An empirical investigation of pronunciation problems of young learners of English as a foreign language - identification and remedial strategies

Anne Bettina von Rekowski

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Human Communication Sciences
University of Sheffield

August 2020

In loving memory of my mum

ACKNOWLEDGEMENTS

First and foremost, I want to express my deepest gratitude to Professor Dr Sandra Whiteside who not only introduced me to the particularities of acoustic phonetics and second language acquisition but whose kind and warm-hearted personality has supported me and my research from day one. I hope you know that if it wasn't for you, I would have not continued down this very long and rocky road. Thank you!

I also want to thank Dr Stuart Cunningham for his kindness, his patient IT advice and for accompanying my first programming steps.

I was very fortunate to meet Lisa Clarkson, Dr Silke Fricke, Dr Frank Herrmann and Dr Blanca Schäfer who have not just literally given me a home in Sheffield but who have supported me not only contentwise but particularly emotionally. I am very grateful to you for your encouragement and your friendship.

I would also like to thank Dr Jenny Thompson and Dr Anna Maria di Betta for providing a fair viva and supporting my PhD project with their comments.

I am also grateful to the TEFL section of the English department at the University of Giessen. In particular, I would like to thank Prof Dr Eva Burwitz-Melzer for her support and guidance and from whom I was able to learn a lot.

However, first and foremost, I owe a very special thanks to Professor Dr Jürgen Quetz who has always believed that this book will one day see the light of day. You did not only find the time to listen to my early ideas, give me honest feedback and provide useful suggestions, you also encouraged me and fully supported this project over the years despite various difficulties and struggles. I owe a huge dept to you and I will always be grateful for your help.

I also which to thank the participants of the TechAG at the Universty of Giessen for providing a very motivating and comfortable discussion forum and very pleasant research dinner meetings.

Furthermore, I would like to thank all of the participating students and teachers who devoted so much time and energy to support my research.

I cannot forget to thank my friends and family who have accompanied me through the years of this intense project. Particularly, I want to thank Dr Silvia Adelhelm, Prof Dr

Anette Fox-Boyer, Edda Karpenstein, Dr Natascha Keunecke, Anne Kraft, Dr Svenja Lauenroth and Dr Katrin Thomas. Thank you very much for being in my life.

Finally, I want to express my deepest gratitude to my father who has always shown a great interest in my studies and especially in this project. Thank you for your love, your constant support and encouragement.

To conclude, I would like to thank my beloved husband Steven and my son Philipp as well as Bruno for being my family.

ABSTRACT

This thesis describes a novel pronunciation training programme devised for young learners of English as a foreign language in schools in Germany. The intervention is grounded in current theories of second language, employs valid and established pronunciation task formats, and targets a core set of significant pronunciation problem areas typical of young learners. This core set was identified from research literature, curricular requirements as well as data from questionnaires completed by 245 English language teachers in the German state of Hesse.

The English pronunciation intervention was implemented over five months within the language education setting of English L2 learners (ages 10 to 12) in two schools in Frankfurt/Main. Quasi-experimental classroom research was conducted using a control group design to examine the effects of the pronunciation intervention. To obtain within-subject measurements, the data were collected at three points: prior to the intervention (pre-test), immediately after the intervention (post-test) and six months later (follow-up). There were two types of analyses of the pronunciation data. First there was auditory evaluation: all test utterances by the students were assessed auditorily using a three-point scale to rate their performance. Secondly, acoustic analysis of the same data was conducted using the computer software programme Praat. The acoustic parameters of interest included fundamental and formant frequencies, vowel and fricative durations, spectral peaks and spectral rate of change, to assess changes in pronunciation. The main aim of the analyses was to measure the impacts, if any, of the pronunciation intervention.

The results revealed that maturation processes took place for both the intervention and control groups, as predicted. Overall, the auditory analysis presented mixed results but showed some training effects for the fricatives /ð/ and /θ/ and the affricate /dʒ/. Most salient of all, there were clear frequency effects. The acoustic data analysis indicated more systematic effects of the intervention on acoustic parameters associated with precision and stability. The scatter plots and Euclidean distances computed from the acoustic data revealed more precision and stability in the production of the vowels /ɑː, ɔː, ə, æ/, the diphthongs /ɪə; eɪ/ and the approximant /w/ in the intervention group. In addition, the centre of gravity values indicated that the training led to a more native-like production of /θ/ and /ð/.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

0. Motivation and rationale for the study

I am a German non-native English language teacher with a Master's degree[1] in English language teaching and a CELTA[2] qualification. I have worked with children and adults from many different cultural and language backgrounds. My work with students at beginner level in particular led me to realise that the language learning process really began as soon as the learners were able to segment the stream of language into smaller components of meaning. Moreover, for learners it was not just their own understanding of the new language which was pivotal: it was being understood by others that marked the really crucial point where learners fully became part of the communicative process. In my experience, this process largely centres around decoding and producing speech with sufficiently accurate English language pronunciation. From a teacher's point of view, I was interested in finding out to what extent I could facilitate this process – indeed, whether as a teacher I could influence it at all. For this reason, I wanted to explore what makes good, effective pronunciation training, and whether it is possible to integrate such training in the everyday lessons of a typical foreign language classroom. Due to there being limited appropriate resources in pronunciation work at the time of the study's inception, I decided to do further research into this topic in order to design and develop a pronunciation training programme and to systematically investigate its effectiveness.

[1] 1. und 2. Staatsexamen
[2] Cambridge Certificate in Teaching English to Speakers of other Languages

1. Introduction

Pronunciation[3] plays a very important role in communicative contexts and it is the first language feature to be noticed. In addition, the perceived strength of the 'foreign accent'[4] correlates significantly with listeners' perceptions of overall oral proficiency as the awareness of errors in grammar or vocabulary is largely influenced by the quality of pronunciation (Götz, 2011; Herbst, 1992). However, even proficient foreign language (FL)[5] learners with years of language input can retain noticeable foreign accents which for some speakers (but certainly not all) compromise their intelligibility. The issues surrounding the difficulty of learning foreign language pronunciation, especially in comparison to other aspects of language, has not yet been fully understood.

The significance of pronunciation teaching in the foreign language classroom in Germany has changed considerably over the last few decades. Until the 1970s, pronunciation was a major focus of foreign language teaching and it was assumed that language learners would only be able to perceive and produce phonetic distinctions and master intonation patterns with the help of explicit instruction (Moyer, 2013). Thus, language labs were frequently used for pronunciation teaching, and exercises were often repetitive (e.g. drills) and without much connection to what is known about the English language learning[6] process in the context of authentic communication. The development of the communicative approaches from the mid-1970s onwards was accompanied by a shift in teaching objectives (cf. Piepho, 1974). Pronunciation teaching became less important (Grotjahn, 1998; Mehlhorn, 2005; Piske, MacKay, & Flege, 2001a) as phonological fluency was no longer treated as a discrete topic but only as "a means to negotiate meaning in discourse" (Dalton & Seidlhofer, 1994). Within this pedagogical trend, the definition of `good pronunciation´ changed from a clear native-speaker model to a more functional approach – that is, for a learner to be able to communicate successfully and intelligibly (Jenkins, 2000; Neri, Cucchiarini, & Strik, 2006). Although there is a consensus that intelligibility and comprehensibility should be the primary goals of second language (L2)[2] pronunciation (Hessisches

[3] In this thesis, pronunciation describes the articulation of specific speech sounds.

[4] Although there is no universally accepted definition of a `foreign accent´, it is widely accepted that the term refers to the deviation from native to non-native pronunciation (Scovel, 2000).

[5] The terms foreign language (FL) and second language (L2) are used to refer to any language used other than the mother tongue(s) (L1) (Edmondson, 1999).

[6] The terms `language acquisition´ and `language learning´ will be treated as synonymous and used interchangeably throughout this study for reasons explained by recent neurolinguistic findings (cf. Abutalebi 2008).

Kultusministerium, 2011; Jenkins, 2000; Neri et al., 2006), the terms do not have widely accepted definitions[7] and, moreover, there is no agreement on how to teach and measure intelligibility (Isaacs, 2008).

Within a communicative approach, students will probably be able to acquire reasonable pronunciation skills (Harmer, 2007). Yet, learning pronunciation is particularly difficult because it requires the learner to develop new perceptual abilities in order to identify the new second language (L2) patterns and to learn new articulatory movements which need to be automated. Moreover, learners typically have to overcome inhibitions and might take on new language identities (Hirschfeld, 1997, p. 69). Despite the fact that foreign language pedagogy postulates that pronunciation is learned without any explicit instruction, second language research has shown that overt pronunciation teaching can strongly increase pronunciation accuracy (Bongaerts, Summeren, Planken, & Schils, 1997; Derwing & Rossiter, 2003; Missaglia & Sendelmeier, 1999; Moyer, 1999). This dispute regarding the importance of pronunciation instruction might indicate that there is not yet a sufficient exchange of ideas between second language (SL) scholars and foreign language pedagogy experts and practitioners.

In the past few years there has been a renewed interest in pronunciation, probably due to the integration of global markets. This is reflected in a huge and burgeoning demand for classroom-based and online pronunciation courses (Moyer, 2013) and especially in the growing number of publications in this area. With communicative competence continuing to be one of the main goals of foreign language teaching, the focus of pronunciation teaching has now shifted more towards the suprasegmental aspects of connected speech (e.g. sentence stress, rhythm and intonation) (Hedge, 2000). Due to the rise of new media, ‘new’ ways to teach pronunciation are on offer. However, the substance of these programmes does not seem to have changed a lot and little is known about their effectiveness (Moyer, 2013).

A foreign accent does not simply mean that a language learner is unable to produce a particular sound or intonation appropriately, but also that he or she might not be able to discriminate or identify aspects of the language input correctly. Several

[7] In this thesis, the term `intelligibility´ is used to describe the formal recognition of the decoding of words and utterances (see Pickering, 2012), and `comprehensibility´ refers to the listener’s ability to understand a word or utterance in a given context.

factors may contribute to flawed discrimination and pronunciation, including the differences in phonological structure between the foreign language and mother tongue of the learner (Trubetzkoy, 1969), restricting aspects in the process of language acquisition, motor-neuronal developments, and individual factors such as aptitude or age amongst others. In order to evaluate pronunciation teaching, these factors and processes underlying pronunciation acquisition must be taken into account.

Although there is a considerable volume of pronunciation research within the area of applied linguistics, most of the studies have focused on single aspects of pronunciation and the investigations were carried out in laboratory settings with only small numbers of adult participants (Archibald & Young-Scholten, 2003; Bohn & Flege, 1992; Gut, 2009; Iverson et al., 2003; Neri et al., 2006; Strange et al., 1998). Therefore, the findings are not easily transferred to a classroom setting, particularly with children. Also, as indicated above, the lack of cross-fertilisation between disciplines has meant that few of the applied linguistics studies have impacted on foreign language pedagogy.

The matters discussed in this section convey the importance of including linguistics research in foreign language teaching and the necessity of classroom-based pronunciation research. For these reasons, this thesis seeks to fill the gaps by pursuing two primary aims: first, to design and develop a pronunciation training programme, firmly grounded in second language learning theory, which can be easily integrated into English pronunciation teaching in German schools; second, to determine whether and to what extent the pronunciation intervention can improve second language pronunciation.

Within the scope of these aims, the research project was carried out in eight fifth-grade classes at a comprehensive school and a grammar school[8] in Frankfurt am Main in the state of Hesse. The pupils studied were aged 10 to 12 years and were learning English at beginner level. Therefore, from the outset, the main limitation of this thesis is that the findings cannot be generalised beyond this age and type of learner and this educational and geographical setting. Moreover, this research project focuses only on selected segmental pronunciation areas while excluding other segments and supra-segmental features. This decision does not reflect different levels of importance, only that seeking to encompass all segments and features would be beyond the scope of

[8] See Section 2.2 for an explanation of the terms ‘comprehensive’ and ‘grammar’ schools in the German system.

one PhD project. The overall structure of the thesis takes the form of nine chapters, including this introductory section.

Chapter 2: Setting the scene: English pronunciation teaching in the foreign language classroom in Germany

Chapter 2 lays the groundwork for this study and describes the status quo of English pronunciation teaching in the foreign language classroom in Germany and in particular in Hesse. Hence, this chapter starts with a discussion of pronunciation models (see Section 2.1. and examines the Hessian curricula (see Section 2.2.) to summarise the guidelines and subject matter of pronunciation teaching. As teachers typically use textbooks as their main frame of reference, the specified English textbooks are analysed with regard to English pronunciation work (see Section 2.3.). To find out how teachers actually implement the guidelines, use their textbooks and carry out pronunciation teaching in English foreign language classrooms in Germany, a questionnaire was completed by 245 English language teachers in Hesse; selected outcomes of the survey are reported in Section 2.4.

Chapter 3: Second language pronunciation

The research literature relevant to the topic was considered and this review is presented in Chapter 3. The chapter begins by looking at previous pronunciation studies (Section 3.1). For insight into pronunciation acquisition, Section 3.2 presents the major theories that have been proposed to explain second language acquisition. The chapter goes on to examine the most widely accepted models of speech perception and production (Section 3.3). In addition, the mechanisms underlying speech motor control (Section 3.3.2) and mirror neurons are outlined (Section 3.3.3). Drawing from the content at hand, the rationale and theoretical paradigm for the new pronunciation treatment is developed (Section 3.4). With reference to the requirements of the school curricula (see Chapter 2) and the presented research, Section 3.5 specifies the core set of speech sounds to be targeted in the intervention. Given that L2 pronunciation acquisition is also affected by non-linguistic factors, the influences of age of learning, gender, language use, motivation in formal instruction, and socio-economic factors, are all discussed in Section 3.6. Looking at individual differences in pronunciation, the

rationale is given for choosing beginner-level language learners aged 10 to 12 years as the subjects for this intervention study (see Section 3.6.6).

Chapter 4: The pronunciation intervention

Chapter 4 introduces the pronunciation intervention programme that forms the centrepiece of this thesis. Over the course of five months, the programme was implemented in a grammar school and a comprehensive school in Germany in the English lessons of children learning English as a foreign language. This chapter also presents the tasks and materials used and describes the procedures for the classroom work.

Chapter 5: Methodological considerations: the acoustic phonetic basis of this current investigation

This research project encompasses second language pedagogy as well as applied linguistics. To evaluate the outcomes of the pronunciation intervention, auditory as well as acoustic analyses of the pronunciation data were carried out. Chapter 5 introduces the acoustic phonetic basis of the current investigation with regard to the speech sounds selected as core to the pronunciation programme. In this account, first the source-filter theory of vowels (Section 5.1) is introduced and then the specific acoustic features needed for the analyses of the vowels (Section 5.2), diphthongs (Section 5.3) and consonants (plosives, fricatives, affricates and approximants; see Section 5.4) are discussed in more detail. Finally, other factors influencing speech production, such as duration, frequency effects, coarticulation and emotions, are considered in Section 5.5.

Chapter 6: Research focus

Chapter 6 draws upon the material introduced so far in the thesis and addresses the following research hypotheses. Given that all subjects participating in this study are subject to maturation processes and ongoing English language input, the first hypothesis (H1) states that the performance of both the intervention and control groups will improve over time (DeCoster, 2001; Mackey & Gass, 2005). It is the purpose of this thesis to show that the pronunciation training will significantly advance the intervention pupils' pronunciation abilities over and above the effects of maturation processes.

Accordingly, this postulation is the basis of the second hypothesis (H2). The third hypothesis (H3) accounts for the assumption that pronunciation performance will show the greatest improvements directly after the treatment (Shadish, Cook, & Campbell, 2001). A number of studies (Best, 1995; Flege, Schirru & Mackay, 2003; Iverson et al., 2003; Kuhl, 2000) have found that a diverse range of L2 speech sounds pose different degrees of difficulty for a language learner. Therefore, the fourth hypothesis (H4) addresses the performance level for each targeted sound. It remains to be seen if the treatment will help develop all target sounds in the same way or whether some acquisitional patterns will emerge. To account for the premise that the improved pronunciation performance of the treatment group will be due not only to the training effects of the intervention but also the intervention itself, the test items that are explicitly trained in the treatment are correlated with untrained test items. Within this context the fifth research hypothesis (H5) addresses the training effects. Frequency effects are known to influence speakers' performances (Bishop & Keating, 2012; Cholin, 2008) as more frequent use leads to increased pronunciation accuracy. Thus, the study controls for higher and lower frequency test items in both the intervention and control groups. Research hypothesis (H6) assumes that higher frequency items will outscore the lower frequency items.

