakers

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots

library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the Received Pronunciation subcorpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\British_English\\Results_Excel\\1-Results_EGA2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Green Line
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\2-Results_GL2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Camden Town
vowels_CT <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\3-Results_CT2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

#one table with all data
vowels_all <- rbind(vowels_EGA, vowels_GL, vowels_CT) %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

#creating subset for monophthongs
my_monophthongs <- subset(vowels_all, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

Creating the necessary subsets using a loop

```
# Create a vector with the IDs of the speakers
speakers <- my_monophthongs$name
```

There are a total of 21 unique speakers in this data set. Here are two loops to create averaged subsets for each speaker.

```
# Create an empty list object to store each individual data frame
all_monophthongs <- list()

# Loop through the speakers vector to extract data on monophthongs
for(i in 1:length(speakers)){
  #creating subset for monophthongs
  df <- subset(my_monophthongs, name == speakers[i]) %>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)   %>%
  #create mean values monophthongs
  group_by(min_lex_set) %>%
  summarise(mean_F1 = mean(m_F1), mean_F2 = mean(m_F2))

  #save dataframe output in a list object
  all_monophthongs[[speakers[i]]] <- df
}
```

Plotting

This next chunk will create and save all plots.

```
make_plot <- function(speaker) {
  # Create the ggplot graph for current speaker
  p <- ggplot(all_monophthongs[[speaker]], aes(x = mean_F2, y = mean
_F1)) +
    geom_point(aes(x = mean_F2, y = mean_F1), alpha = 0.7) +
    geom_text_repel(data = all_monophthongs[[speaker]], aes(x = mean
_F2, y = mean_F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +

    scale_x_reverse() + scale_y_reverse() +
    labs (title = speaker, x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  # define the filename with the current speaker's name
  filename <- paste0("Plot", speaker, ".png")

  # Save the file
  ggsave(filename, plot = p,
         width = 12, height = 7,
         units = "in", dpi = 300)
}

lapply(speakers, make_plot)
```

Australian English

Plotting Monophthongs and Diphthongs for AusE speakers

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots

library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the Australian English subcorpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\Australian_English\\Results_Excel\\1-Results_Praat_EGA.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

# Data from Green Line
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\Australian_English\\Results_Excel\\2-Results_Praat_GreenLine.xlsx")
%>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

#one table with all data
vowels_all <- rbind(vowels_EGA, vowels_GL) %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

#creating subsets for diphthongs
my_diphthongs <- subset(vowels_all, type == "diphthong") %>%
  select (id, name, word, min_lex_set, type, o_F1, o_F2, g_F1, g_F2)

#creating subset for monophthongs
my_monophthongs <- subset(vowels_all, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

Creating all the necessary subsets using two loops

```
# Create a vector with the IDs of the speakers
speakers <- my_diphthongs$name
```

There are a total of 13 unique speakers in this dataset.

Here are two loops to create averaged subsets for each speaker.

```
# Create an empty list object to store each individual data frame
all_diphthongs <- list()

# Loop through the speakers vector to extract data on diphthongs
for(i in 1:length(speakers)){

  #creating subset diphthongs
  df <- subset(my_diphthongs, name == speakers[i]) %>%
    select (id, name, word, min_lex_set, type, o_F1, o_F2, g_F1, g_F
2) %>%

  #create mean values per diphthong per speaker
  group_by(min_lex_set) %>%
  summarise(o_F1 = mean(o_F1, na.rm=TRUE),
            o_F2 = mean(o_F2, na.rm=TRUE),
            g_F1 = mean(g_F1, na.rm=TRUE),
            g_F2 = mean(g_F2, na.rm=TRUE),
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.)) %>%

  #turn diphthongs into tall data
  gather("measurement_formant", "hz", o_F1, o_F2, g_F1, g_F2) %>%
  separate(measurement_formant, into = c("measurement", "formant"))
%>%
  spread(formant, hz) %>%
  arrange(id)

  #save dataframe output in a list object
  all_diphthongs[[speakers[i]]] <- df
}


# Create an empty list object to store each individual data frame
all_monophthongs <- list()

# Loop through the speakers vector to extract data on monophthongs
for(i in 1:length(speakers)){
  #creating subset for monophthongs
  df <- subset(my_monophthongs, name == speakers[i]) %>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)   %>%
  #create mean values monophthongs
```

```
  group_by(min_lex_set) %>%
  summarise(mean_F1 = mean(m_F1), mean_F2 = mean(m_F2))

  #save dataframe output in a list object
  all_monophthongs[[speakers[i]]] <- df

}
```

Plotting

This next chunk will create and save all plots.

```
make_plot <- function(speaker) {
  # Create the ggplot graph for current speaker
  p <- ggplot(all_monophthongs[[speaker]], aes(x = mean_F2, y = mean
_F1)) +
    geom_point(aes(x = mean_F2, y = mean_F1), alpha = 0.7) +
    geom_text_repel(data = all_monophthongs[[speaker]], aes(x = mean
_F2, y = mean_F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    geom_path(data = all_diphthongs[[speaker]], aes(x = F2, y = F1,
linetype = min_lex_set), size = 1, arrow = arrow(angle = 10, ends =
"first", type = "closed", length = unit(0.2, "inches")), show.legend
= FALSE) +
    geom_path(data = all_diphthongs[[speaker]], aes(x = F2, y = F1,l
inetype = min_lex_set), size = 1) +
    scale_linetype_manual(breaks=c("FACE", "PRICE", "MOUTH"), values
=c("solid", "dotted", "twodash")) +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = speaker, x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  # define the filename with the current speaker's name
  filename <- paste0("Plot", speaker, ".png")

  # Save the file
  ggsave(filename, plot = p,
         width = 12, height = 7,
         units = "in", dpi = 300)
}

lapply(speakers, make_plot)
```