Chapter 7: Methods

Chapter 7 lays out the methodology used in this thesis. To investigate the pronunciation intervention, quasi-experimental classroom research was conducted using a control group design (see Section 7.1). After introducing the ethics approval for this study, this section goes on to introduce the participating schools, students and teachers (Section 7.2) and explains the choice of speech materials used in the pronunciation analyses (Section 7.3). Section 7.4 presents the data collection methods and focuses in particular on the elicitation techniques as well as on the recording of the stimuli. The assessment of the pronunciation data included two levels of analyses (see Section 7.5): firstly, all stimuli were assessed auditorily using a three-point scale to evaluate the students' performance (Section 7.5.1). Secondly, acoustic data analysis (Section 7.5.2) using the same data was carried out using the Praat software. To explore the changes in pronunciation, Section 7.5 provides a detailed description and examples of specific acoustic parameters including fundamental and formant frequencies, duration,

amplitude, spectral peaks and spectral rate of change. Finally, the methods chapter concludes by providing the results of the interrater reliability test.

Chapter 8: Results

The first section of Chapter 8 presents the results of the auditory analysis (see Section 8.1) showing the main effects of time, training, frequency, sounds and group as well as interaction effects. These results are discussed in detail in Section 8.2. In contrast to the auditory analysis which involves the data set as a whole, the results of the acoustic analyses are presented separately for the vowels, diphthongs, approximants, plosives, fricatives and affricates due to the specific, inherent features of each sound group (see Section 8.3). Chapter 8 concludes with an analysis of the acoustic results (Section 8.4).

Chapter 9: Conclusions and implications

The final chapter draws upon the entire thesis, tying up the various theoretical and empirical strands in order to critically evaluate the auditory and acoustic outcomes of the intervention study. It offers some conclusions with reference to the results and discusses the limitations of this study.

2. Setting the scene: English pronunciation teaching in the foreign language classroom in Germany

It is the aim of this thesis to empirically investigate English pronunciation teaching in the foreign language classroom in Germany. Therefore, the purpose of the next section is to examine how English pronunciation is treated in the foreign language classroom in order to set the scene for a pronunciation intervention study. The introductory Section 2.1 discusses the aim of pronunciation teaching in foreign language acquisition. As the guidelines and subject matter of pronunciation teaching are anchored within the respective educational curricula, these are evaluated in Section 2.2. Teachers typically use their textbooks as their main frame of reference, so in Section 2.3 the textbook contents are analysed with regard to English pronunciation work. To find out how teachers actually implement the guidelines, use their textbooks and carry out pronunciation teaching in English language classrooms in Germany, a questionnaire was completed by 245 English language teachers in Hesse and selected outcomes are presented in Section 2.4. Finally, Section 2.5. summarises the information presented in this chapter.

2.1 Pronunciation model

Non-native and native accents are regarded as the principal factors in impaired intelligibility and fluent communication respectively (Cruz-Ferreira, 2009). Thus, it is generally accepted that L2 learners should aspire to a standard pronunciation model in order to be intelligible to their interlocutor and vice versa (Richardson, 2008). Nevertheless, there has been much debate and controversy about the nature of such a language model. Traditionally, the speech of educated native speakers in one of the `inner circle´ of long-established, English-speaking countries, such as `received pronunciation´ (RP) in the UK and to a lesser extent `general American´ (GA) in the USA, were seen as the standard prestige models of English language pronunciation (cf. Kachru, 1985; Richardson, 2008, p. 23). Today, this picture is not so clear. The use of English as a *lingua franca* invites contentious discussion of speech standards. As a consequence, there are now several approaches adopting `world English´ as a model and favouring intelligibility over specific speech model standards. This `world English´ pronunciation is

> formed by an amalgam of features from various native-speaker standards (including RP and GA); and that this amalgam may be further altered by reducing the number of contrasts of sounds and changing the usual (the "default") realization of sounds, to take into account the L1 transfer and to form a possible international English lingua franca (Cruttenden, 2008, pp. 317–18)[9].

Despite the considerable appeal of this approach, teaching `world English´ pronunciation in the foreign language classroom appears to be rather challenging because the pronunciation parameters are difficult for teachers to specify as well as for learners to adopt. In contrast, a standard model of pronunciation provides a uniform point of reference for L2 speakers with all kinds of different L1s. Thus, a standard model offers common ground for teachers and learners and opens up a large communicative radius which affords greater intelligibility and comprehensibility (Jenkins, 2000). As Peter Ladefoged (2007, p. 27) puts it: “it is best to teach some standard form of speech. It’s like helping people to dress appropriately, whatever the occasion”. There are several acceptable language models and it is beneficial for teachers to discuss these with their students. Nevertheless, with reference to pronunciation teaching, it is reasonable for teachers ultimately to decide on a model that is suitable and which they are able to speak themselves (Cruttenden, 2008).

In Germany most English teachers are not English native speakers and are thus often limited by their own foreign accent (Hedge, 2000). As acquiring a near-native pronunciation can be an unrealistic goal for many learners, excessively high expectations might also produce negative effects such as frustration in students and teachers (Grotjahn, 1998; Morley, 1991). Nevertheless, it is a principal aim for learners to pronounce the target language in a fully comprehensible manner (Kanellou, 2009).

The educational infrastructure in Germany is in a transitional state. In Hesse, new educational ‘standards’ began to replace the existing school curricula in the 2011/2012 school year (Klieme et al., 2004) including the English language curricula. These curricula referred to one of the standard varieties of English as a goal for pronunciation teaching (Hessisches Kultusministerium, 2010b). However, there is no longer any reference to a specific standard speech model in the educational standards, nor in the Common European Framework of References (CEFR) (cf. Council of Europe, 2001). These developments might be linked to recent debates about `world English´ (see discussion earlier in this section).

[9] See Gnutzmann (1999) and Jenkins (2000) for an overview of English as an international language.

2.2 Pronunciation curricula

To gain clearer insight into current pronunciation learning and teaching in Germany, this section looks at the curricula for foreign language teaching. However, there are numerous different specifications for different types of schools, different federal states and various age groups. As the aim of this thesis is to evaluate a pronunciation intervention delivered to grade five pupils (ages 10 to 12) in two schools in Frankfurt, this section will focus only on the guidelines for early secondary education in the relevant types of school types in the state of Hesse. As background, a short overview of the German school system is necessary before turning to the curricula analyses.

After four to six years of primary school, most parts of Germany offer a tripartite secondary school system: The *Gymnasium* (~ grammar school) takes eight or nine years and prepares students for university. The school leaving diploma is the *Abitur* (~ A-levels). The *Realschule* takes six years and offers an intermediate high school leaving diploma whereas the *Hauptschule* also offers an intermediate high school diploma after five years but prepares its students for a non-academic education. There are also *Gesamtschulen* (~ comprehensive school) which comprise all three school types[10]. In order to recruit a wide range of pupils with different backgrounds and abilities, the research project was carried out in the fifth grades of both a grammar school and a comprehensive school in Frankfurt am Main, Hesse. The German education system is regulated by the 16 German federal states. At the time of the intervention and data collection for the research in 2010, school-based teaching and learning in Hesse was regulated by the compulsory requirements of the Hessian Curriculum (Hessisches Kultusministerium, 2011b). In order to establish the nature of English L2 pronunciation teaching in the respective schools and grades, the Hessian curricula for G8[11] grammar and comprehensive schools were examined.

2.2.1.1 Pronunciation in the comprehensive school curriculum

The comprehensive school curriculum gives quite a detailed description of the pronunciation requirements in grade five. It states that the English sound system should

[10] To enhance readability, `grammar school´ will be used for *Gymnasium* and `comprehensive school´ for *Integrierte Gesamtschule*, but the author is mindful that these terms have different meanings in different countries.

[11] Between 2004 and 2007, Hessian grammar and comprehensive schools reduced the period from the beginning of secondary education to the Abitur (~A-levels) from nine to eight years (Hessisches Kultusministerium, 2010a). Thus, the new curricula are called G8 instead of G9.

be introduced with the help of pronunciation and intonation work as well as discrimination tasks and authentic listening comprehension. Moreover, pupils should learn to read phonetic transcription. As the comprehensive school includes all three school types *Hauptschule, Realschule,* and *Gymnasium* (see Section 2.2.), the comprehensive school curriculum encompasses all three corresponding curricula (Hessisches Kultusministerium, 2011b), which are summarised in the following sections[12].

2.2.1.2 Pronunciation in the grammar school curriculum

The grammar school curriculum for grade five for English as the first foreign language recommends listening comprehension tasks and pronunciation work with the help of simple sentences (Hessisches Kultusministerium, 2010b).

2.2.1.3 Pronunciation in the *Realschule* curriculum

In the *Realschule* curriculum, the topics of pronunciation, stress and intonation are seen as part of vocabulary teaching and basic language skills such as listening and speaking (Hessisches Kultusministerium, 2011d, p. 5). So, within the scope of vocabulary teaching, the curriculum specifies that the students should:

- understand and recognise the sounds of familiar English words or phrases and thus be able to divide the sound stream into entities of meaning;
- be able to identify short forms, weak forms and linking forms;
- identify the meaning of word homophones (e.g. by – bye) with the help of context;
- be able to identify similar sounding entities as structural units or words with reference to the context (e.g. he's/his, it's/its, you're/your, they're/their/there);
- identify the sentence type due to English intonation patterns (e.g. falling tune: declarative sentence, imperatives, etc.; rising tune: questions, etc.); and
- understand simple English utterances in the established pronunciation standards of `General British´ (Received Pronunciation (RP); modified standard) and `General American´.

With reference to speaking and listening skills, the *Realschule* curriculum specifies that students should be able to correctly articulate the following English phonemes withspecial attention to:

- English sounds that are unfamiliar to the German language /r,T,D,w,eI,@U/;[13]

[12] The author translated the content of the curricula.
[13] Faulty phonetics in all curricula versions published by the ministry.

- specific distinguishable sounds: /w/:/v/, /T/:/f,s/, /D/:/d,z/, /tS/:/dZ/, /i:/:/I/, /e/:/&/.[13]
- final voicing of plosives (/b/,/d/,/g/) and fricatives (/v/), such as in `job´, `bed´, `food´, `bag´, `leg´, and `live´;
- the levelling of functional final consonants, e.g. plural-s (as /z/ in `bags´, /s/ in `books´), 3rd person singular in the simple present (e.g. /z/ in `reads´, /s/ in `sits´), past-tense endings (/d/ in `cleaned´, /t/ in `looked´, /id/ in `started´);
- difficult consonant clusters (e.g. /vz/ in `lives´, /ts/ in `table-cloths´, /Dz/ in `clothes´, */ksT/*[13] in `sixth´, */tT/*[13] in `eighth´);
- the English alphabet and the spelling of English words; and
- the changing pronunciation of the determiner depending on the following sound (e.g. /ðə/ `the book´, /ðɪ:/ `the animal´).

Moreover, students should be able to[14]:

- use English phonetic transcription receptively with the help of key words;
- identify familiar English words and phrases in their written form, associate them with their sound structure, and utter them with the correct pronunciation and intonation; be aware of silent letters (“b” in `climb´, “d” in `sandwich´, “k” in `know´, “gh” in `eight´, “l” in `talk´, “w” in `write´, “u” in `guest´, “h” in `hour´);
- read familiar texts with correct pronunciation and intonation;
- read re-arranged texts with familiar words adequately;
- orientate their pronunciation and intonation of words and sentences towards `General British´ and `General American´ (Hessisches Kultusministerium, 2011d);
- check the pronunciation with the help of the phonetic transcription of key words and the word lists of the textbooks; and
- use about 650 words with adequate command of their meaning, use, pronunciation and orthography (Hessisches Kultusministerium, 2011d).

2.2.1.4 Pronunciation in the *Hauptschule* curriculum

The *Hauptschule* curriculum for English in grade five states that listening comprehension depends on the discrimination and recognition of sounds. Therefore, sound structures of the spoken language need to be correctly identified and the learners need to master pronunciation to the level of intelligibility. Right from the beginning, teaching and learning should make use of authentic language situations and train the correct pronunciation with the help of playful elements, tongue training (see below), rhymes, songs, reading aloud to the class and on tape, as well as intonation and discrimination exercises. Motivational imitation activities are recommended, along with the frequent use of recordings by native speakers which should convey intonation

[14] The curriculum also demands the of use multimedia teaching aids (CD-ROM) to train pronunciation and vocabulary (Hessisches Kultusministerium, 2011d). Nevertheless, the students participating in this research did not use these aids so they are not reviewed in this text.

patterns and colloquial weak and short forms and should especially help weak learners to develop the receptive competence needed for listening comprehension practice (Hessisches Kultusministerium, 2011c).

More specifically, the students should be able to[15]:

- correctly perceive and reproduce the sounds of presented vocabulary;
- recognise the different intonation patterns of various sentence types (e.g. rising tune, falling tune);
- correctly pronounce English sounds with special attention to the English sounds that are not part of the German sound inventory (e.g. `this´, `these´, `bag´, `job´);
- read words and short texts with the corresponding intonation;
- read and understand some phonetic transcription;
- know about 400 items of vocabulary including idioms, cardinal numbers, irregular plural forms and classroom phrases.

2.2.1.5 Curricula requirements and pronunciation teaching

In contrast to the grammar school curriculum, which gives only very broad guidelines in recommending listening comprehension tasks and pronunciation work with the help of simple sentences (Hessisches Kultusministerium, 2010b), the curricula for *Realschule* and *Hauptschule* offer quite detailed information about the contents of pronunciation teaching in grade five. On further inspection, the information set out in those two curricula is partly unclear, e.g. the *Hauptschule* curriculum (see Section 2.2.1.4.) states that "students should master pronunciation to the level of intelligibility", but there is no definition or elaboration of what intelligibility involves. Additionally, students should be able to read and understand `some´ phonetic transcription. Again, this statement is vague and more ambiguities can be found upon closer scrutiny. The *Hauptschule* curriculum does offer some suggestions for teaching methods, but these statements are also rather imprecise: for example, to teach pronunciation with the help of "tongue training" and "playful elements" or to use "motivational imitation activities". The *Realschule* curriculum (see Section 2.2.1.3) gives more detailed information than the *Hauptschule* curriculum. Moreover, in the curricula there are few guidelines on methods to implement and systematise pronunciation teaching in the classroom. However, the Hessian Ministry of Education states in its curricula that relevant textbooks provide the basis for instruction and implementation of the curricula contents for grade five (cf. Hessisches Kultusministerium, 2010b). Accordingly, an analysis of the textbooks was undertaken and this is reported in Section 2.3 below.

[15] List directly translated from Hessisches Kultusministerium (2011c).

2.3 Textbook and pronunciation

The textbook is regarded as the most important medium of teaching and learning in German schools and it functions as a set of guidelines for teachers. The choice of a textbook has direct consequences for the quality and content of the lesson (Kieweg, 1998, p. 27). A review of all English textbooks for the fifth grade in all school types would be beyond the scope of this thesis, so this section is confined to an analysis of the textbooks used by the participants in the intervention study, namely Cornelsen's *English G 21 A1* (Schwarz, 2006) and Diesterweg's *Notting Hill Gate 1* (Edelhoff, 2007b), to seek insights into the daily practice of pronunciation teaching. A textbook (TB) is usually published with a set of components such as a student workbook (WB), a teachers' manual (TM), an audio CD and other additional materials, which are also included in the following analysis where necessary. To evaluate the pronunciation practice in these textbooks and materials, Kieweg's (1998) checklist of criteria for evaluation of textbooks is used. With reference to pronunciation, Kieweg offers eight criteria (see Table 2-1). As some aspects were outdated (e.g. use of audio cassette) or not specific enough (e.g. audio examples mentioned only with reference to imitation), some minor amendments to the original list have been made.

Table 2-1: Textbook analysis checklist: pronunciation practice (Kieweg, 1998)

	Criteria	Content
1.	Extent of phonetic activities	How much (if any) pronunciation practice can be found in the textbook?
2.	Integration of textbook and audio material	Is there enough audio material? Does the textbook support the audio material?
3.	Phonetic transcription	Is there systematic progression in developing the phonetic transcription abilities of the students?
4.	Stress and prosody	How are the students introduced to the concepts of stress and prosody?
5.	Reference to standard varieties	Does the textbook refer to the differences between American English (AE) and British English (BE) standard varieties?
6.	Phoneme-grapheme correspondence	Does the textbook offer explanations with reference to the different phonetic realisations of graphemes (e.g. /i/ → <he, see, sea, believe, key>)?
7.	Phonetic characteristics of the spoken language	Does the textbook offer information regarding the characteristics of the spoken language, such as devoicing, contractions, elisions, etc.?
8.	Discrimination of sounds	Is there an introduction to the discrimination of sounds at the beginning of the pronunciation instruction?

Regarding the first criterion: "Extent of phonetic activities", both textbooks *English G 21* and *Notting Hill Gate 1* include many pronunciation activities throughout all units and materials, as shown in the following Table 2-2:

Table 2-2: Texbook analysis: Criterion 1: Extent of phonetic activities

Cornelsen English G21 A1	Diesterweg Notting Hill Gate 1
Unit 1 TB, p. 30, 14: a /ə/ or an /ən/? (+ audio) → TM, p. 94f. Solutions 14, (no explanation) → WB, p.20, 20, a /ə/ or an /ən/? (+ audio) TB, p. 31, 15: `the´ with /ə/, /i/? (+ audio) → TM, p. 95, Solutions 15 (no explanation) → WB, p. 20, 21: `the´ with /ə/, /i/? (+ audio)	**Unit 1** TB, p. 26. P2: Rhyming pairs (no explanation) (+ audio) → TM, p. 61, Solutions P2 (no explanation) WB, p. 15, C9: Sound check /ɪ/ vs. /i/ (+ audio) → TM, p. 60, Solutions C9 (no explanation)
Unit 2, TB, p. 42, 2: Plural `s´ (+ audio) → TM, p. 122, Solutions 2, (explanation voiced/ voiceless) → WB, p. 26, 5, Plural `s´ (+ audio) → Grammar File 2, p. 131f., Plurals TB, p. 43, 5: The `-s´ in the simple present (+ audio) (voiced/voiceless used, but no explanation) → TM, p. 123, Solutions 5, (no explanation) → Grammar File 2, p. 133	**Unit 2** TB, p. 46, P3: Juice and jam /dʒ/ (+ audio) → TM, p. 97: Solutions P3 (no explanation) WB, p. 28, C8: Match the rhyming words (+ audio) → TM, p. 96, Solutions C8 (no explanation)
Unit 3 TB, p. 62, 12: /æ/ and /eɪ/ (+ audio) → TM, p. 156, Solutions 12 (no explanation) → WB, p. 40: /æ/ and /eɪ/ (+ audio)	**Unit 3** TB, p. 66, P2: Sound Check /əʊ/ and /ɒ/ (+ audio) → TM, p. 129, Solutions P2 (no explanations) WB, p. 38, B7: Rhymes (+ audio) → TM, p. 120, Solutions B7 (no explanation) WB, p. 38, B8: Sounds /dʒ/, /ʃ/ (+ audio) → TM, p. 120, Solutions B8 WB, p. 43, C9: Sound check /i/ (+ audio) → TM, p. 128, Solution C9 (no explanation)
Unit 4 TB, p. 80, 15: /əʊ/ and /ɒ/ (+ audio) → TM, p. 188, Solutions 15, (no explanation) → WB, p. 55, 16: /əʊ/ and /ɒ/ (+ audio)	**Unit 4:** TB, p. 76, Tipp Mr. /mɪstə/ Mrs /misɪz/ TB, p. 83, C9: Weather Poems /v/, /w/ (+ audio) → TM, p. 165, Awareness of pronunciation and intonation (no explanation) TB, p. 84, P5: How to say `u´? (+ audio) → TM, p. 167, Solutions: P5
Unit 5 TB, p. 92, 8: Past tense forms -/d/,-/t/,- /ɪd/ (+ audio) → TM, p. 218, Solutions 8 → WB, p. 61, 8: Past tense forms -/d/,-/t/,- /ɪd/ (+ audio) → Grammar File, p. 142 (no explanation)	**Unit 5** TB, p. 105, P6: A tongue twister /θ/, /ð/ (+ audio) → TM, p. 203, drilling should be used, (no explanation on how to pronounce the “th”) WB, p. 62, A5, Sound Check, silent letters (+ audio) → TM, p. 176, Solutions: A5 (no explanation)
Unit 6 TB, p. 108, 8: /ʃ/,/tʃ/, and /dʒ/ (+ audio) → TM, p. 245, Solutions 8 (explanation) → WB, p. 71, 8: /ʃ/,/tʃ/, and /dʒ/ (+ audio)	**Unit 6** TB, p. 120, P4, /æ/ and /e/ (+ audio) → TM, p. 231, Solution P4 (no explanation) WB, p. 78, A8, Past tense forms -/d/,-/t/,- /ɪd/ (+ audio) → TM, p. 218, Solutions A8, (no explanation) → Language in Focus 21, p. 171, -/d/,-/t/,- /ɪd/ WB, p. 82, B4, Sound check, Find words in a jumble of letters (+ audio) → TM, p. 224, Solutions B4
English Sounds, p. 147 (sounds with example words)	English Sounds, p. 173 (sounds with example words)
The English alphabet, p. 147	The English alphabet, p. 173
TM, p.57f, Drilling	TM, p. 29f. Drilling
Vocabulary, p. 149: Pronunciation Vocabulary, p. 150: Intonation Vocabulary, p. 152, Linkings TM, KV4, master copy, English sounds (pictures with sounds)	Language in Focus 7, p. 156, a/an/the Language in Focus 8, p. 158, Plural of nouns

For both textbooks there is a teacher's CD that includes all the main textbook texts, dialogues, songs and poems. In total, *Notting Hill Gate 1* includes about 145

minutes and *English G 21* about 76 minutes of audio material, which is recorded by English native speakers (Schwarz, 2007). Moreover, a student's CD is provided containing the audio files for the pronunciation activities in the workbook. On the whole, there is strong and effective integration of textbook and audio material (see Criterion 2, Table 2-1). Additionally, both textbooks offer a wide variety of accompanying online resources which include a large amount of audio and video material.

Criterion 3 (see Table 2-1) relates to whether *Notting Hill Gate 1* and *English G 21* provide for systematic progression in developing the phonetic transcription abilities of the students. Both textbooks offer an overview of the English sounds with example words (*English G21* also supports the examples with pictures on the master textbook copy). Additionally, there is phonetic transcription in the vocabulary and dictionary sections of the two textbooks. Both books also concentrate on some specific English sounds in the pronunciation tasks (see Table 2-2). So, there are several activities that use phonetic transcription. However, there is no reference in the teachers' manual, the textbook, the workbook, or on the audio CD to how to introduce phonetic transcription to the students. Moreover, there is no rationale given for the choice of content nor any explanation of the articulation of the sounds. So, with regard to phonetic transcription (Criterion 3) there is not adequate progression in developing the phonetic transcription abilities of the students in any of the teaching materials and there is no background information about the concepts underpinning pronunciation teaching.

English G 21 introduces the phonetic symbols for the transcription of stress and these are linked to two info-boxes on the 'English sounds' page (Schwarz, 2006) and also to the vocabulary section. *Notting Hill Gate 1* also introduces the stress mark diacritics on the 'English sounds' page. Although the stress and linking diacritics are used frequently in the phonetic transcription throughout the book, there is no further reference to the concepts of stress and prosody (Criterion 4, Table 2-1). In their storylines, *English G 21* and *Notting Hill Gate 1* refer to Great Britain with its multicultural society (Edelhoff, 2007a; Schwarz, 2007). Although there is no explicit reference to a standard model of English (see Criterion 5, Table 2-1), British English seems to be the focus of the books. The choice of topics, the British spelling, and the choice of narrators on the audio CD all support this view[16].

[16] None of the speakers on the recorded material shows a trace of regional, ethnic or other varieties of English.

Phoneme-grapheme correspondences (see Criterion 6, Table 2-1) are implicitly trained in tasks where students have to match words according to their sounds (see Table 2-2). Nevertheless, neither *English G 21* nor *Notting Hill Gate 1* offer explanations of the different phonetic realisations of graphemes.

The seventh criterion (see Table 2-1) investigates whether the textbook offers information regarding the characteristics of the spoken language, such as devoicing, contractions, elisions, etc. Although teaching materials in both textbooks use the terms 'voiced' and 'voiceless' in the teachers' manuals and apply contractions in the phonetic transcription, no further information is given about the nature of voicing or its relevance for pronunciation teaching. Generally, there is minimal information on the characteristics of the spoken language – for both teachers and students.

This lack of information is also reflected in the pronunciation exercises where the students are required to match or listen for specific sounds, but the exercises offer barely any explanation regarding the pronunciation task in focus. Moreover, there is no introduction given to the discrimination of sounds at the beginning of the pronunciation instruction or to the use of phonetic transcription (see Criterion 8, Table 2-1).

Applying Kieweg's (1998) pronunciation criteria, both of the analysed textbooks include a large number of pronunciation activities and, notably, there is a lot of audio material. Nevertheless, in the textbooks there is no statement on the concept of pronunciation teaching, no identification of problematic pronunciation areas and/or how to deal with them in class. Although phonetic transcription is frequently used, there is no statement on how to implement it in the classroom or any reference to a language standard. Although the exercises tackle several problematic pronunciation areas, there is no systematic progression in the pronunciation training.

With regard to the curricula (see Section 2.2), both textbooks put the official specifications into action. There are numerous tasks with speech production by native speakers using the `General British´ pronunciation standard. The phonetic transcription is used receptively and there are many tasks that deal with the identified pronunciation problem areas, such as the sounds /θ, ð/, 3rd person singular, plural and past tense endings, etc. Nevertheless, as already mentioned, there is still a lot of important information missing concerning the pronunciation teaching approach, progression through tasks, and background information regarding the addressed pronunciation areas. So, with reference to the curricula and the analysed textbooks, the planning and

implementation of pronunciation teaching is largely the responsibility of the respective teacher. To shed light on how the teachers actually do the pronunciation work in the English foreign language classroom, Section 2.4.presents the outcomes of a teacher questionnaire on pronunciation teaching.

2.4 Questionnaire data

This section describes the daily practice of English L2 pronunciation teaching in the lower secondary grades in Hesse, Germany. The information was gathered using an online questionnaire[17] for teachers which was designed and piloted by the author and approved by the Hessian Ministry of Education[18]. Using the mailing list of the Centre for Teacher Education in Frankfurt (Zentrum für Lehrerbildung, (ZFL, 2011) the online questionnaire was sent via email to all secondary schools in Hesse[19]. The complete questionnaire is presented in the appendix (see p. 217). Given the large amount of content covered in the survey, only selected topics will be summarised in this section.

In total, 245 English language teachers from all types of schools in Hesse completed the questionnaire. Almost one quarter of the respondents (24.9%) taught at an integrated or non-integrated comprehensive school; 38.3% taught at a grammar school; 8% taught at a *Realschule* and 3.4% had a *Hauptschule* background. The participating teachers mainly taught students between the ages of 10 and 15 (see Table 2-3).

Table 2-3: Taught student age (Teacher questionnaire)

	Students at the ages of about 10-12	Students at the ages of about 13-15	Students at the ages of about 16-19
Valid Percentages	41.8%	49.4%	28.7%

Almost three quarters (73.9%) of the participants are female and about one quarter (26.1%) male and they differed quite markedly with respect to their teaching experience (see Table 2-4).

[17] The questionnaire was carried out with Limesurvey software (Limesurvey, 2012)

[18] The approval to conduct a teacher survey at Hessian secondary schools was granted by the head of the division of primary schools as well as the data protection commissioner of the *Hessian Ministry of Education* on August 26th, 2009 (Reference number: 660.003.000-304).

[19] To maximise the level of participation, two iPod Nano were raffled among all participants.

Table 2-4: Years of teaching experience (Teacher questionnaire)

Years (teaching experience)	Less than 5	6-10	11-14	15-19	20-24	25-29	More then 25
Valid Percentages	32.4%	14.8%	15.9%	6.8%	4.5%	6.8%	18.8%

The respondents' answers were given with reference to the pronunciation work in grades five and six – that is, English as Foreign Language (L2) students at ages 10 to 12. This corresponds to the researched age group. Information was retrieved on the following topics:

- How big is the students' and teachers' interest in pronunciation teaching?
- How much and how often does pronunciation work take place?
- What kinds of activities are used to introduce and practise pronunciation?
- Which resources do teachers use to teach pronunciation?
- How do teachers deal with speech standards and varieties?
- What are the goals of pronunciation teaching?
- How do teachers introduce new items and correct pronunciation errors?
- What features of pronunciation are perceived as particularly difficult?
- What impact does the student's language background have on pronunciation errors?

Most of the participating teachers showed a strong or even very strong interest in integrating pronunciation teaching into their English lessons (see Table 2-5).

Table 2-5: Importance of pronunciation work (Teacher questionnaire)

	Very important	Important	Neither	Unimportant	Very unimportant
Valid Percentages	40.8%	40.4%	15.9%	2.4%	.4%

With regard to their students, 90% of the teachers were convinced that their students also have a strong or very strong interest in pronunciation exercises. This interest was also reflected in the frequency of pronunciation work in the classroom. Nearly half of the participating teachers reported doing pronunciation exercises at least once a week, and one quarter said they integrate them in every English lesson (see Table 2-6).

Table 2-6: Frequency of pronunciation work (Teacher questionnaire)

	Very often (in every English lesson)	Frequently (at least once a week)	From time to time (more than once a month)	Rarely (about once a month)	Hardly ever
Valid Percentages	25.1%	47.0%	23.3%	3.2%	1.4%

In general, the duration of these exercises ranged between five minutes (51.2%) and ten minutes (30.2%). Nevertheless, more than half of the teachers (55.1%) thought that

pronunciation exercises should be more strongly integrated into the English lessons. It was reported that pronunciation is very often taught with the help of words, phrases and sentences. Reading text passages and the use of dialogues were also very common methods. The textbook and the related materials are the main resources in the English lessons (Kieweg, 1998). However, 34.5% of the participating teachers did not think that these sources contained sufficient pronunciation exercises and almost half (44.5%) of the teachers said they would like to have more pronunciation tasks in their textbooks and accompanying materials. Additional materials were used by 27.9% of respondents to teach pronunciation.

The participating teachers indicated a very positive attitude towards their own language skills. Of the group, 65.8% referred to their own language level as near-native and one fifth (20.6%) even rated themselves as native-like. With reference to the variety of English they speak, nearly 70% of the participants classified it as British English; one third even identified it as `received pronunciation´ (see Table 2-7). Also, British English varieties were the most widely used in the classroom, followed by American English varieties.

Table 2-7: Speech standards (Teacher questionnaire)

	British English				American English	Australian English	New Zealand English	Other
	RP	***General***	***Scots***	***Irish***				
Valid Percentages	31.3%	37.9%	.4%	1.2%	25.5%	.8%	.4%	2.5%

For about three-quarters of the participating teachers, intelligibility was seen as the major goal of pronunciation teaching (see Table 2-8):

Table 2-8: Goal of pronunciation teaching (Teacher questionnaire)

	Native-speaker competence	Intelligibility	As long as the students talk at all, I'm happy
Valid Percentages	13.9%	75.4%	10.7%

With regard to phonetic transcription, 85.3% of the participating teachers felt competent enough to use it and almost two thirds of the teachers believed that it helps at least some students to improve their pronunciation competence (see Table 2-9). Despite these results, only 55.9% of the teachers reported integrating phonetic transcription into their lessons, where the students mainly had to understand it receptively and not productively.