Script for the grouped textbook vowel plots

Received Pronunciation

Plotting grouped mean value plots per textbook for the RP speakers

Lisa Scheiwe

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots

library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the General American subcorpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\British_English\\Results_Excel\\1-Results_EGA2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Green Line
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\2-Results_GL2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Camden Town
vowels_CT <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\3-Results_CT2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)


#creating subsets for monophthongs
my_monophthongs_EGA <- subset(vowels_EGA, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)

my_monophthongs_GL <- subset(vowels_GL, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

```
my_monophthongs_CT <- subset(vowels_CT, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

Calculating the mean values

This chunk calculates first the mean F1 and F2 values per vowel per speaker, and then the pooled mean value across all speakers per textbook.

```
#mean values monophthongs per speaker (TB EGA) rounded to one decima
l point
mean_monophthongs_EGA <- my_monophthongs_EGA %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

## `summarise()` regrouping output by 'name' (override with `.groups
` argument)
## `mutate_if()` ignored the following grouping variables:
## Column `name`

#mean values monophthongs across all speakers (TB EGA) rounded to on
e decimal point (the mean of the mean)
means_EGA <- mean_monophthongs_EGA %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

## `summarise()` ungrouping output (override with `.groups` argument
)

#mean values monophthongs per speaker (TB GL) rounded to one decimal
point
mean_monophthongs_GL <- my_monophthongs_GL %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

## `summarise()` regrouping output by 'name' (override with `.groups
` argument)
## `mutate_if()` ignored the following grouping variables:
## Column `name`

#mean values monophthongs across all speakers (TB GL) rounded to one
decimal point (the mean of the mean)
```

```
means_GL <- mean_monophthongs_GL %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

## `summarise()` ungrouping output (override with `.groups` argument
)

#mean values monophthongs per speaker (TB CT) rounded to one decimal
point
mean_monophthongs_CT <- my_monophthongs_CT %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

## `summarise()` regrouping output by 'name' (override with `.groups
` argument)
## `mutate_if()` ignored the following grouping variables:
## Column `name`

#mean values monophthongs across all speakers (TB CT) rounded to one
decimal point (the mean of the mean)
means_CT <- mean_monophthongs_CT %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

## `summarise()` ungrouping output (override with `.groups` argument
)
```

Plotting the grouped data for each textbook group

```
# Plot for EGA group
ggplot(means_EGA, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = means_EGA, aes(x = mean_m_F2, y = mean_m_
F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "EGA RP speakers", x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_EGA_speakers.png", plot = last_plot(), scale = 1, wi
```

```
dth = 12, height = 7, units = "in", dpi = 300)

  # Plot for GL group
ggplot(means_GL, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = means_GL, aes(x = mean_m_F2, y = mean_m_F
1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "GL RP speakers", x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_GL_speakers.png", plot = last_plot(), scale = 1, wid
th = 12, height = 7, units = "in", dpi = 300)

  # Plot for CT group
ggplot(means_CT, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = means_CT, aes(x = mean_m_F2, y = mean_m_F
1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "CT RP speakers", x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_CT_speakers.png", plot = last_plot(), scale = 1, wid
th = 12, height = 7, units = "in", dpi = 300)
```

Australian English

Plotting grouped mean value plots per textbook for the AusE speakers

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots
library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the Australian English sub-corpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\Australian_English\\Results_Excel\\1-Results_Praat_EGA.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

# Data from GL
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\Australian_English\\Results_Excel\\2-Results_Praat_GreenLine.xlsx")
%>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

#creating subsets for monophthongs
my_monophthongs_EGA <- subset(vowels_EGA, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
my_monophthongs_GL <- subset(vowels_GL, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
#creating subsets for diphthongs
my_diphthongs_EGA <- subset(vowels_EGA, type == "diphthong") %>%
  select (id, name, word, min_lex_set, type, o_F1, o_F2, g_F1, g_F2)
my_diphthongs_GL <- subset(vowels_GL, type == "diphthong") %>%
  select (id, name, word, min_lex_set, type, o_F1, o_F2, g_F1, g_F2)
```