Table 2-9: Impact of phonetic transcription (Teacher questionnaire)

	Yes, definitely	It helps for some students	No	I don't know
Valid Percentages	17.5%	52.1%	22.3%	8.1%

Of the respondents, 63.6% reported that they correct only their students' profound mistakes (see Table 2-10). There appeared to be some discrepancy here with regard to the teachers' opinions on how often their students wanted to be corrected, since the teachers thought that the students wanted correction less often (see Table 2-10, Table 2-11).

Table 2-10: Frequency of error correction (teacher) (Teacher questionnaire)

	At every mistake	Only at profound mistakes	From time to time	Rarely	Rather not
Valid Percentages	23.8%	63.6%	11.7%	0.9%	0%

Table 2-11: Frequency of error correction (Teacher questionnaire: teachers' opinions on students' preferences)

	At every mistake	Only at profound mistakes	From time to time	Rarely	Rather not
Valid Percentages	12.1%	42.4%	28.1%	10.4%	6,9%

The most frequent corrective action reported by the teachers was to model the correct version themselves. Other frequently used feedback options were to offer listening examples and to refer to similar-sounding words and to phonetic transcription.

76.1% of the teachers thought that the specific language background of an English learner accounts for specific pronunciation problems. The participating teachers identified various sounds as the most difficult ones for learners of English coming from German, Turkish, Russian and Polish language backgrounds, and these sounds are ranked in Table 2-12:

Table 2-12: Top five most difficult sounds by language background (Teacher questionnaire)

Top 5	German (1st question[20])	German (2nd question)	Turkish	Russian	Polish
1.	/θ/ (n=163)	/θ/ (n=94)	/θ/ (n=26)	/r/ (n=15)	/θ/ (n=10)
2.	/ð/ (n=141)	/ð/ (n=75)	/ð/ (n=22)	/θ/ (n=12)	/ð/ (n=8)
3.	/ɔ:/ (n= 93)	/v/ (n=31)	/eə/ (n=14)	/ð/ (n=10)	/r/ (n=5)
4.	/v/ (n=76)	/w/ (n=30)	/n/ (n=13)	/w/ (n=5)	/u:/ (n=3)
5.	/ʒ/ (n=75)	/eə/ (n=28)	/ɔ:/ (n=12) /r/ (n=12)	/ɔ:/ (n=4)	/əʊ/ (n=3)

[20] The teachers were asked to answer this question at two times in the questionnaire. Thus, both answers are presented in Table 2-12.

Of the teachers in this sample, 55.1% thought that pronunciation exercises should be more strongly integrated into English lessons, while 40.8% did not feel adequately educated in pronunciation teaching and 61.9% said they would like to take part in pronunciation teacher trainings.

Studies have shown that pronunciation teaching plays only a marginal role in standard foreign language teaching (Haß, 2006; Mehlhorn, 2005). However, the teacher questionnaire conducted for this study, although geographically restricted to Hesse, clearly showed a different picture. To summarise the key findings of the questionnaire: most of the participants had a very strong interest in pronunciation teaching and reported carrying out pronunciation exercises in lower secondary L2 English classes about every week and very often in every English lesson. This evident interest is supported by the survey finding that almost half of the participating teachers wanted more pronunciation exercises in the textbooks and accompanying materials and the fact that many teachers reported using additional materials. Nevertheless, it is important to reiterate that 40% of the teachers indicated that their own education in pronunciation teaching and learning was not sufficient and, even more striking, over 60% wanted to take part in pronunciation teacher training. These findings might bear a relationship to the flaws found in the textbooks and teachers' manuals, such as the lack of explanation of key concepts and the inadequate progression within the exercises (see Section 2-3 above).

2.5 Summary

This chapter intended to set the scene for the intervention study. First, it described current language pedagogy and the objectives of pronunciation teaching. It became apparent that pronunciation is taught only as a by-product to negotiate meaning in discourse and is not treated as a discrete objective. With the development of communicative approaches to language learning and the rise of global English, intelligibility has surpassed 'nativeness' as a principal objective. An examination of the relevant curricular guidelines offered a mixed picture. The former curricula referred to one of the standard English varieties as a speech standard. However, the new educational standards and the CEFR, which replaced the curricula, do not refer to a speech standard. However, it might still be advisable to use a standard variety of English as a reference point for students and teachers in order to give language

learners a pronunciation framework and yardstick. The former curricula also contained comprehensive guidance on which pronunciation features to teach, but they did not prescribe or recommend best methods to implement and structure pronunciation teaching in the classroom. The same issue holds true for the textbook analysis. Despite the fact that there is a significant number of exercises and pronunciation activities in the textbooks and numerous audio and video materials, there is little systematic structure to the tasks and minimal information on how to teach pronunciation. So, the responsibility for actual planning, methods and implementation of pronunciation teaching again rests predominantly with the class teacher. Importantly, the outcomes of the teacher questionnaire showed that although pronunciation was regarded as an integral part of English foreign language teaching, teachers did not feel adequately educated in pronunciation teaching and desired training in this area. This gap might be connected to the lack of explanatory detail in the textbooks and teachers' manuals.

It was asserted in the introduction to this thesis that in order to fully understand this field, an investigation of pronunciation teaching needs not only to examine closely the *status quo* of pronunciation teaching, it must evaluate the success of applied pronunciation trainings and consider the findings of second language acquisition research. Relevant research in applied linguistics is therefore the focus of the next Chapter 3.

3. Literature: Second Language Pronunciation

In this thesis a pronunciation training programme is developed (see Chapter 4) that is implemented within the language education of English L2 learners at the ages of 10-12 in Frankfurt, Germany. It is for this reason, that the literature chapter first looks at prior L2 pronunciation intervention studies (see Section 3.1) and then looks at explanations for what causes the difficulties in achieving accurate pronunciation. For this purpose, Section 3.2 provides an overview of several basic linguistic theories of 2nd language pronunciation such as native language transfer, language universals and the role of the similarity between the L1 and the L2. Then, Section 3.3 looks at linguistic theories of L2 speech perception and production with regards to the context at hand and also takes the physiological and neurological background of pronunciation learning such as speech motor control as well as mirror neurons into account. Deriving from the outlined theories and models, the theoretical paradigm which provides the linguistic foundation for the pronunciation programme investigated in this study is presented (see Section 3.4). Based on the presented findings, reasons for the choice of the pronunciation objective targeted in the intervention are given in Section 3.5. As non-linguistic factors such as age of learning, gender, language use, formal instruction, motivation and socio-economic factors also affect L2 pronunciation training (Major, 2001; Wieden & Nemser, 1991) they are presented in the last section 3.6 which also provides the rationale for the choice of research subjects. The literature chapter thus provides the framework for the conception of the pronunciation intervention, that is presented in Chapter 4.

3.1 Previous pronunciation instruction studies

Most of second language pronunciation research still centres around the theoretical processes of speech acquisition and even more on single isolated areas which are not directly applicable to language teaching. This might be due to the fact that research tends to focus on testing a theoretical proposal about the speech acquisition process itself and is less concerned about evaluating the efficacy of pedagogical interventions (Derwing & Munro, 2015). As outlined in the introduction to this thesis, the mostly form-focused pronunciation instruction has played a minor role for several decades due to the rise of communicative language teaching in the 1980s (see Chapter 1). However, in the past 10 to 15 years conference proceedings and research have shown a renewed

interest in pronunciation teaching that is rapidly growing (Thomson and Derwing 2014). Within this scope, there have been an increasing number of studies that have tried to test the efficacy of pronunciation training programmes (Thomson & Derwing 2015). However, the studies vary greatly, as pronunciation learning and language teaching in themselves are highly complex, and there are vast differences with regards to the training scope, choice of subjects, target language and elicitation and evaluation of the data. In their review paper, Thomson and Derwing (2014: 2f) look at the conception of pronunciation instruction studies and provide some benchmark with regards to a "good" pronunciation training study:

- "pronunciation research should be primarily concerned with helping learners to become more understandable;
- provide enough detail about participants and procedures to allow replication;
- have large enough samples to conduct statistical analyses, including effect sizes;
- employ a control group to verify that improvement is a result of instruction;
- not limit assessment stimuli measuring learners' pronunciation ability to reading aloud; extemporaneous or spontaneous speech that better reflects natural communication is important;
- include a delayed post-test to determine whether the intervention had a lasting effect;
- to address concerns regarding ecological validity, the ideal study should be conducted in the classroom;
- complementary qualitative analyses should be conducted to provide insights in learning, such as motivation, the nature of interactions in the L2, and other social influences".

Thomson and Derwing (2015) provide a meta-analysis and look at 75 pronunciation studies including peer-reviewed as well as unpublished manuscripts. However, only six studies present data for classroom-based interventions with younger learners (ages ranging from 8.4 to 20) in primary or secondary school education despite the high ecological validity (Cardoso 2010; Chen and Goswami 2011; Kennedy 2003; Lima 2010; Trofimovich et al. 2009; Tsiartsioni 2010) for reference see Table 10-1 in the appendix). One of the reasons why rather few researchers have chosen classroom-based learners in primary and secondary school education might be the complex nature of this kind of research: Ethics approvals are not only needed from the board of the researcher's university but also from the responsible body of the institution at hand. Moreover, parental consent is usually needed for younger students and the cooperation of the responsible teachers and departments is required. In addition, the relevant language curriculum has to be considered to ensure the ecological validity of the study. As it is usually not possible to research the students outside the classroom it might be

hard to find enough matching subjects in the intact classroom and to get a large enough sample size. Also, the attrition rate is rather high due to teachers and students falling ill and the school routine with its field trips and term restrictions (Derwing and Munro 2015).

In the following, those six studies dealing with younger classroom-based learners identified by Thomson and Derwing (2014) will be summarised, highlighting their individual approaches to research designs and analyses. One of these studies was carried out by Cardoso (2010) who looks at 30 monolingual speakers of Brazilian Portuguese (BP) at the ages from 15 to 20 (M = 16.3) who attend a secondary public school in Brazil. However, the intervention does not take place in a regular foreign language classroom but offers an extra-curricular five weeks course to learn an invented language called Slavir, which is designed to consist of many homorganic onset clusters (e.g. /sl/, /sn/, /st/). These often appear in foreign languages but are particularly difficult for BP learners as these clusters are not part of their sound inventory. Cardoso uses a quasi-experimental, within groups pre-test/ post-test design to test three different types of instructions: Teachability (Pienemann 1984), Projection of Markedness (Zobl 1983) and a combination of these two. To evaluate each of the three theoretical paradigms, three 30 minutes Slavir teaching sessions conducted in BP were offered. The Projection group was taught exclusively /st/ initial words, the Teachability group was taught one /s/ + consonant onset cluster (SC) per session following the natural order of acquisition (/sl/</sn/< /st/) and the mixed group was taught all three SC sequences throughout the Slavir course. The pronunciation sessions followed Thornbury's (2002) guidelines to teach vocabulary. After each of the three sessions a word reading aloud task was administered which was recorded, transcribed and analysed using accuracy scores. The result showed that the Projection group focusing on the more marked /st/ cluster had the best overall performance, followed by Teachability and the mixed group. Although Cardoso's study has some implication for language teaching as his study supports the Markedness Hypothesis and the author concluded that the instructional effects of mastering the most marked /st/ cluster projects to the acquisition of the less marked forms /sl/ and /sn/, the ecological validity can be questioned as there is no connection to a real foreign language classroom or curriculum or even a real language.

Similar to Cardoso, Chen and Goswami (2011) set out to test the impact of a theoretical paradigm on pronunciation learning. In their intervention study they

investigated how Cooperative Learning (CL) affects the pronunciation skills of 44 Mexican Spanish L2 learners at the ages 15 to 19. They focussed on seven English consonants (/t/, /d/, /v/, /z/, /ð/, /θ/, /š/) that are deemed particularly difficult for Spanish learners of English. The subjects were enrolled in two different pronunciation classes at a private, residential high school. One class received conventional teaching (n=25) while the other class was taught using a CL approach (n=19). Despite some indication of mostly task-oriented activities that were applied in the training programme of the CL classroom (e.g. Power Point, Jeopardy, Bingo, etc.), the paper does not outline in detail how both approaches were implemented. The phonetic and phonological instruction of how to pronounce the sounds and words including those seven English consonants took place for 90 minutes a day from Monday to Friday over the course of six weeks which amounts to 45 hours of teaching time in total. The pronunciation of the target sounds was audio taped and video recorded before and after the intervention. The researchers themselves recorded and evaluated the pronunciation data as either correct or incorrect. The overall performance for both groups showed a significant increase in the pronunciation scores which was assigned to the implementation of phonetic and phonological instruction in general. Nevertheless, there was no statistical difference in the performance between the intervention and control group, supporting the view that CL did not positively impact on the development of pronunciation skills.

According to Lima (2010) Brazilian English L2 textbooks usually focus on a worldwide audience with different L1's and thus do not specifically take problems of Brazilian learners of English into account. Thus, the study (2010) investigated whether the explicit teaching of segmental features of English sounds that are problematic to Brazilian EFL learners (i.e. /θ/,/ð/, /æ/, /ɬ/,/uː/, final /t, d, ɪd/,/iː/, /h/, /t/, /iː/, /h/, /r/, /ɪ/,/ə/, /ə/, /ɽ/) might enhance pronunciation proficiency. Lima applied interventionist action research and looked at 28 11- to 13-year-old students with a basic level of English in two intact classes. Both groups received the same amount of English instruction (59 hours) that also included pronunciation work (4 hours) over the course of one semester. In contrast to the control group (n = 11) for which the exact time and content of pronunciation work was not specified, the intervention group (n = 17) dedicated 15 minutes from the regular class each week on explicit teaching of the identified segments. The pronunciation sessions were structured according to the communicative framework to teach pronunciation as proposed by Celce-Murcia et al. (1997). First, the sounds were introduced and it was explained how the sound is articulated and minimal

pair activities were used. After that the sounds were practised in controlled tasks and finally communicative activities with meaningful and authentic discourse took place. The pronunciation performance was tested before and after the intervention and a delayed post-test took place 11 months after the intervention was completed. Reading aloud diagnostic tests that included words and sentences containing the target sounds were carried out and the results were presented with the help of the reduction in error occurrences between the tests. The evaluation of the data showed a higher decrease in error occurrence for the intervention group (56% vs. 46%) between pre- and post-test and for each participant in comparison to the control group (57% vs. 54%). The delayed post-test applied 11 months later demonstrated even slightly lower error occurrences (46% vs. 43%) between post- and follow-up test in the intervention group. Lima argues that this supported the durability of the explicit pronunciation teaching. However, no statistical analysis of the presented data was provided and there was also a slight decrease in error occurrence in the control group between the post- and follow-up test (54% vs. 51%).

Kennedy (2003) looked at the effects of corrective feedback provided by peers on the pronunciation of the participants. 47 L1 Canadian French speaking junior college students between the ages of 17 to 20 in an English as a second language classroom participated in the study. The students had the same level of English. The assignment to experimental (n = 22) and control group (n = 25) took place via intact classes. For both groups language teaching took place three hours per week for 15 weeks while nine weeks were dedicated to the treatment phase in the intervention group. After a pre-treatment phase in which language samples were collected to target problematic sound, the experimental group was then trained on how to pronounce target words and instructed how to provide corrective feedback on the “th” sounds /θ/ and /ð/ which are not part of the French sound inventory. The nine-week treatment phase then allowed the participants to provide oral peer corrective feedback to each other. The training material consisted of Power Point slides showing how to correctly produce /θ/ and /ð/ and examples on how to deliver peer corrective feedback. A pre- / post-test design was applied in which a prepared dialogue was presented by the participants and audio recorded. The students’ production of the “th” was rated by the researcher with the help of accuracy scores (0. -1, -2) of the number of produced “th”-sounds, eg. if the transcript showed 42 obligatory productions of “th” and produced it 9 times native like, the

accuracy score was 9/42 or .21. The results did not show significant differences between the two groups.