Calculating the mean values

This chunk calculates first the mean F1 and F2 values per vowel per speaker, and then the pooled mean value across all speakers per textbook. This is done for the monophthongs and diphthongs separately

```
#mean values monophthongs per speaker (TB EGA) rounded to one decima
l point
mean_monophthongs_EGA <- my_monophthongs_EGA %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

#mean values monophthongs across all speakers (TB EGA) rounded to on
e decimal point (the mean of the mean)
means_EGA <- mean_monophthongs_EGA %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

#mean values monophthongs per speaker (TB GL) rounded to one decimal
point
mean_monophthongs_GL <- my_monophthongs_GL %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

#mean values monophthongs across all speakers (TB GL) rounded to one
decimal point (the mean of the mean)
means_GL <- mean_monophthongs_GL %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

#mean values diphthongs per speaker (TB EGA) rounded to one decimal
point
mean_diphthongs_EGA <- my_diphthongs_EGA %>%
  group_by(name, min_lex_set) %>%
    summarise(o_F1 = mean(o_F1, na.rm=TRUE),
              o_F2 = mean(o_F2, na.rm=TRUE),
              g_F1 = mean(g_F1, na.rm=TRUE),
              g_F2 = mean(g_F2, na.rm=TRUE),
    ) %>%
    mutate_if(is.numeric, round, digits = 1)

#mean values diphthongs across all speakers (TB EGA) rounded to one
decimal point (the mean of the mean)
means_diphthongs_EGA <- mean_diphthongs_EGA %>%
```

```
  group_by(min_lex_set) %>%
  summarise(o_F1 = mean(o_F1, na.rm=TRUE),
            o_F2 = mean(o_F2, na.rm=TRUE),
            g_F1 = mean(g_F1, na.rm=TRUE),
            g_F2 = mean(g_F2, na.rm=TRUE),
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

#turn diphthongs into tall data - for all speakers (TB EGA) as one g
roup
  my_diphthongs_tall_EGA  <- means_diphthongs_EGA %>%
        gather ("measurement_formant", "hz", o_F1, o_F2, g_F1, g_F2)
%>%
    separate (measurement_formant, into = c("measurement", "formant"
)) %>%
    spread (formant, hz) %>%
    arrange (id)

#mean values diphthongs per speaker (TB GL) rounded to one decimal p
oint
mean_diphthongs_GL <- my_diphthongs_GL %>%
  group_by(name, min_lex_set) %>%
    summarise(o_F1 = mean(o_F1, na.rm=TRUE),
              o_F2 = mean(o_F2, na.rm=TRUE),
              g_F1 = mean(g_F1, na.rm=TRUE),
              g_F2 = mean(g_F2, na.rm=TRUE),
    ) %>%
    mutate_if(is.numeric, round, digits = 1)

#mean values diphthongs across all speakers (TB GL) rounded to one d
ecimal point (the mean of the mean)
means_diphthongs_GL <- mean_diphthongs_GL %>%
  group_by(min_lex_set) %>%
  summarise(o_F1 = mean(o_F1, na.rm=TRUE),
            o_F2 = mean(o_F2, na.rm=TRUE),
            g_F1 = mean(g_F1, na.rm=TRUE),
            g_F2 = mean(g_F2, na.rm=TRUE),
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

#turn diphthongs into tall data - for all speakers (TB GL) as one gr
oup
  my_diphthongs_tall_GL  <- means_diphthongs_GL %>%
        gather ("measurement_formant", "hz", o_F1, o_F2, g_F1, g_F2)
%>%
    separate (measurement_formant, into = c("measurement", "formant"
)) %>%
    spread (formant, hz) %>%
    arrange (id)
```

Plotting the grouped data for each textbook group

```
# Plot for EGA group
ggplot(means_EGA, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = means_EGA, aes(x = mean_m_F2, y = mean_m_
F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    geom_path(data = my_diphthongs_tall_EGA, aes(x = F2, y = F1, lin
etype = min_lex_set), size = 1, arrow = arrow(angle = 10, ends = "fi
rst", type = "closed", length = unit(0.2, "inches")), show.legend =
FALSE) +
    geom_path(data = my_diphthongs_tall_EGA, aes(x = F2, y = F1,line
type = min_lex_set), size = 1) +
    scale_linetype_manual(breaks=c("FACE", "PRICE", "MOUTH"), values
=c("solid", "dotted", "twodash")) +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "EGA AusE speakers", x = "F2 (Hz)", y = "F1 (Hz)")
+
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_EGA_speakers.png", plot = last_plot(), scale = 1, wi
dth = 12, height = 7, units = "in", dpi = 300)

  # Plot for GL group
ggplot(means_GL, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = means_GL, aes(x = mean_m_F2, y = mean_m_F
1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    geom_path(data = my_diphthongs_tall_GL, aes(x = F2, y = F1, line
type = min_lex_set), size = 1, arrow = arrow(angle = 10, ends = "fir
st", type = "closed", length = unit(0.2, "inches")), show.legend = F
ALSE) +
    geom_path(data = my_diphthongs_tall_GL, aes(x = F2, y = F1,linet
ype = min_lex_set), size = 1) +
    scale_linetype_manual(breaks=c("FACE", "PRICE", "MOUTH"), values
=c("solid", "dotted", "twodash")) +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "GL AusE speakers", x = "F2 (Hz)", y = "F1 (Hz)")
+
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_GL_speakers.png", plot = last_plot(), scale = 1, wid
th = 12, height = 7, units = "in", dpi = 300)
```

Script for the grouped vowel plots across all speakers

Received Pronunciation

Plotting grouped mean value plots for the RP speakers

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots

library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the Received Pronunciation sub-corpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\British_English\\Results_Excel\\1-Results_EGA2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Green Line
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\2-Results_GL2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Camden Town
vowels_CT <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\British_English\\Results_Excel\\3-Results_CT2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

#one table with all data
vowels_all <- rbind(vowels_EGA, vowels_GL, vowels_CT) %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

#creating subset for monophthongs
my_monophthongs <- subset(vowels_all, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

Calculating the mean values

This chunk calculates first the mean F1 and F2 values per vowel per speaker, and then the pooled mean value across all speakers.

```
#mean values monophthongs per speaker rounded to one decimal point
mean_monophthongs <- my_monophthongs %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

#mean values monophthongs across all speakers rounded to one decimal
point (the mean of the mean)
mean_monophthongs_group <- mean_monophthongs %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))
```

Plotting the grouped data for all speakers of the sub-corpus

```
ggplot(mean_monophthongs_group, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = mean_monophthongs_group, aes(x = mean_m_F
2, y = mean_m_F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "All RP speakers", x = "F2 (Hz)", y = "F1 (Hz)") +
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_all_RP_speakers.png", plot = last_plot(), scale = 1,
width = 12, height = 7, units = "in", dpi = 300)
```