Trofimovich, Lightbown, Halter and Song (2009) designed a longitudinal study examining whether comprehension practice in listening and reading in the absence of speaking could help to develop L2 pronunciation skills. They hypothesed that a learner's success in mastering a second language depended on the knowledge about the language itself. They looked at 74 students (mean age 8.4) in year three, in which the French L1 students started to learn English, and year four from 20 intact ESL classes in Canada. 12 classes implemented the training programme consisting entirely of listening and reading activities. The remaining eight classes received "regular" comprehension based English lessons, mostly focusing on communication and also including a minimal amount of reading and writing skills. All subjects were followed for two years and tested at the end of grade three and four using an elicited imitation task (repeating six simple sentences with four to nine syllables in length). 20 native speakers of English transcribed the sentences and graded them with regards to accentedness, comprehensibility and fluency. The analysis was carried out using repetition accuracy scores in which the total number of words in the prompt was divided by the number of words correctly produced by the subjects. The results did not show a difference between the two groups at year three, however, the listener ratings of comprehensibility and fluency showed higher proficiency for the regular course in year four.

In her study, Tsiartsioni (2010) examines the production and acquisition of English L2 speech rhythm within a formal school setting. It included 72 native speakers of Greek, aged 10- , 13- and 16-years. From each age group half of the learners were enrolled in the training programme (50 10 to 15 minutes lessons), which were embedded in their regular English classes at school. The control group received the same amount of teaching time in their regular class sessions. The training programme presented by the researcher focused on the pronunciation of English stops, speech rhythm (word and sentence stress) as well as reduced vs. stressed patterns. Similar to Lima (see above), Tsiartsioni applied Celce-Murcia et al's (1997) communicative framework to teach pronunciation as a teaching method. Participants' speech samples comprised of reading-aloud texts which were recorded before and after the intervention. As the perception of speech rhythm depends on the differences in the variability between the duration in vocalic and consonantal intervals, the Pairwise Variablity Index (PVI) was used to evaluate the data. Therefore, the duration of 125 vocalic and 125

intervocalic intervals were measured with the acoustic speech analysis programme Praat. The analyses showed a change in the PVI values towards the native target in the intervention group with regards to the sounds taught in the intervention. However, the results were not always statistically significant.

The above summary of the studies highlights the heterogeneity of research approaches applied to study the efficacy of different ways to teach English pronunciation. In addition, with reference to the criteria outlined by Thomson and Derwing 2014), several limitations of the studies can be identified. For example, apart from Cardoso's study, which offered an extra-curricular course to learn an invented language, all of the other presented studies were conducted in a foreign language classroom demonstrating a high ecological validity. However, the presented training programmes appear to be stand-alone isolated pronunciation tasks, as none of them actually mentions how the applied interventions can be integrated in the foreign language curriculum of the researched participants. In order to verify that improvement is a result of instruction, all presented studies employ a control group design but only Lima (2010) and Trofimovich et al. (2009) apply a delayed post-test to determine whether the intervention had a long-term effect. None of the presented studies included complementary qualitive analyses to provide insights in the learning process. To assess the learners' pronunciation ability, most of the studies (Cardoso 2010; Chen and Goswami 2011; Kennedy 2003; Lima 2010; Tsiartsioni 2010) used reading aloud tasks in contrast to spontaneous speech. Although Chen and Goswami (2011), Lima (2010) and Tsiartsioni (2010) provide some general information on the applied methods or activities implemented in the training programme, only Kennedy (2003) provides enough detail about the procedures to allow replication. Apart from Lima (2010), who presents decreasing error occurrence numbers as a result of the training programme but does not provide any statistical tests, none of the presented studies could show statistically significant evidence of improved intelligibility and comprehensibility as a result of the pronunciation intervention. However, there was usually an increase in performance for both groups over time that could be assigned to ongoing foreign language learning impact and maturation.

These presented challenges of classroom-based intervention might be the reason why most pronunciation intervention studies focus on adult learners outside a classrooms context (for an overview see Table 10-2, Table 10-3, Table 10-4 and Table 10-5 in the Appendix, p. 200f.) and the training scope also varies greatly among those

studies. With communicative competence being the main goal of 2nd language teaching (see Chapter 1 and Section 2.1), intelligibility plays a major role. Therefore it does not come as a surprise that a lot of studies focus on linking words and suprasegemental elements, e.g. (Henderson, 2008; Hincks & Edlund, 2009; Ingels, 2011; Perlmutter, 1989; Sardegna, 2011). However, research on the development of L2 suprasegementals is scarce (Pickering 2012). This might be the reason why most of the time segments (see Table 10-2) are at the linguistic focus of the pronunciation studies (Elliott, 1995; Garcia, 2005; Huthaily, 2008; Liu & Fu, 2011; Warsi, 2001) as there is a huge amount of research and lab studies that focus on segments in the field of second language acquisition which can be used as a theoretical framework (Chen & Goswami, 2011; Derwing & Munro, 2015, p. 110; Lima, 2010). Very often salient segments chosen on the basis of contrastive analysis are at the core of the investigation (e.g. Chen & Goswami, 2011; Huthaily, 2008; Warsi, 2001).

To evaluate the pronunciation performance there is a variety of test instruments. Very often rating scales (see Section 7.5.1) are used to validate the auditory impression (see Counselman, 2010; Munro & Derwing, 2008; Perlmutter, 1989). Despite its rather subjective nature, the huge advantage of this kind of assessment is that even larger amounts of data can be analysed and in particular that it focuses on the intelligibility and comprehensibility of the language which reflects the purpose of pronunciation teaching (see Section 2.1; Derwing & Munro, 2015). In order to get objective measurements, some research looks at acoustic cues such as voice onset time, formant frequencies, pitch and duration (Counselman, 2010; Hicks & Edlund,2009, Suarez, 2013; see Section 7.5.2). Probably due to the immense amount of possible acoustic data points and the time-consuming evaluation most of the studies that use acoustic analysis, focus on a modest number of participants and a manageable training scope (see Table 10-5).

A lot of studies have provided information that pronunciation instruction is successful and can improve learners' productions (see Section 3.6.2; Derwing & Rossiter, 2003; Derwing & Munro, 2015; Flege, Frieda, & Nozawa, 1997; Flege et al., 1995; Missaglia, 1999; Missaglia & Sendelmeier, 1999; Piske et al., 2001a; Wode, 1993). In addition, it can even surpass the effects of just being exposed to the target language (see Section 3.6.1). Some studies (Couper, 2011; Ingels, 2011; Nagamine, 2011; Sardegna, 2011) have shown that awareness training can increase pronunciation learning. However, few studies clearly address the theoretical paradigm that forms the

basis for the interventions (see Table 10-3) or have even linked the outcomes to the pedagogical concept of the training programmes. This does not only make it difficult to evaluate the impact of the intervention itself and poses a gap in the research but it is also unsatisfactory for second language teachers who are in need of successful pronunciation concepts. It is for this reason that the following sections look on research in 2nd language pronunciation in order to compile a suitable concept for the pronunciation intervention that forms the basis of this thesis.

3.2 Theories of second language pronunciation

This section aims to give a brief consideration to the most common second language theories in order to establish a basis for the pronunciation intervention. Learning a second language means that there is already the linguistic system of the L1 in place. On this basis, the influence of the L1 on the learner's L2 is still today one of the most relevant issues in the study of second language acquisition (Gut, 2009).

Research into the study of L2 phonology can be divided into two areas. The first one explains pronunciation problems with the help of two linguistic systems in place; the native language (NL) and the target language (TL) (Eckman, 2012). The second area takes the postulation of the interlanguage hypothesis (IL) into account. This means that learners create their own mental system of the target language "that enables them to produce and understand utterances of the target language (Eckman, 2012, p. 94)."From the 1950's onwards, `Contrastive Analysis´ (CA) has tried to predict language learners' performances with regards to learning new segments (Archibald & Young-Scholten, 2003) and to explain the nature of L2 accents (Pickering, 2012). In his book `Linguistics Across Cultures: Applied Linguistics for Language Teachers´ Lado (cf. 1957) postulates the `Contrastive Analysis Hypothesis´ (CAH) in which he claims that the language learner's mistakes are due to transfer from the L1 to the L2. In this context, pronunciation errors could be explained by the comparison between the native language and target language phoneme inventories. In the 1970's the `creative constructionists´ criticised the idea of L1 language transfer with regards to the acquisition of morpho-syntax as many errors predicted by the CA could not be observed in the learner's L2 language and, moreover, learners produced some uniform errors regardless of their L1. They argued that cognitive processes generating patterns of morphology and syntax could explain these errors. Nevertheless, with regards to L2

phonology, errors were still explained by automatic processes such as L1 transfer, articulatory problems and perceptual filtering (see Section 3.3.1; Archibald & Young-Scholten, 2003). In addition to Lado's Contrastive Analysis Hypothesis, which tries to explain the learning difficulties only as a result of transfer, Eckman (cf. 1977) introduces the idea of `markedness´ to L1 transfer. The `Markedness Differential Hypothesis´ (MDH) describes the asymmetrical relationship between two propositions. Eckman argued that in a marked/ unmarked relationship, one element is more basic and dominant (e.g. certain) than the other. This element is described as "unmarked", whereas the "marked" term refers to a segment that can be deduced from the unmarked form (e.g. uncertain) and is rather more complex. In this asymmetrical relationship, if the marked proposition is true, then the unmarked proposition is true as well. However, if the unmarked proposition is true, nothing can be deduced about the truth of the marked proposition (Young, 2011). Eckman claims that only those L2 segments that differ from the L1 are difficult to acquire if they are more marked than the L1 segment (Eckman, 2008). However, there are some L2 patterns in which the structure adheres to the markedness principle, but which were independent of the L1 and L2 and thus cannot be explained by the MDH. So, with the development of the interlanguage hypothesis in 1984, Eckman puts forward the `Structural Conformity Hypothesis´ (SCH) (amended in 1991) in which he eliminates the difference between the native language and the target language as a criterion for markedness (Eckman, 2008, 2012). The term language universals (U) is used differently in the literature. In the Chomskian way, language universals consist of a whole set of universal principles, including `universal grammar´ (UG). They are defined by an innate language module that operates in all human languages (Carr, 2008; Major, 2001). According to Prince & Smolensky (2002, p. 2) "Universal Grammar consists largely of a set of constraints on representational well-formedness, out of which individual grammars are constructed." In the 1970's a closer investigation of the learner language through error analysis began and there was a shift from the Contrastive Analysis towards the recognition that the learner creates his or her own version of the target language, the so-called `interlanguage´ (IL). An interlanguage refers to a mental system that allows the learner to produce and understand utterances of the target language and it can be seen as a stable, transitional grammar on its own (Eckman, 2012; Pickering, 2012; cf. Selinker, 1972). Interlanguage systems are constrained by general linguistic principles, which interact with the native language phonology (Eckman, 2012, p. 96). In contrast to the Contrastive Analysis, it

also allows for novel structures which are neither present in the L1 or the L2. Major's Ontogeny Phylogeny Model (OPM) postulates that the interlanguage consists of three components, the L1, the L2 and language universals, that are not already part of either the L1 or L2 (2001). The impact of each of the three components changes throughout the different language acquisition stages. In the beginning of the language learning process there is no L2 and the language universals are dormant. With the increase of L2 input the L1 decreases, and the amount of the universals increases as well. Finally, when the L2 is completely mastered the language universals are not needed any longer (Major, 2001).

The presented theories above are commonly used to explain second language acquisition with regards to pronunciation, but there is no theory that is completely able to account for all encompassing language learning processes (Archibald & Young-Scholten, 2003). Moreover, the perception and production abilities of the learner play a major role in the acquisition of L2 pronunciation. Therefore, the following sections take a closer look at the theories of speech perception and speech production (see Section 3.3).

3.3 Theories of speech perception and production

Traditionally, speech perception was looked at as an independent process and researchers did not take the muscular activity of the articulators into account (Hayward, 2000). However, language learners do not only have to learn to perceive fine phonemic differences that might contradict the rules of their L1, they also have to learn how to articulate new sounds and sound sequences (Catford, 1988; Moyer, 2013). In this regard, the following paragraphs first present theories of speech perception, before having a closer look at speech production and the underlying motor-control processes as well as mirror neurons. Taking these findings into account, the theoretical paradigm of pronunciation intervention is developed and presented in the next section 3.4.

3.3.1 Cross-language speech perception

Every language differs in its sound system and it often seems that reaching native-like competence in L2 pronunciation is nearly unattainable (cf. Flege, 1995) due to the constraints of perceived similarities between non-native sounds and native categories

(Best, 1993; Cleary & Pisoni, 2001). Learning a foreign language means to encounter the new sound inventory of the target language. Not only does the production of a new sound cause problems, but, in particular, problems might arise from the learner's inaccurate perception of an L2 sound (cf. Flege, 1995). Research has shown that the inability to correctly hear the L2 sounds is strongly influenced by the internal structure of maternal language categories (Bosch, Costa, & Sebastián-Gallés, 2000).

Newborns have the ability to learn any of the thousand different languages in the world (Gopnik, Kuhl, & Meltzoff, 2001, p. 127) and can identify any speech sound independent of language or speaker. After about six months of age babies start to lose this ability and at about one year they can only differentiate between the sounds of their own language (Best, 1993; Gopnik et al., 2001). It seems that after having been exposed to their native language/s babies start to develop prototypes of the sounds of their language/s. On the basis of these prototypes, babies filter the speech input and develop sound categories, build up the sound inventory of their native language and start to lose the ability to discriminate between the sound differences of other languages (Gopnik et al., 2001; Hardison, 2012).

Getting competent in the native language/s, therefore means losing the openness for non-native languages. In addition, mature learners show difficulties in discriminating many non-native speech contrasts that are not part of their L1 sound inventory (Avery & Ehrlich, 2002; Best, 1993, Best & Tyler, 2007). It seems that the experience gained as a young infant constitutes the groundwork for the adult speech perception (Cleary & Pisoni, 2001, p. 511).

According to Strange (1995, p. 4f.) there is no

> "one-to-one correspondence between phonemes as perceived and the acoustic patterns generated by speech gestures that constitute the stimuli for speech perceptions. Thus, many physically different acoustic patterns may be categorised as the same phoneme (many-to-one correspondence)."

Several theories try to account for cross-language speech perception. Two of the most frequently mentioned models are Best's `Perceptual Assimilation Model´ (PAM) and Flege's `Speech Learning Model´ (SLM) (Bosch et al., 2000). In contrast to the contrastive analysis hypothesis which concentrates on the differences between the native and target language, Flege's SLM (1995) and Best's PAM (1995) argue that the reasons for pronunciation problems lie in the similarities between the native language and the target language (Eckman, 2012; Mayr, 2005). Whereas the PAM concentrates

on perception patterns of non-native contrasts at the initial state of adult L2 learners, the SLM attempts "to predict perception and production patterns as they change with experience with the L2 (Strange et al., 1998, p. 313)."

In her PAM, Best (cf. 1995) describes possible cross-language category assimilation patterns and predicts their consequences.

"According to the PAM, when perceiving non-native segments, listeners can

- assimilate them to a native category (either as good or as bad examplars of it)
- perceive them as uncategoriseable speech sounds
- perceive the segments as non-speech (noise)

Depending on which type of treatment the L2 sound has received, L2 listeners degree of difficulty in perceiving it will vary" (Strange et al., 1998, p. 313f.).

Flege introduces the term `equivalence classification´ proposing that L2 learners have already established phonetic categories in their L1's and are likely to assimilate perceptually similar but phonetically distinct L2 phones into the same category of those of the native category (Pickering, 2012, p. 336). Flege argues: "The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned" (1995, p. 239). This hypothesis postulates that misperception leads to the accented inaccurate production and that changes in perception also results in the alteration of the production. However, not all inaccurate production is necessarily a result of missed perceptual cues.

Building on his idea that the L1-L2 similarity of sounds rather than the difference leads to problems in speech acquisition (cf. Flege, 1987), and according to the aforementioned assumptions, Flege formulates the `Speech Learning Model´ (SLM), which he describes as follows (1995, p. 239):

> "An assumption we make is that the phonetic systems used in the production and perception of vowels and consonants remain adaptive over the life span, and that phonetic systems reorganize in response to the sounds encountered in an L2 through the addition of new categories, or through the modification of old ones."

With more experience the learner will be able to identify the distinct L2 phonetic segments and be able to produce and perceive them more accurately (cf. Flege, 1995).

However, Flege's SLM approach focuses on the production of segments[21] and does not refer to suprasegmental features.