Australian English

Plotting grouped mean value plots for the AusE speakers

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse) #for the scatter plots

library(readxl)
library(ggrepel)
library(ggConvexHull) #to create the vowel spaces
```

Visualising the data from the Australian English sub-corpus

Loading and pre-processing the data

```
# Data from EGA
vowels_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks
\\Australian_English\\Results_Excel\\1-Results_Praat_EGA.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

# Data from GL
vowels_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbooks\
\Australian_English\\Results_Excel\\2-Results_Praat_GreenLine.xlsx")
%>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

#one table with all data
vowels_all <- rbind(vowels_EGA, vowels_GL) %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, o_F1, o_F2,
m_F1, m_F2, g_F1, g_F2)

#creating subsets for diphthongs
my_diphthongs <- subset(vowels_all, type == "diphthong") %>%
  select (id, name, word, min_lex_set, type, o_F1, o_F2, g_F1, g_F2)

#creating subset for monophthongs
my_monophthongs <- subset(vowels_all, type == "monophthong")%>%
  select (id, name, word, min_lex_set, type, m_F1, m_F2)
```

Calculating the mean values

This chunk calculates first the mean F1 and F2 values per vowel per speaker, and then the pooled mean value across all speakers. First for the monophthongs, then for the diphthongs.

```
#mean values monophthongs per speaker rounded to one decimal point
mean_monophthongs <- my_monophthongs %>%
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

#mean values monophthongs across all speakers rounded to one decimal
point (the mean of the mean)
mean_monophthongs_group <- mean_monophthongs %>%
  group_by(min_lex_set) %>%
  summarise(mean_m_F1 = mean(mean_m_F1, na.rm=TRUE),
            mean_m_F2 = mean(mean_m_F2, na.rm=TRUE)
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

## `summarise()` ungrouping output (override with `.groups` argument
)

#mean values diphthongs per speaker rounded to one decimal point
mean_diphthongs <- my_diphthongs %>%
  group_by(name, min_lex_set) %>%
    summarise(o_F1 = mean(o_F1, na.rm=TRUE),
              o_F2 = mean(o_F2, na.rm=TRUE),
              g_F1 = mean(g_F1, na.rm=TRUE),
              g_F2 = mean(g_F2, na.rm=TRUE),
    ) %>%
    mutate_if(is.numeric, round, digits = 1)

#mean values diphthongs across all speakers rounded to one decimal p
oint (the mean of the mean)
mean_diphthongs_group <- mean_diphthongs %>%
  group_by(min_lex_set) %>%
  summarise(o_F1 = mean(o_F1, na.rm=TRUE),
            o_F2 = mean(o_F2, na.rm=TRUE),
            g_F1 = mean(g_F1, na.rm=TRUE),
            g_F2 = mean(g_F2, na.rm=TRUE),
  ) %>%
  mutate_if(is.numeric, round, digits = 1)%>%
  mutate(id = 1:nrow(.))

## `summarise()` ungrouping output (override with `.groups` argument
)
```

```
#turn diphthongs into tall data - for all speakers as one group
  my_diphthongs_tall  <- mean_diphthongs_group %>%
        gather ("measurement_formant", "hz", o_F1, o_F2, g_F1, g_F2)
%>%
    separate (measurement_formant, into = c("measurement", "formant"
)) %>%
    spread (formant, hz) %>%
    arrange (id)
```

Plotting the grouped data for all speakers of the sub-corpus

```
ggplot(mean_monophthongs_group, aes(x = mean_m_F2, y = mean_m_F1)) +
    geom_point(aes(x = mean_m_F2, y = mean_m_F1), alpha = 0.7) +
    geom_text_repel(data = mean_monophthongs_group, aes(x = mean_m_F
2, y = mean_m_F1, label = min_lex_set)) +
    geom_convexhull(alpha = 0.3, fill = "grey") +
  geom_path(data = my_diphthongs_tall, aes(x = F2, y = F1, linetype
= min_lex_set), size = 1, arrow = arrow(angle = 10, ends = "first",

type = "closed",

length = unit(0.2, "inches")), show.legend = FALSE) +
    geom_path(data = my_diphthongs_tall, aes(x = F2, y = F1,linetype
= min_lex_set), size = 1) +
     scale_linetype_manual(breaks=c("FACE", "PRICE", "MOUTH"), value
s=c("solid", "dotted", "twodash")) +
    scale_x_reverse() + scale_y_reverse() +
    labs (title = "All AusE speakers", x = "F2 (Hz)", y = "F1 (Hz)")
+
    theme_classic() +
    theme(legend.position = "top", legend.title = element_blank())

  ggsave ("Plot_all_AusE_speakers.png", plot = last_plot(), scale =
1, width = 12, height = 7, units = "in", dpi = 300)
```

Script to calculate mean formant values and SD

Set-up

Built with R4.0.2

```
knitr::opts_chunk$set(echo = TRUE)

library(tidyverse)
library(readxl)
library(ggrepel)
```

Calculating the mean F1 and F2 values per speaker and the standard deviation

Loading and preprocessing the RP data

```
# Data from EGA
vowels_RP_EGA <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolbo
oks\\British_English\\Results_Excel\\1-Results_EGA2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Green Line
vowels_RP_GL <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolboo
ks\\British_English\\Results_Excel\\2-Results_GL2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

# Data from Camden Town
vowels_RP_CT <- read_excel("C:\\Users\\Admin\\sciebo\\PhD\\Schoolboo
ks\\British_English\\Results_Excel\\3-Results_CT2.xlsx") %>%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)

#one table with all data
vowels_all_RP <- rbind(vowels_RP_EGA, vowels_RP_GL, vowels_RP_CT) %>
%
  mutate(id = 1:nrow(.)) %>%
  select(id, name, word, lexical_set, min_lex_set, type, m_F1, m_F2)
```

Calculating the mean values in the individual sub-corpora

The mean F1 and F2 values are calculated per vowel per speaker and the results are rounded to one decimal point. The frequency per token per speaker is counted as well.