In 1995, Kuhl and Iverson introduce the concept of the `native language magnet´ showing that the exposure to a specific L1 early in life "results in a distortion of the perceived distances between stimuli; in a sense, language experience *warps* the acoustic space underlying phonetic perception (Kuhl & Iverson, 1995, p. 121f.)." On this account, L2 sounds that are similar to a specific L1 sounds, are attracted by the L1 magnet and the learner cannot detect any perceptual difference between the two sounds (Best & Tyler, 2007; Hardison, 2012, p. 349).

The presented theories of speech perception show that L2 sounds, which are similar to a L1 sound, might be assimilated to a native L1 category and that sounds which are different enough might be perceived as a new category. However, in order to acquire a native-like pronunciation central perceptual representations for all of the L2 sounds need to be established. On this basis, pronunciation training has to focus on developing these perceptual representations. Yet, learning L2 pronunciation does not only involve speech perception, but also requires the ability to develop motor routines to produce the new physically different phones (Flege, 1997). Thus, the following section addresses the motor theory of speech perception.

3.3.2 Motor theory of speech perception and production

In order to produce intelligible speech, rapid muscular movements have to be carried out to coordinate the respiratory movements as well as the articulation of the glottis, pharynx, velum, jar, tongue, and lips (Pickett, 1999). In the 1950's, Alvin Liebermann and Franklin Cooper originally proposed the motor theory of speech perception. They argued that speakers transfer their linguistic intentions into a series of speech movements of the articulatory tract, which are called `articulatory gestures´. According to the `motor theory´, there is a neural representation of which sound belongs to which `articulatory gesture´. In a similar fashion, the perceived acoustic signals are decoded by articulatory representations that synthesize the vocal tract shapes needed to produce a given speech pattern (Eckman, 2012; Pickering, 2012; cf. Selinker, 1972) instead of identifying the sound patterns of speech. The detection and production of

[21] Research on the development of L2 suprasegmentals is scarce (Pickering, 2012).

sounds is driven by the motor system in the central nervous system. Learning a new language means that the learner becomes able to "read" the articulatory gestures of an L2 speaker and that he also learns how to properly use the musculature of his mouth to produce a certain L2 sound (Avery & Ehrlich, 2002).

The `motor theory´ hypothesized that the processes to decode speech differ from those used to perceive non-speech acoustic signals (Strange, 1995). The `motor theory´ is also supported by the following findings: MacLeod and Summerfield showed that speech perception in noise increases if the subjects were able to see the speaker (cf. 1987). In their paper "Listening with hand and eye", Fowler and Dekle (cf. 1991) found out that the perception of syllables improved for those participants who could haptically feel the syllable production. Moreover, the `McGurk effect´ shows that visual speech information strongly influences auditory cues (cf. McGurk & MacDonald, 1976).

3.3.3 Mirror neurons

In 1992 di Pellegrino et al. (1992:176) recorded that neurons in the inferior motor cortex of macaque monkeys not only discharged during goal-directed hand movements but were also activated when observing these meaningful hand movements carried out by the experimenters (Fogassi and Ferrari 2009:348). Accordingly, the neurons were called "mirror neurons" as they respond not only to performed actions but also to observed actions 3.3.1.

Brain imaging experiments with humans showed that Broca's area is also activated when observing hand and mouth movements confirming the existence of mirror neurons for action understanding in humans (Fogassi and Ferrari 2009:348). Due to this "mirror effect" persons are able to understand the actions of others, as they are able to internally simulate the action for themselves. This also holds true for language perception and the re-production of articulatory gestures (Bauer, 2008, p. 76). Research showed that when individuals observe biting actions or other individuals performing silent speech the inferior frontal gyrus is activated (Fogassi and Ferrari 2009:348). With regard to pronunciation teaching the existence of mirror neurons means that learners might profit from seeing the production of new sounds in order to establish new perceptual representations for themselves (see Section 3.3.1). Hence, with the discovery of these mirror neurons, the motor theory of speech (see Section 3.3.2) gained renewed interest.

3.4 The theoretical paradigm of the pronunciation intervention

From the discussion of the previous sections of this chapter, several principal issues and suggestions with regards to second language pronunciation have arisen. The key aspects are summarised in the following paragraph (see Figure 3-1):

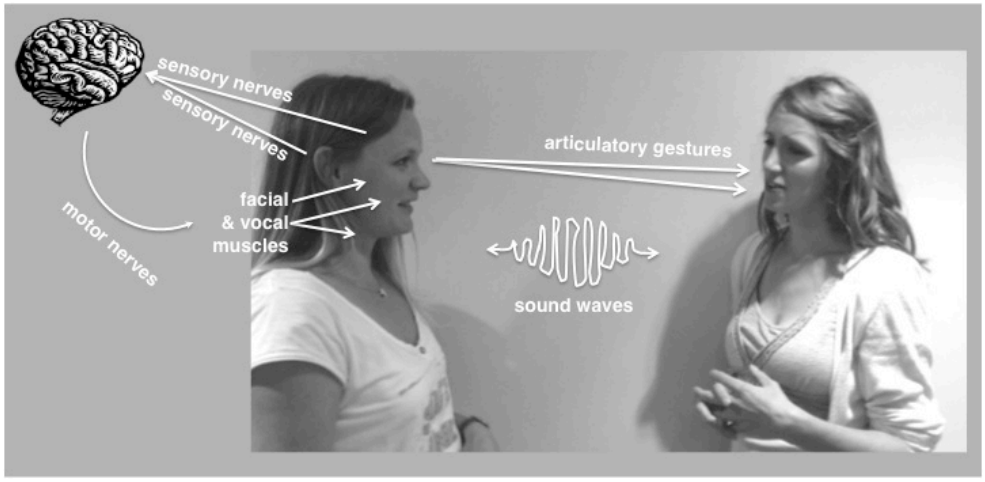


Figure 3-1: Key aspects of pronunciation learning (adapted from Denes & Pinson, 1993)

Each segment in the native language has neural representations of a sound target, which includes the oral perception of the sound as well as the articulatory gestures needed to produce a specific speech sound (cf. Eckman, 2012, p. 94f.). For language to be maximally efficient and to facilitate the intelligibility between different speakers, e.g. children, males and females with different vocal tract sizes as well as between speakers of different dialects, the concept of a language magnet provides a plausible framework. In this regard, small discrepancies in sound production and perception are not detected as the language experience warps the acoustic space of a sound target (Hardison, 2012; Kuhl & Iverson, 1995) and intelligible communication can take place.

Second language sounds, that are different enough, are perceived as a new sound category. Nonetheless, the novel articulatory representations and gestures have to be acquired. If this learning process does not take place, the novel sounds might be substituted with other sounds, such as the English /θ/, which is frequently pronounced as an /s/ or /z/ by German speakers of English despite the fact that they are usually able to discriminate between the two sounds. However, L2 sounds that are similar to a native category might be assimilated to the latter. In this case, the pull of the L1 language magnet has to be broken and the learner must be made aware of the

existence of the new sound category in order to acquire it competently. Regardless of the sound being similar or different, the learner has to be able to decode and produce the L2 articulatory gestures driven by the motor system in the central nervous system (Avery & Ehrlich, 2002). To facilitate this process, according to the notion of mirror neurons learners should focus on the articulatory gestures of a model L2 speaker to adopt these gestures for themselves.

The presented concept builds the theoretical paradigm of the pronunciation intervention (for more information see Chapter 4) in which the students are first made aware, see and haptically feel a new L2 target before they are asked to identify and discriminate it from similar or different sounds. As the training programme can not encompass all pronunciation areas and is set out to be a generic part of the English lesson, the following Section 3.5 provides a rationale for the choice of the intervention sound core.

3.5 The intervention sound core

It is a primary concern of the pronunciation intervention, that it can be easily integrated into English lessons at a beginner level over the course of one school term (about five months) without taking up too much teaching time. On this basis, achievable goals have to be set and specific pronunciation areas need to be selected. As already indicated in the introduction to this thesis (see Chapter 1), the author regards the segmental as well as suprasegemental areas of pronunciation as equally important. However, due to the results of the curricula and textbook analyses (see Section 2.2 and Section 2.3), teacher questionnaire (see Section 2.4) and the presented theories of speech perception and production which mainly focus on segments, a sound core is selected that is already part of the students' curriculum (see Section 2.2.) and thus ensures the ecological validity of the study. In addition, the choice should also allow to test the hypothesis whether different or similar sounds are harder to learn.

Learners that have already mastered their L1 are bound to draw comparisons between the L2 and L1 (see Section 3.2). Learning a new language means that they need to learn to perceive and produce differences in sounds, in terms of quality, duration and allophonic realisation. Moreover, learners need to learn about distributional differences in sound categories of the target language (Moyer, 2013) and their articulatory production, e.g. tongue height and placement and lip movement.

Taking the theoretical framework presented in the previous section into account, sounds or patterns in the target language that are different to the L1 will be easily perceived but still need practice in mastering them. However, it can be assumed that language learners will confuse target sounds that are similar in the target language with sounds of their L1 despite detectable acoustic differences (Flege, 1997), and they might not be able to notice the differences without any feedback from a teacher. Although English and German share some similar phonological patterns due to a shared historical relationship, there are many differences as well (Smith & Peterson, 2012). Therefore, sounds as well as phonological patterns that are both similar or very different between English and German are included in the pronunciation intervention.

3.5.1 Vowels (Monophthongs)

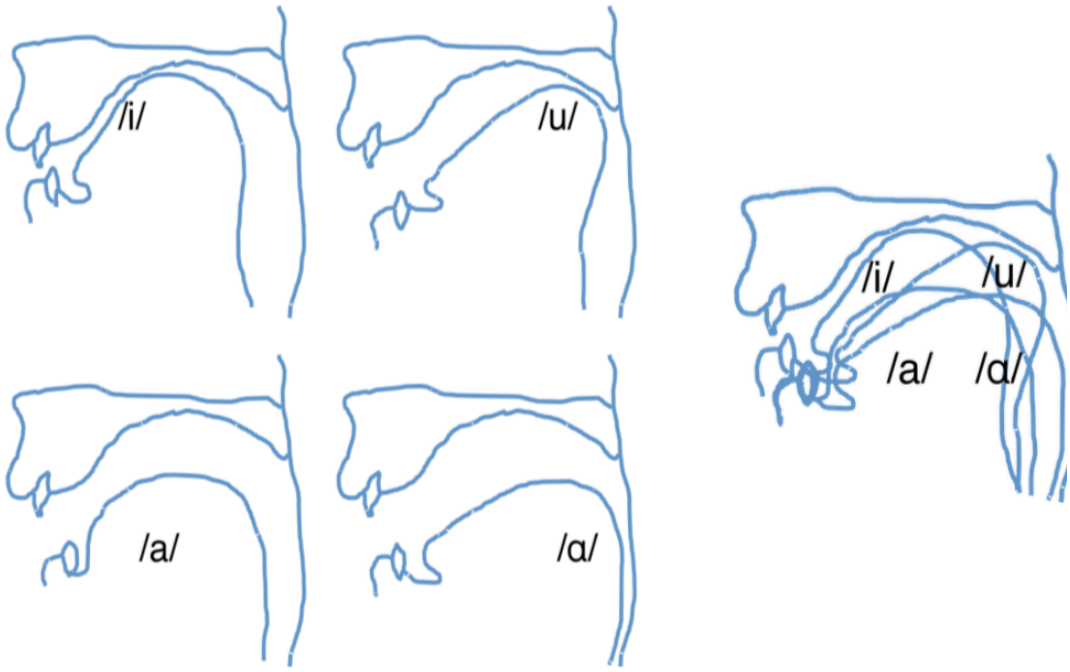
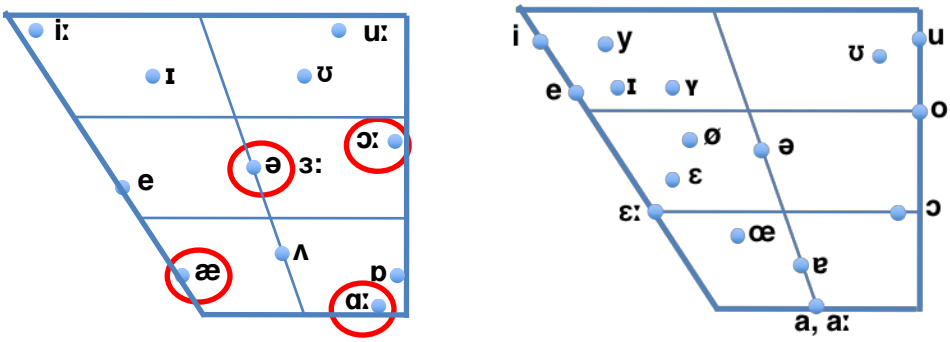


Figure 3-2: Four peripheral vowels produced by different tongue positions displayed in the mid-sagittal section of the vowel tract (adapted from the International Phonetic Association, 1999)

Vowels are produced by the resonances of the vocal tract shaped by various tongue positions. Therefore, they can be described in terms of an abstract vowel space (the `Vowel Quadrilateral´ (see Figure 3-3 and Figure 3-4), which bears relation to the four extreme positions of the tongue in vowel production (see Figure 3-2). In case of a fronted and high position of the tongue, the oral cavity is rather closed and the front closed vowel /i/ is produced. Similarly, the open and backed position of the tongue

narrowing the pharynx, leads to the production of the open back vowel /ɑ/. If the mouth is as open as possible and the tongue is fronted, the front open vowel /a/ is produced. Lip rounding and constriction results in the production of the close back vowel /u/. These four peripheral vowels provide the space in which the other vowels can be produced (International Phonetic Association, 1999).



English German

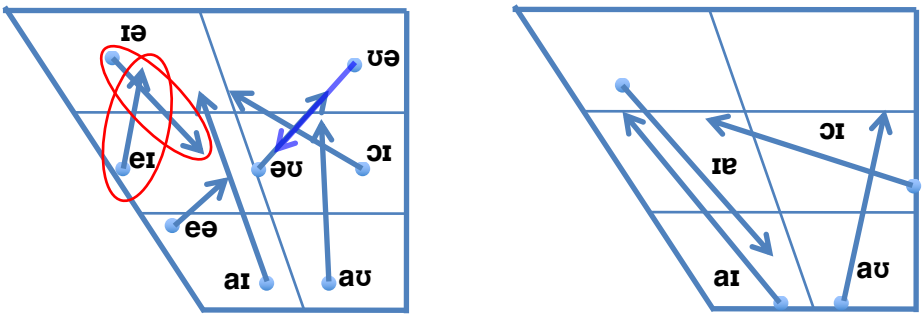
Figure 3-3: IPA vowel chart for English Received Pronunciation (adapted from Wikimedia Commons, 2008b) and German monophthongs (adapted from International Phonetic Association, 1999)

A look at the IPA vowel spaces (see Figure 3-3) for English (RP) and German (educated speakers in the North) reveals several differences between the vowel positions of the two languages. It is evident, that the extreme open front and open back positions of the German vowel space are not occupied. As discussed above, pronunciation errors are often explained by the sound differences between the languages (see Section 3.3.1). Moreover, the theories of speech perception and motor control state that although new perceptual categories might be easily spotted, there is still some learning and practice required to automate the precise articulatory gestures. It is for these reasons, that the front open vowel /æ/ and back open vowel /ɑː/ which are not part of the German sound inventory are included in the intervention. However, the presented theories of speech perception and production such as Flege's `Speech Learning Model´ (SLM) propose that sounds that are similar but not the same are hardest to pronounce (see Section 3.3). Thus /ɔː/ is also included in the sound core because the English counterpart of the German open-mid vowel /ɔ/ is pronounced as a long vowel in a rather more closed-mid position (/ɔː/). Moreover, these vowels are also targeted in the students' textbook (see Section 2.3). Although other vowels would have served the same purpose, they are not included due to the limited scale of the intervention. In order to have a reference point for the vowel space, the neutral vowel

/ə/ was also chosen as part of the intervention as the neutral vowel /ə/ is shared in both languages. In summary, the following English vowel sounds were included in the vowel core:

- Long Vowels: /ɑ:/, /ɔ:/
- Short Vowels: /æ/, /ə/

3.5.2 Diphthongs



English German[22]

Figure 3-4: IPA diphthong chart for English Received Pronunciation (adapted from Wikimedia Commons, 2008a) and German diphthongs (adapted from International Phonetic Association, 1999)

In German, a diphthong ending with an /ɐ/ can be formed with every vowel except /ə/ and /ɐ/. For a better visibility those diphthongs are not included in the vowel space presented in Figure 3-4 with the exception of /ɪɐ/ as the English diphthong /ɪə/ shows a similar direction of movement to the German /ɪɐ/. However, the German diphthong displays a longer trajectory length, and ends in a near open central vowel in contrast to the central schwa of the English diphthong /ɪə/.