```
# Calculating the mean values for the RP sub-corpus
mean_vowels_RP <- vowels_all_RP %>%
```

```
  group_by(name, min_lex_set) %>%
  summarise(mean_m_F1 = mean(m_F1, na.rm=TRUE),
            mean_m_F2 = mean(m_F2, na.rm=TRUE)
            ) %>%
  mutate_if(is.numeric, round, digits = 1)

# count tokens per vowel and speaker in the RP sub-corpus

count_vowels_RP <- vowels_all_RP %>%
  group_by(name, min_lex_set) %>%
  summarise(n = n())

# calculating the SD per vowel and per speaker, rounded to one decim
al point

SD_vowels_RP <- vowels_all_RP %>%
  group_by(name, min_lex_set) %>%
  summarise(sd_m_F1 = sd(m_F1, na.rm=TRUE),
            sd_m_F2 = sd(m_F2, na.rm=TRUE)
               ) %>%
  mutate_if(is.numeric, round, digits = 1)
```

Exporting the results in text files

This next chunk of code creates three text documents per sub-corpus to save the mean values, the standard deviation, as well as the token numbers per speaker per vowel.

```
#export into txt - RP sub-corpus results

write.table(mean_vowels_RP, file = "mean_values_RP.txt", append = FA
LSE, sep = "\t ", dec = ".",
            row.names = TRUE, col.names = TRUE)

write.table(SD_vowels_RP, file = "SD_values_RP.txt", append = FALSE,
sep = "\t ", dec = ".",
            row.names = TRUE, col.names = TRUE)

write.table(count_vowels_RP, file = "count_values_RP.txt", append =
FALSE, sep = "\t ", dec = ".",
            row.names = TRUE, col.names = TRUE)
```

Appendix 5

Quantitative Data

This Appendix includes the mean values and standard deviations of the features of analysis per speaker.

Mean values and standard deviations for the speakers from English G Access 2

Speaker	GOOSE				LOT				CLOTH				TRAP				BATH				START			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
F15E2	F1	513.5	112.8	4	F1	733	116	8	F1	745	9.5	3	F1	1131.3	142.2	14	F1	1140	N/A	1	F1	968	215	2
	F2	2287	461.1		F2	1449.8	219.4		F2	1521	123.4		F2	1815.1	101.7		F2	1412	N/A		F2	1353.5	29	
F16E2	F1	473.8	109.3	4	F1	764.5	78.5	2	F1	N/A	N/A	0	F1	971	118.1	4	F1	976	87.7	2	F1	N/A	N/A	0
	F2	2226.5	651.8		F2	1151.5	96.9		F2	N/A	N/A		F2	1625.5	130.9		F2	1407.5	65.8		F2	N/A	N/A	
M20E2	F1	433.4	107.9	9	F1	648.7	143.4	11	F1	900.7	159.1	3	F1	987.6	156.9	5	F1	894.2	31	4	F1	816.3	97.9	3
	F2	1784.9	492.1		F2	1175.2	168.8		F2	1369	137.2		F2	1564.8	389.8		F2	1529.8	280.1		F2	1206.7	102.8	
M21E2	F1	394.2	68.9	4	F1	724.2	33.5	4	F1	687.7	34.4	7	F1	913.6	117.3	9	F1	951.5	132.2	2	F1	747.4	68.6	14
	F2	1299.5	171.7		F2	1015.8	95.1		F2	1047.4	53.7		F2	1470.4	38		F2	1380.5	101.1		F2	1112.3	53.2	
M22E2	F1	388.8	55.8	11	F1	647.6	56.1	10	F1	678	58.5	4	F1	884.3	183.6	26	F1	840.1	59.2	8	F1	787.6	97.2	5
	F2	1567.5	207.2		F2	1177.4	104.2		F2	1198.5	83.2		F2	1594	81.6		F2	1289	78.9		F2	1237.6	86.1	
M23E2	F1	315.6	38.5	12	F1	608.5	219.1	11	F1	518.3	68.7	3	F1	956.4	234.4	10	F1	809.4	229.7	15	F1	814.3	209.2	6
	F2	1480.8	375.1		F2	1110.6	234.9		F2	1052.7	142		F2	1532.8	98.7		F2	1154.5	96.6		F2	1120.3	38.6	
M24E2	F1	362.9	52.9	16	F1	608.5	88.4	11	F1	583	75	2	F1	990.4	148.1	31	F1	825.1	111.7	19	F1	841.7	113.8	22
	F2	1374	347		F2	1100.9	104.8		F2	1220.5	33.2		F2	1626.4	115.8		F2	1206.7	43.5		F2	1207.7	69.8	
F17G2	F1	342.4	39.5	5	F1	658.9	153.5	13	F1	605.8	79.3	5	F1	1050.7	206.2	9	F1	834	NA	1	F1	771	NA	1
	F2	2188.6	227.1		F2	1196.3	143.7		F2	1226.8	76		F2	1659.6	283.3		F2	1289	NA		F2	1260	NA	

Mean values and standard deviations for the speakers from Green Line 2

Speaker	GOOSE				LOT				CLOTH				TRAP				BATH				START			
		Mean				Mean				Mean				Mean				Mean				Mean		
		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*
F18G2	F1	406.5	24.7	2	F1	757.3	99.3	6	F1	722	N/A	1	F1	943.3	17.2	3	F1	894.5	74.2	2	F1	864.5	48.8	2
	F2	1481	752.4		F2	1145.7	112		F2	1064	N/A		F2	1578.3	366.2		F2	1345.5	109.6		F2	1348	50.9	
F19G2	F1	456.8	65.4	5	F1	721.2	122.6	8	F1	N/A	N/A	0	F1	1021.4	176.2	5	F1	983.7	19.1	3	F1	N/A	N/A	0
	F2	2153	439.2		F2	1216	158.9		F2	N/A	N/A		F2	1674.6	440.1		F2	1439	47.1		F2	N/A	N/A	
M25G2	F1	335.4	40.7	5	F1	351.5	177.5	4	F1	605.3	368.4	3	F1	823.6	255.8	11	F1	551.5	27.6	2	F1	330.5	19.1	2
	F2	1211.2	98.7		F2	952	88		F2	1435.7	824.8		F2	1819.7	145.8		F2	1127.5	9.2		F2	1135.5	37.5	
M26G2	F1	372.1	31.5	10	F1	606.8	142.1	10	F1	668	NA	1	F1	845	92.4	8	F1	728.7	21.1	3	F1	715	29.1	3
	F2	1763.5	451.1		F2	1163.6	391.8		F2	1169	NA		F2	1443.2	86		F2	1088.3	41.6		F2	1095.7	71.2	
M27G2	F1	355.6	60.8	13	F1	441.2	68.1	5	F1	521	72.7	3	F1	920.3	101.5	13	F1	627.4	176.6	9	F1	721.5	26.2	2
	F2	1764.4	364.9		F2	1063.8	90.6		F2	1068	48.5		F2	1587	132.6		F2	1180.6	181.9		F2	1166.5	19.1	
M28G2	F1	332	54.6	18	F1	616.8	49.7	22	F1	629.8	59.3	5	F1	858.1	92.4	18	F1	704.2	39.6	12	F1	685.4	71.7	10
	F2	2042.8	466.2		F2	1104.4	85		F2	1156.2	82.5		F2	1526.3	109.1		F2	1147.2	55.7		F2	1207.7	164.9	

Mean values and standard deviations for the speakers from Camden Town 2

Speaker		GOOSE				LOT				CLOTH				TRAP				BATH				START		
		Mean				Mean				Mean				Mean				Mean				Mean		
		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*
F20C2	F1	413	72.1	2	F1	628	N/A	1	F1	672	66.1	3	F1	1110.5	4.9	2	F1	N/A	N/A	0	F1	1051	19.8	2
	F2	2621.5	55.9		F2	985	N/A		F2	1169	100.7		F2	1669	33.9		F2	N/A	N/A		F2	1376.5	12	