Comparing the vowel spaces of the English and German diphthongs, it becomes apparent that all of the German diphthongs tend to have quite long trajectory lengths compared to their English counterparts. In addition, there are no German diphthongs similar to /eɪ/ that begin in a close-mid front position and whose trajectory ends in a close front position. For these reasons, the above mentioned /ɪə/ and /eɪ/ and were chosen due to their places in the vowel space and their short trajectory lengths. Moreover, /eɪ/ is also targeted in the students' textbook (see Section 2.3). Although

[22] The diphthong /ʊɪ/ as in `Pfui´ is also missing from the vowel chart, as it only plays a marginal role (Mangold, 2005). Moreover, German diphthongs, which are a typical marker for foreign words, such as /oa/ in `Croissant´, are also not included.

other diphthongs might have been equally suitable, they were not included due to the limited timescale of the pronunciation intervention.

3.5.3 Consonants

Consonants are shaped by a narrowing or stricture of the vocal tract which is typical for each consonantal group. Therefore, they can be identified by their place and manner of articulation, and also by whether they are voiced or voiceless. Table 3-1 shows a comparison of the distribution of the English and German consonant inventories.

Table 3-1: German and English consonant phonology (adapted from Fox 2000)

		German	English
Initial consonants	plosives	p, b, t, d, k, g	p, b, t, d, k, g
	nasals	m, n	m, n
	fricatives	f,v s,zʃ ʁ h	θ,ð,f,v,s,z,ʃ,ʒ,h
	approximants	j	w, j
	laterals	l	l, r
	affricates	ts, pf	tʃ, dʒ
Inter-syllabic consonants	plosives	p, b, t, d, k, g	p, b, t, d, k, g
	nasals	m, n, ŋ	m, n, ŋ
	fricatives	f,v s,z,ʃ, ç, x, ʁ, h	θ,ð,f,v,s,z,ʃ,ʒ
	laterals	l	l, r
	affricates	ts, pf	tʃ, dʒ
Final consonants	plosives	p, t, k	p, b, t, d, k, g
	nasals	m, n, ŋ	m, n, ŋ
	fricatives	f, s, ʃ, ç, x	θ,ð,f,v,s,z,ʃ,ʒ
	laterals	l	l, r
	affricates	ts, pf	tʃ, dʒ

Most learners of English do not have any problems in pronouncing the plosives /b/, /d/, and /g/ correctly, but especially German learners of English typically encounter difficulties with the voiced plosives in word final position (Quetz, 1998). While in German final devoicing is a highly regular process, this phonological pattern does not exist in English (König & Gast, 2007; Smith & Peterson, 2012), and thus /b/, /d/ and /g/ do not appear in word final position (see Table 3-1). As a consequence, German learners of English tend to devoice voiced plosives (as well as fricatives and affricates) in word-final position (Kelly, 2000; Weinberger, 1997). In addition, /b/, /d/ and /g/ are targeted in the corresponding school curriculum (see Section 2.2.) as well as in the related textbooks (see Table 2.2). It is for this reason that the English plosives /b, d, g/ in final position were included in the pronunciation intervention of the current study. The German sound inventory lacks the interdental fricatives /θ/ and /ð/ and they are often substituted by /s/ and /z/ (Avery & Ehrlich, 2002; Kelly, 2000; Kenworthy, 1987; Weinberger, 1997) which is less effortful (Weinberger, 1997). As a result, the

mispronunciation of the `th´-sounds is probably the most prominent feature of the accent of a German learner of English. Probably due to this reason, they are identified as one of the major pronunciation difficulties in the teacher questionnaire (Table 2.12.) and are included in the curricula and textbooks. On this basis, /z/, /θ/ and /ð/ are also chosen as a part of the sounds core for the pronunciation intervention. Similarly, the affricate /dʒ/ does not exist in the German language and is often substituted with its voiceless counterpart /tʃ/ (Avery & Ehrlich, 2002; Kelly, 2000) and it is also targeted in the students' textbook. Thus, it is also included in the pronunciation intervention. A typical German accent often includes the mispronunciation of the /w/ because German learners of English often replace the English approximant /w/ with the German /v/. This might be due to the fact that /w/ does not equal its German counterpart (Haß, 2006; Piepho, 1974). Moreover, curricula and textbooks also target the /w/. For these reasons, the approximant /w/ is also chosen for the pronunciation intervention in the current study.

Taking the discussion above into account, the selection of the sound core was carried out according to prominence in foreign accent, the difference and similarity of sounds or phonological patterns and due to reasons of ecological validity. The author is aware, that other sounds could have served the same purpose, and that the choice was partly arbitrary. However, a selection had to be made due to the limited timescope of the intervention. In summary, the following sounds were included in the intervention study:

- Monophthongs: /ɑ:/, /ɔ:/, /æ/, /ə/
- Diphthongs: /eɪ/, /ɪə/
- Consonants:
    - Word-final plosives /b/, /d/, /g/
    - Fricatives: /θ/, /ð/, /z/
    - Voiced affricate: /dʒ/
    - Approximant: /w/

So far, this chapter has looked at the linguistic factors that influence second language pronunciation. However, pronunciation acquisition also depends upon individual factors, which are presented in the following section.

3.6 Other factors influencing second language pronunciation

There are a number of individual learner variables that influence the degree of foreign accent. These include age of L2 learning, gender, formal instruction, motivation, and psychosocial factors (Piske et al., 2001a). The following section looks at these factors with regard to the research project at hand. However, age of L2 learning seems to be the most important predictor of the degree of foreign language accent and the relative impact of the other variables often remains uncertain as the data is often confounded due to a lack of experimental control (Moyer, 2013; Pickering, 2012; Piske et al., 2001a) and due to the fact that all individual factors interact with each other as well (Edmondson, 1999). Along the same line, one has to bear in mind that age factors are also confounded by cognitive and social factors (Moyer, 2011).

3.6.1 Age of L2 learning

The assumption that adult L2 learners are biologically less capable of learning a native-like accent in their L2 due to their declining cognitive plasticity is one of the most widely accepted truisms about language learning (Flege et al., 2003; Pickering, 2012) and in fact, research into the overall degree of foreign accent of L2 non-native speakers shows strong effects of age (Bosch et al., 2000; Flege et al., 1995; Flege et al. 2003; Wode, 1993).

In 1967 Lenneberg (1967) stated that L2 acquisition after puberty needs a very high effort and that it is hardly possible to acquire a native-like pronunciation. He argues that the reasons for the declining phonological abilities might be due to the brain maturation and the increasing inflexibility of neuromuscular processes.

The `critical period hypothesis´ (CPH) proposes a developmental period after which a native-like language acquisition is impossible (Lenneberg, 1967; Scovel, 2000). With regard to pronunciation, the age threshold for the CPH ranges - depending on the respective research - from five to 15 (Birdsong, 1999; Pickering, 2012). The results of a significant number of L2 foreign language studies researching the age of arrival in the L2 target country support the view of “the earlier, the better” (Flege & Fletcher, 1992; Flege et al., 1995; Moyer, 1999; Patkowski, 1990). However, no sharp discontinuities can be found at a certain age. Therefore, Oyama (1976) and Long (1990) suggested a

`sensitive period´ instead of a `critical period´ for L2 learning (Piske et al., 2001a, p. 196).

During the early 1980's, Flege and his colleagues challenged the presumption that post-pubescent L2 learners would maintain accented speech and questioned the role of CPH as the main factor explaining the differences between children's and adult's production of L2 phonetic segments (Flege, 1981). The differences in pronunciation performance might be the result of a number of other confounding factors apart from, or in addition to the CPH, such as previous linguistic experience, motivation, formal instruction or social factors (Pickering, 2012). All in all, research results with regards to the CPH remains ambiguous (Flege et al., 2003; Singleton, 2001; Wode, 1993).

Nevertheless, there is agreement on the following substantial findings (Edmondson, 1999; Krashen, Long, & Scarcella, Robin, 1979):

- Children and teenagers seem to have better imitation abilities than adults. Imitation abilities play a major role in the acquisition of L2 pronunciation;
- Learners who have access to the L2 as children or teenagers are very likely to have higher pronunciation and supra-segmental skills

Moreover, the neural functions that are involved with the motor control development seem to be very active in childhood but are said to decline with age (Moyer, 2013).

Most age-related research has been carried out on non-native speakers without any formal education in their L2. It was suggested that for these naturalistic L2 learners there is a sensitive or critical period of age as only few if any individuals manage to speak their L2 without a detectable foreign accent (Cruttenden, 2008; Flege et al., 1995). However, there is a strong variation among adult learners regarding the degree of pronunciation accuracy and some are even able to achieve native-like pronunciation (Flege, Frieda, & Nozawa, 1997; Flege et al., 1995; Wode, 1993). These `exceptional learners´ often received formal language instruction and often did not even move to the target country (Gut, 2009). However, it is difficult to compare and to evaluate different pronunciation studies as there is no standardized way of data elicitation, there are different pronunciation tasks and ways to analyse the data in pronunciation research (Moyer, 2013).

A literature review suggests that `length of residence´ (LOR) in a foreign country and `age of arrival´ (AOA) seem to influence L2 accent (Bosch et al., 2000; DeKeyser, 2012; Edmondson, 1999; Flege et al., 2003). However, as the current research project

is carried out in German classrooms, the LOR variable cannot be applied to the research subjects and is therefore not discussed any further in this section (see Piske et al., (2001a) for a review on LOR).

3.6.2 Formal instruction

A number of studies researching the influence of formal instruction with regard to foreign accent do not show significant results (Flege, 1995; Flege, Yeni-Komshian, & Liu, 1999). However, in most studies, `formal instruction´ is measured by the `length of English language instruction´. So, while `length of instruction´ does not seem to be a significant predictor for L2 accent, special pronunciation training in the perception and production of L2 sounds may have an effect on L2 pronunciation accuracy (Derwing & Rossiter, 2003; Missaglia, 1999; Missaglia & Sendelmeier, 1999; Piske et al., 2001a). In their research, Bongaerts et al. (1997) looked at five late English L2 learners who received not-specified intensive pronunciation training in the perception and production of English sounds. In a rating test, their results mirrored the ones of English native speakers. Moyer (1999) researched English learners of German who received training on the segmental and suprasegmental level. Their results were rated very close to the range of German native speaker rankings.

Although the presented results are by no means conclusive, they suggest that pronunciation instruction can have a strong impact on the pronunciation abilities.

3.6.3 Sex

There has been lots of research in how far the speaker sex plays a role in the acquisition of phonological properties of the L2. Contrary to the saying: “Girls are better at languages”, a lot of studies (Elliott, 1995; Park, 2009; Piske et al., 2001a) have revealed no significant effect of sex on the pronunciation proficiency. On this basis, both boys and girls are included in the study.

3.6.4 Motivation

It might be obvious to suggest that learners who are very motivated to achieve a good L2 pronunciation will also be very likely to achieve good results. However, motivation

is not readily observable. The impact of motivation on the degree of L2 accent has been measured in most cases with the help of rating scales, where the subjects had to rate the importance of good L2 pronunciation for their work as well as for their social life. Most of those studies identified motivation as a predictor for L2 pronunciation success (Elliott, 1995; Flege et al., 1995, 1999; Purcell & Suter, 1980). Bongaerts et al. (1997) and Moyer (1999) looked at L2 learners who had a very high personal or professional motivation to achieve native-like pronunciation. However,

> "the results obtained so far clearly suggest that factors like professional motivation, integrative motivation or strength of concern for L2 pronunciation accuracy do not automatically lead to accent-free L2 speech. Apparently, they are rarely so strong that late learners will still be able to attain a native-like pronunciation in the L2 (Piske et al., 2001a, p. 202)."

3.6.5 Psychosocial factors

Several psychosocial factors are suggested to have an influence on the degree of the L2 learner's pronunciation. Amongst others, these are talent, personality traits such as field independence, risk-taking, extroversion and the relationship between the L1 and L2 country (Mayr, 2005).

The `egopermeability model´ (Guiora, Beit-Hallahmi, Brannon, Dull, & Scovel, 1972) means that language acquisition is always accompanied by the transfer of the learner's social L1 identity towards the social identity of the new culture. Engaging and becoming competent in an L2 involves a psychosocial alienation from the identity of the L1 speaker. The willingness to identify with the L2 learner identity declines with age and especially after puberty due to the stabilisation of the L1 identity (Celce-Murcia, 1997; Grotjahn, 1998) Schuhmann's `acculturation model´ (1976) proposes that the learner's social distance to the target country can also account for the success in language learning.

The CPH explains the ease or difficulties in learning due to a shift in neurological processes that take place in puberty but other cognitive developments and affective factors like inhibition and empathy also correlate with age. Language learning cannot take place without cultural learning. Brown (1980, p. 158f.) states that the interaction between the language and the culture reaches a "certain stage during which language learning achieves an optimal distance". In his `optimal distance model´ Brown explains his proposal with the help of four related parameters:

- Stages of Acculturation;
- Anomie;
- Social Distance; and
- Perceived Social Distance.

Brown describes four stages of acculturation when individuals assimilate themselves into a new culture. First, the learner might undergo a period of excitement over the newness of the surrounding. Secondly, the cultural shock about the cultural differences emerges. Thirdly, a recovery period commences until they are finally accepting the new culture and the "self-confidence of the new person that has developed in this culture" (Brown, 1980, p. 159). When adapting to a new culture, the learners begin to lose some ties to their native culture and might experience `anomie´ - a feeling of homelessness. Only until a person has mastered to deal with the `social distance´ (see (Schuhmann, 1976) above, the feeling of anomie decreases. The final stage is reached when learners "see themselves maintaining some distance between themselves and both cultures" (Brown, 1980, p. 161).

With regards to language teaching, teachers have to bear in mind that they are not only teaching the target language itself, but they are also introducing the target culture. Moreover, they need to be sensitive about the fact that some learners might experience unease when they are asked to use sound patterns of the target language that are very different to their L1 inventories, and "put themselves in the shoes" of a native speaker, and try to imitate the foreign language sounds.

3.6.6 Choice of subjects for the pronunciation intervention

It is for the reasons mentioned above, that children who receive formal language instruction were selected as research subjects. However, as these children usually begin to learn English as a foreign language in the third year of primary school and a certain amount of basic language proficiency and pronunciation experience is required to fulfil the tasks of the intervention, pre-pubescent learners at the ages of 10 to 12 with at least two years of English instruction at primary school were selected for the study. In addition, equal numbers of girls and boys were included and all subjects were asked to rate their motivation to learn English as well as their interest in doing pronunciation exercises. To include a variety of subjects, students from a grammar and a comprehensive school and different socio-economic areas of Frankfurt were included in the study (see Section 7.2.4).

3.7 Summary

It was the purpose of this section to give an overview of the theoretical background of second language pronunciation. Therefore, section 3.2 presented theories that have been proposed to explain second language acquisition and looks at the most widely established models of speech perception (see Section 3.3). In addition, the mechanisms underlying speech motor control (see Section 3.3.2) and mirror neurons are outlined (see Section 3.3.3). Drawing from the content at hand, the rationale for the pronunciation treatment and the theoretical paradigm of this study is presented in Section 3.4). Following this, the English and German sound inventories are introduced and the intervention sound core is selected (see Section 3.5). As L2 pronunciation acquisition is also affected by non-linguistic factors (see Section 3.6) the influences of age of learning, gender, language use, formal instruction motivation, socio-economic factors, are discussed. Looking at individual differences in pronunciation, reasons for choosing 10 to 12-year-old language learners at a beginner level as participants for this intervention study are outlined (see Section 3.6.6).