F21C2	F1	391.3	15.8	3	F1	703.5	136.3	10	F1	789	N/A	1	F1	920.5	103.9	2	F1	899.4	79.3	5	F1	990.3	58.7	3
	F2	1667.7	513.5		F2	1366.1	152.7		F2	1099	N/A		F2	1741	12.7		F2	1361.6	57.2		F2	1395.7	16.1	
F22C2	F1	402.7	77.9	6	F1	858.9	206.1	13	F1	914.7	77.2	3	F1	1116	52.9	5	F1	1098.8	38.5	4	F1	1012.3	202.4	7
	F2	1770.8	738		F2	1420.3	201.9		F2	1309.7	82		F2	1764.2	166.2		F2	1656	39.3		F2	1559.6	180.5	
F23C2	F1	384	60.4	4	F1	800.8	171.6	9	F1	837	7.1	2	F1	1241	143.9	5	F1	972.3	58.9	3	F1	749	NA	1
	F2	1581	453.4		F2	1201.8	185.3		F2	1156	15.6		F2	1740.2	107.7		F2	1538	95.2		F2	1993	NA	
F24C2	F1	555.2	464.5	18	F1	848.1	144.1	18	F1	889.7	45.5	6	F1	1103.4	72.8	20	F1	1017.2	101.6	8	F1	1015	51.9	4
	F2	1902.3	647.1		F2	1411.7	538.5		F2	1352.7	176.4		F2	1719.6	134.3		F2	1513.6	161.6		F2	1428	57.1	
M29C2	F1	406.8	34.7	6	F1	595.8	93.8	5	F1	749	N/A	1	F1	1056.3	168	10	F1	1098	N/A	1	F1	N/A	N/A	0
	F2	1804.3	502.8		F2	1233	196.6		F2	1233	N/A		F2	1754	232.5		F2	1877	N/A		F2	N/A	N/A	
M30C2	F1	351.1	84.5	9	F1	610.1	92.3	13	F1	847.5	188.8	2	F1	890.8	96.2	20	F1	746.3	17.2	3	F1	743	45.3	2
	F2	1901.6	299.6		F2	1132.2	84.3		F2	1604	767.9		F2	1476.8	77.9		F2	1201	22.6		F2	1100.5	43.1	

General American

English G Access 4

Speaker		BATH				LOT				CLOTH			
		Mean				Mean				Mean			
		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*	
F5E4	F1	1036.9	102.5	7	F1	861.2	170	11		683	283.1	1	
	F2	1761.1	136.6		F2	1299.6	133		F2	1166.5	89		
F6E4	F1	1022.9	98.3	11	F1	914.5	129	24	F1	1002	45.4	2	
	F2	1909.5	171.1		F2	1410.5	165.1		F2	1380	172.3		
F7E4	F1	1001.6	202.4	5	F1	871.2	160.7	20	F1	782	234	12	
	F2	1507.4	208.7		F2	1543.6	227.4		F2	1200.8	183.5		
M10E4	F1	720.6	188	5	F1	663.5	124.5	21	F1	521	N/A	1	
	F2	1624.2	280.1		F2	1093	222.9		F2	886	N/A		
M11E4	F1	723.4	233.5	16	F1	808.9	220.3	21	F1	790.3	345	3	
	F2	1847.3	440.2		F2	1248.4	172		F2	1372.7	414.2		
M12E4	F1	1005.8	41	4	F1	730.7	161.6	20	F1	585.4	52.6	5	
	F2	1713.8	30.4		F2	1152.4	91.4		F2	1012.4	54.9		
M13E4	F1	731.4	167.6	7	F1	850.8	221.6	10	F1	615.4	154	5	
	F2	1528.6	67		F2	1148.4	130.3		F2	1045	93		

Green Line 4

Speaker	BATH				LOT				CLOTH			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
F8G4	F1	917	379	2	F1	850.3	166.9	10	F1	817.5	229.8	4
	F2	1984.5	19.1		F2	1527.7	126.7		F2	1519	148.5	
F9G4	F1	1043	N/A	1	F1	1017.1	96	10	F1	878	N/A	4
	F2	1776	N/A		F2	1532.3	194.1		F2	1403	N/A	
F10G4	F1	959.5	36.1	2	F1	990.1	114.2	9	F1	470	N/A	4
	F2	1764.5	269.4		F2	1377.1	182.6		F2	1054	N/A	
M14G4	F1	796.9	103.3	13	F1	787	124.5	48	F1	728.7	149.2	6
	F2	1986.7	358.6		F2	1315.4	142.4		F2	1157	122.3	
M15G4	F1	879	226.4	4	F1	720.8	49.2	12	F1	698.6	38.3	7
	F2	1487.8	98.6		F2	1045.4	77.1		F2	976.3	105.3	
M16G4	F1	N/A	N/A	0	F1	709.3	44.7	6	F1	688	N/A	1
	F2	N/A	N/A		F2	1270.2	60		F2	1317	N/A	
M17G4	F1	552.5	56	4	F1	672.8	104	15	F1	N/A	N/A	0
	F2	1911.8	188.7		F2	1268.3	108.5		F2	N/A	N/A	

Camden Town 4

Speaker		BATH				LOT				CLOTH		
		Mean				Mean				Mean		
		Hz	SD	*n*		Hz	SD	*n*		Hz	SD	*n*
F11C4	F1	1029.8	211.7	4	F1	1024.7	111.3	15	F1	813	274.4	6
	F2	2032.2	241.9		F2	1359	77.8		F2	1095.5	208.6	
F12C4	F1	1027.5	6.4	2	F1	950.1	141.7	11	F1	763.1	191.9	2
	F2	1810.5	118.1		F2	1536.5	163.6		F2	1142.9	199	
F13C4	F1	968.7	45.5	3	F1	847.1	85.9	9	F1	816.2	49.1	1
	F2	1891.7	103		F2	1325.8	81.6		F2	1287.5	105.7	
F14C4	F1	N/A	N/A	0	F1	874	107.6	8	F1	N/A	N/A	0
	F2	N/A	N/A		F2	1614.6	305.5		F2	N/A	N/A	
M18C4	F1	850.8	34.3	6	F1	801	67.3	24	F1	816.3	45.7	7
	F2	1611.3	138.4		F2	1235.5	113.3		F2	1182.1	71.2	
M19C4	F1	778.7	158.5	6	F1	774.1	92.3	20	F1	722.6	96.2	5
	F2	1753.5	228		F2	1156.8	95.1		F2	1097.2	76.5	

Australian English

English G Access 5 and Green Line 5

Speaker	FACE				PRICE				MOUTH			
		Mean Hz	SD	*n*		Mean Hz	SD	*n*		Mean Hz	SD	*n*
F1E5	F1	980.6	162.2	8	F1	866.4	93.9	7	F1	897.8	124.1	4
	F2	1677.8	142.3		F2	1375.9	82.8		F2	2123.8	76.4	
M1E5	F1	774.3	78.3	50	F1	749.9	83.9	36	F1	818.7	65.5	27
	F2	1442.2	146.9		F2	1082.1	111.1		F2	1698.6	110.0	
M2E5	F1	703.8	48.6	35	F1	652.0	58.8	24	F1	718.1	40.9	20
	F2	1362.9	107.8		F2	1087.5	77.5		F2	1560.6	134.5	
M3E5	F1	697.5	45.8	8	F1	611.1	110.1	10	F1	679.0	N/A	1
	F2	1345.0	65.6		F2	987.2	121.7		F2	1592.0	N/A	
M4E5	F1	731.0	70.2	30	F1	760.0	43.8	20	F1	811.1	34.9	19
	F2	1497.7	206.2		F2	1078.0	68.4		F2	1501.4	137.6	
M5E5	F1	679.7	66.9	24	F1	664.3	79.8	47	F1	649.1	98.4	19
	F2	1491.2	131.9		F2	1126.7	103.9		F2	1544.0	104.2	
F2G5	F1	838.8	110.7	9	F1	873.2	135.0	27	F1	869.5	128.7	4
	F2	1816.4	196.0		F2	1312.5	171.3		F2	2059.8	96.1	
F3G5	F1	834.2	117.3	30	F1	969.8	137.2	20	F1	934.6	80.4	9
	F2	1780.5	240.4		F2	1353.8	124.7		F2	1781.9	313.4	
F4G5	F1	860.9	166.2	18	F1	999.0	116.8	24	F1	916.6	181.5	11
	F2	2059.8	314.4		F2	1398.2	144.9		F2	1687.4	346.2	
M6G5	F1	777.7	98.4	15	F1	790.1	84.4	27	F1	771.1	124.9	15
	F2	1632.6	146.9		F2	1079.3	57.4		F2	1809.9	114.4	
M7G5	F1	621.3	128.8	18	F1	856.1	133.9	37	F1	864.3	77.6	7
	F2	1777.3	148.5		F2	1211.3	87.7		F2	1696.7	90.7	
M8G5	F1	724.8	78.0	18	F1	748.1	121.7	18	F1	612.7	102.0	9
	F2	1483.3	110.7		F2	1084.7	93.8		F2	1670.3	107.0	
M9G5	F1	694.7	97.8	17	F1	738.6	80.1	10	F1	613.3	52.4	6
	F2	1528.6	121.2		F2	1197.7	282.9		F2	1668.0	183.9	

Appendix 6

Quantitative Data of the additional monophthongs

This Appendix includes the token numbers for the monophthongs additionally analysed for the Received Pronunciation and General American sub-corpora.

Token numbers for the additional monophthongs per speaker in the Received Pronunciation sub-corpus

Speaker	FLEECE	KIT	DRESS	PALM	NURSE	STRUT	THOUGHT	NORTH	FOOT
F15E2	5	5	5	4	1	5	2	1	2
F16E2	5	5	5	0	0	1	1	5	2
M20E2	5	5	5	1	1	5	1	3	0
M21E2	5	5	5	0	0	5	5	2	5
M22E2	5	4	5	4	5	4	4	4	0
M23E2	5	5	5	0	2	5	5	5	5
M24E2	5	5	5	0	3	5	5	5	3
F17G2	5	5	5	0	1	5	5	5	5
F18G2	2	5	5	0	0	7	3	2	1
F19G2	3	5	5	0	3	5	0	0	1
M25G2	5	5	5	1	4	9	4	0	5
M26G2	5	5	5	0	5	3	5	4	2
M27G2	5	5	5	0	5	5	4	5	5
M28G2	5	5	5	0	4	5	3	5	3
F20C2	5	4	5	1	1	4	2	1	3
F21C2	5	5	5	0	1	5	2	5	3
F22C2	5	5	5	1	4	4	2	4	3
F23C2	5	5	5	0	1	4	0	3	1
F24C2	5	5	5	0	5	5	5	5	5
M29C2	5	5	5	0	3	5	5	3	2
M30C2	5	5	5	0	5	7	3	5	2
Total	100	103	105	12	54	103	66	72	58

Table 1. Token numbers per speaker for the additional monophthongs in the Received Pronunciation sub-corpus.

Token numbers for the additional monophthongs per speaker in the General American sub-corpus

Speaker	FLEECE	KIT	DRESS	TRAP	PALM	START	NURSE	STRUT	THOUGHT	NORTH	FORCE	FOOT	GOOSE
F5E4	5	5	5	5	3	5	1	4	3	1	1	4	5
F6E4	5	5	5	5	1	5	5	6	5	1	5	4	5
F7E4	5	5	5	5	0	2	1	5	3	1	1	0	5
M10E4	5	5	5	6	0	1	2	5	5	2	2	5	5
M11E4	5	5	4	5	4	2	4	5	5	0	5	5	5
M12E4	5	5	5	5	0	5	5	5	5	0	3	5	4
M13E4	3	7	5	5	0	5	5	5	5	1	2	4	5
F10G4	5	5	5	5	0	3	2	5	4	2	1	1	4
F8G4	3	0	3	3	0	2	3	2	2	1	1	4	4
F9G4	5	5	5	5	1	6	0	5	5	1	3	4	5
M14G4	5	5	5	5	1	5	4	6	5	5	5	5	5
M15G4	5	5	5	5	0	5	5	5	0	5	5	5	5
M16G4	3	5	5	2	0	0	0	5	2	2	0	1	5
M17G4	5	5	5	5	0	2	3	5	5	5	1	2	5
F11C4	5	5	5	5	0	3	5	5	5	4	4	5	5
F12C4	5	5	5	5	0	5	5	5	1	6	5	5	3
F13C4	5	5	5	5	0	5	5	5	5	5	5	2	5
F14C4	5	5	5	5	0	4	1	5	2	0	2	0	5
M18C4	6	5	5	5	5	5	5	5	4	2	5	2	5
M19C4	5	5	5	4	0	5	5	5	5	5	4	5	5
Total	**95**	**97**	**97**	**95**	**15**	**75**	**66**	**98**	**76**	**49**	**60**	**68**	**95**

Table 2. Token numbers per speaker for the additional monophthongs in the General American sub-corpus.

Appendix 7

Digital Appendix

The Digital Appendix includes an annotated Praat file and the Excel files with the results of the acoustic analyses.

Overview of files included:

Annotated Praat file: EGA5-5

Excel files:

Received Pronunciation

Results from English G Access 2

Results from Green Line 2

Results from Camden Town 2

General American

Results from English G Access 4

Results from Green Line 4

Results from Camden Town 4

Australian English

Results from English G Access 5

Results from Green Line 5