4. The pronunciation intervention

The pronunciation intervention is based on the theoretical framework mapped out in Section 3.4. This chapter specifies the programme's constituent tasks and procedures in sequence together with their rationales. Given the evidence that speech sound awareness training can facilitate pronunciation (Couper, 2011; Ingels, 2011, Nagamine, 2011; Sardegna 2011, see Section 3.1), the intervention seeks to promote speech sound awareness by presenting new sound targets to learners helping them attend to the phonetic features. Each intervention stage tackles a specific sound or group of related sounds and begins with a presentation phase in which new sound characteristics as well as novel phonological patterns are highlighted. Supports include visual stimuli such as face close-ups in videos, models or mirrors and haptic exercises to stimulate the mirror neurons and to learn the perception and motor production of the new articulatory gestures. Given that the L1 language experience of a second language learner influences their ability to detect L2 sounds (Hayes-Harb, 2007, p. 65; Hirschfeld, 2007, p. 277; Neri et al., 2006, p. 358), identification and discrimination skills need to be trained. This includes the ability to perceive differences in similar sounds as well as the decoding of new articulatory gestures (see Section 3.3). Finally, a significant amount of practice is required in order to automate the articulatory gestures and establish the neural representations of new sounds and phonological patterns. Reflecting the outline above, the pronunciation training involves three stages for each of the pronunciation targets:

1. Presentation;
2. Identification and discrimination; and
3. Production.

To help ensure a similar delivery of the pronunciation intervention, a comprehensive teachers' manual was developed (see Appendix, p. 231ff.) and the teachers were provided with the required materials, resources and media. The intervention was conducted in four fifth-grade classes – two classes at a comprehensive school and two classes at a grammar school, all with learners at age 10 to 12 (see Section 7.2.4). Based on a curriculum analysis (see Section 2.2) and a teacher questionnaire (Section 2.4) as well as research on L2 development (Section 3.3) the following sounds were selected for the intervention (Section 3.5) and thus tackled in the training programme:

- Monophthongs: /ɑ:/, /ɔ:/, /æ/, /ə/
- Diphthongs: /eɪ/, /ɪə/
- Word-final plosives /b/, /d/, /g/
- Fricatives: /θ/, /ð/, /z/
- Voiced affricate: /dʒ/
- Approximant: /w/.

The following paragraphs describe the intervention in more detail. It should be noted that the structure of the pronunciation programme was newly set up to accommodate the theoretical framework mapped out in Section 3.4. However, although some tasks were actually novel, other tasks were adapted from existing, well-established pronunciation task formats.

In order to work on the pronunciation tasks, the students need to be introduced to the technical pronunciation terms and phonetic symbols and code (for reference see the teachers' manual in the Appendix, p. 231ff.). Thus, in the introductory phase the students get to know "Andy the Pronunciation Android" on a worksheet. This introduces the new vocabulary in a visual format and for the purpose of listening and matching tasks the students have to "touch their voice box", "push air through the mouth", "spread their lips", etc. Finally, the new words are trained in an activity adapted from the "Simon says" format, e.g. "Andy the Android says put your tongue between your teeth!" (Bowler, 2005, p. 45). Andy also uses an alien language which is written in phonetic code. The students are informed that a phonetic code is used because the same English (or alien) letter can be pronounced in various ways (e.g. **a**rm, b**a**d, sm**a**ll) and that it is often very different to its German counterpart. To learn Andy's language, the students are given a phonetic code chart (see Figure 10-12 in the Appendix, p. 260) which they can use as and when required over the course of the intervention. Although the order can be varied, the teachers' manual starts off with the pronunciation of the vowels and proceeds to the diphthongs, consonants and the approximant. The following descriptions use the same order.

4.1 Monophthongs: /ɑ:/, /ɔ:/, /æ/ and /ə/

In the presentation phase, the students begin with the long vowels. First the teacher displays a poster describing the /ɑ:/ (see Figure 10-7 in the Appendix, p.237) and with the help of their books the students learn that /ɑ:/ is one of the phonetic codes for the letter [a]. A video (BBC, 2010) shows an oversized close-up of the face of a native

speaker modelling the sound /ɑ:/ in isolation and in the context of some reference words, e.g. bath. The students are asked to repeat after the speaker. The teacher tells the students that /ɑ:/ is a long sound, indicated by the colon, and that they must relax their lips and half-open their mouths to produce this sound. The teacher also models and drills the sound and the example words (arm, far, grass, aunt) on the poster and the students have to write one reference word for /ɑ:/ in their own sound chart (see Figure 10-12 in the Appendix, p. 260). The long sound /ɔ:/ is introduced in the same fashion with the difference that to produce /ɔ:/, the lips need to be opened and rounded. Similarly, the short vowels /æ/ and /ə/ are presented with the help of videos, close-ups and drills. However, to produce /æ/ the students are asked to spread their lips wide and open their mouths as if they were eating an apple; to make /ə/ they have to relax their faces and open their mouths a little to produce a sound like Tarzan. To practise the new articulatory gestures, the students use mirrors. First, they produce the sounds in isolation. Then they produce one of the example words for /ɑ:/, /ɔ:/, /æ/ and /ə/ from the posters out loud, and then, they model them silently while watching their mouths in the mirror. They are asked to describe and discuss the movements of their mouth, tongue and teeth with a partner in the classroom. Finally, they have to work with a student peer; one student silently produces an example word while their partner has to guess it, and vice versa.

In the identification and discrimination phase the students are handed two ‘smiley’ cards – one positive and one negative. The teacher reads a list of words containing long and short vowels. The students´ task is to identify the long sounds and every time they hear a long sound they should raise the positive smiley. Similarly, the students are asked to identify each of the “new” sounds from a list of words and required to do the procedure with a student peer.

To practise /ɑ:/, /ɔ:/, /æ/ and /ə/, the posters with the example words are put on the blackboard and the students have to write down one meaningful sentence that contains two instances of each of the four vowels. They are also required to underline the vowels in the sentence and check them and their pronunciation with their partners and teacher. After a short phase in which the teacher introduces and models the Queen’s accent (RP), the students are asked to put on their English shoes and try to make themselves sound like the English Queen. Then they should walk around the classroom and present their sentence to other students who are asked to repeat it. Finally, the funniest sentences are selected and presented in the class.

4.2 Diphthongs: /eɪ/ and /ɪə/

A diphthong is a "vowel sound in which there is a transition from one vowel quality to another within a single syllable nucleus" (Carr, 2008, p. 43). Accordingly, in the sound presentation phase, first each of the vowel elements of the diphthongs /eɪ/ and /ɪə/ are presented to the learners. Both diphthongs contain the short vowel /ɪ/. In order to teach the students how to pronounce this short vowel, highlighting the contrast with the long /i:/ helps the student to understand the articulatory gestures needed. The teacher displays a poster for /ɪ/ (see D1, Appendix, p. 242) and asks the students for words that contain the sound. Examples are collected on the blackboard and the students are reminded that /ɪ/ is a short vowel, as long vowels are indicated with a colon. Then the poster for the long vowel /i:/ (see D2, Appendix, p. 242) is displayed and again students are asked to come up with examples. The teacher then shows two large illustrations to the class: one showing a drawing of a neutral mouth position and the other a drawing of a spread mouth position. The students are asked to imitate the mouth positions to determine which picture matches which vowel and to come up with an example word for each condition, e.g. sheep, ship. The teacher then shows the illustrations to the students with varying speed and every time the students see the spread mouth position they are asked to say one example word with a long /i:/ and one with the neutral mouth position for /ɪ /. Next, a repetition drill takes place with different settings: loud, quietly, whisper, fast, slowly, etc. As the students have already been introduced to /ə/, the teacher displays a poster for the /ə/ (see D3, Appendix, p. 237), reminding the students that it is the sound made by Tarzan, and collects example words on the blackboard. Then the teacher puts a poster for /e/ (see D4, Appendix, p. 242) on the board and the class is asked to come up with example words (e.g. egg, head) which are then repeated with the group (see Appendix, Section 10.4.2 on p. 238) for an outline of the use of the blackboard. The students are then asked to describe the articulatory gestures needed for this sound. Finally, they have to supply an example word for each of the four sounds in their phonetic chart (see Appendix, p. 260).

To understand how to pronounce diphthongs the students are shown the video and song "When two vowels go walking" (Between the Lions, 2009). With this catchy and memorable song, the students are helped to deduce that diphthongs are made up of two vowels, that the first one can be louder and longer and that there is a glide

between the two sounds[23]. The two posters for the /ɪ/ and the /ə/ are handed to two students. The two volunteers are asked to come to the front of the classroom, face the other students and stand about two metres apart from each other. First the student with the /ɪ/ moves towards the student with the /ə/ and the class says /ɪ/; then the student with the /ə/ moves closer to the other volunteer and accordingly the class says /ə/. The closer the students get to each other, the louder the first vowel should be pronounced and the closer the transition gets between the two vowels until a glide between /ɪ/ and /ə/ is established. After that the teacher tells the students that they should pretend not to hear someone and say: “eh”. With a bit of drama in their voice the students should repeat the “eh” (/eɪ/) several times. Then the teacher follows the same procedure for /eɪ/ as for /ɪə/ with two students moving towards each other at the front of the classroom and the class shouting the two elements of the diphthong. Next, the teacher puts the posters for /eɪ/ and /ɪə/ (see D5 and D6, Appendix, p. 242) on the blackboard and introduces some example words (e.g. `year´, `dear´, `page´, `grey´) which should be copied into the students’ sound charts.

The students are asked to pronounce the example words silently while watching their mouths in mirrors to identify the different sounds. They are asked to describe and discuss the movements of their mouth, tongue and teeth with their partners and in class. Finally, in pairs the students must silently pronounce a word containing one of the diphthongs while another student has to guess it, and vice versa. After that the teacher puts posters of the mouth position for /æ/ and /i:/ on the blackboard next to the posters of the diphthongs. The students should remember these illustrations from the introduction to the vowels (see Section 4.1). The teacher asks the students to copy the illustrated mouth positions and to say what sounds are made by these positions. Comparing the four English sounds in these ways requires the students to pay close attention to clear production and to listen carefully in order to identify the differences and discriminate between the sounds. The teacher then writes the sounds /æ/ and /i:/ next to the corresponding posters on the blackboard and collects example words from the students, e.g. `clap´, `hat´, `sheep´, `me´, etc.

Finally, the teacher hands the students the worksheet D1 which contains a table with columns for /eɪ/, /æ/, /ɪə/ and /i:/. The teacher slowly reads a list of words that the students have to identify and discriminate and put in the correct columns (see

[23] Although the first vowel in a diphthong is not usually pronounced with more volume, this approach might help the students to make the first vowel a little longer and put stress on it.

Appendix, p. 238 for the worksheet with solutions). To practise the diphthongs and presented vowels, the students are asked to write a poem with at least four lines. Each of the diphthongs and vowels have to be used at least once per line. After correction by a peer and the teacher, the students are asked to present their poems to the class paying special attention to the pronunciation of the trained sounds.

4.3 Word-final plosives: /b/, /d/ and /g/

The technical meanings of the terms ‘vowel’ and ‘consonant’ might not be completely clear to the students, so the teacher introduces the concept of consonant sounds by writing examples of two contrasting speech sounds on the blackboard, e.g. /ɑː/ and /d/. The students are asked to pronounce the sounds while concentrating on the movements of their tongue and mouth. In this way, they discover that unlike the production of vowels there is an obstruction in the mouth when pronouncing consonants. The students learn that vowels are speech sounds formed by air from the lungs travelling through the mouth without being impeded or interrupted, whereas consonants are speech sounds produced when the air from the lungs encounters an obstacle or block in the mouth area which interrupts the airflow (see Transparency P1, Appendix, p. 243). The students are asked to copy this information into their exercise books.

As the concept of voicing is also likely to be new to the students, the teacher begins the explanation by writing plosives on the board: /p/, /t/, /k/, /b/, /d/, /g/ (see Appendix, p. 243) outlines the use of the blackboard). Each student is then given a small piece of paper and instructed to hold it in front of their lips and to produce the sounds on the blackboard with a loud voice. The students learn that the paper moves for the voiceless plosives /p/, /t/ and /k/. Next, the students are asked to produce the sounds very slowly while touching their voice box: they should feel vibration for the voiced plosives /b/, /d/ and /g/. The teacher marks the voiced and voiceless consonants on the blackboard and writes the definition for ‘voiced’ and ‘voiceless’ on the board which the students copy into their exercise books.

Most German learners of English do not have problems with pronouncing the voiced plosives /b/, /d/ and /g/ correctly as they are also part of the German sound inventory (see Section 3.5.3). Therefore, the presentation stage for the voiced plosives focuses on providing the theoretical background, given that the German language

always requires devoicing of plosives in final position. Consequently, when /b/, /d/ and /g/ occur in final position in English, German learners tend to produce /p/, /t/ and /k/ respectively. It is therefore important to emphasise in the classroom that the final /b/, /d/ and /g/ are pronounced very "heavily" in English.

To build the identification and discrimination skills of the students, the teacher shows the class a picture of a "cap" and a "cab". He then pronounces one of the words and the students must discern whether it was "cap" or "cab". Following that, the students close their eyes and the procedure is repeated. Usually, students find it very difficult to discern the correct word, so then the class discusses why it might be so challenging to hear the difference. With the help of the German greeting "Guten Tag", in which the teacher pronounces the devoiced /k/ instead of the voiced /g/, the concept of final devoicing in the German language is explained and the students are made aware that they need to voice the plosives /b/, /d/ and /g/ when they occur at the end of a word. The class has to copy the rule into their exercise books (Rule 1 in Appendix, p. 243). Then they are asked to produce words with final voiced plosives. The students know that they are voicing it correctly when they feel the vibration of their voice box and no puff of air is pushed out.

Next, the students are given a worksheet with the minimal pair words `cab´ and `cap´, `bag´ and `back´, `bed´ and `bet´ (Worksheet P1 in the Appendix, p. 247). An audio file (Bowler, 2005) presents the word pairs several times. The students are asked to pay particular attention to the final consonants and the length of the pronounced words. The students learn that the vowels before the voiced consonants are longer than before the voiceless consonants. Again, the students copy this rule into their exercise books (see Rule 2, Appendix, p. 247). Then the class and the teacher together pronounce the six example words while touching their voice boxes and checking for puffs of air.

To practise their developing identification skills, the class continues work on worksheet P1. The students listen to the audio file and are required to discern which of the minimal pair is presented by the speaker, e.g. `back´ or `bag´. Then the teacher puts the three posters P1, P2 and P3 on the wall, showing the voiced plosives /b/, /d/ and /g/ and the corresponding example words (see Appendix, p. 247). The example words are drilled with the class in a funny and/or dramatic way. Finally, the students need to put an example word into their own sound charts (see Appendix, p. 260). For further practice, the teacher puts a transparency on the overhead projector which

shows the sentence: “The sad bad dog did his job and went after the fat cat in the bed.” The students practise the sentence together with their peers while checking for the correct pronunciation by feeling their voice boxes and for puffs of air. Then they are asked to say the sentence as fast as possible.

4.4 Fricatives: /θ/, /ð/ and /z/

German speakers of English usually do not have a problem with pronouncing the fricatives /s/ or /z/ correctly but due to word-initial voicing in the German language (see Section 3.5.3), /s/ at the beginning of a word is always pronounced as /z/, e.g. the English name Susan /su:zn/ is pronounced as /zu.zn/. However, the fricatives /θ/ and /ð/ are not part of the German sound inventory (see Section 3.5.3). Thus, learners frequently substitute them with /s/ and /z/ (Kenworthy, 1987).

To introduce the sounds /s/ and /z/ to the students, the teacher puts up a transparency with pictures of a bee (zzzzzzz), a snake (sssssss) and a person touching his voice box. In addition, there are pictures of a `zoo´ and a `Sue´ and the corresponding phonetic transcription is provided (see transparency F1 in the Appendix, p. 248). The students are asked to put their finger on their voice box and make the sound of a bee and then a snake and to determine which of the two sounds is voiced and which is devoiced. Then, the example words `zoo´ and `Sue´ are trained with the help of a repetition drill. The students are made aware that English words like `Sue´ are pronounced with an initial devoiced /s/. After that, the learners are instructed to alternate the two sounds without stopping, e.g. /s…s…s…s…z…z…z...z…s…/.

To introduce the fricatives /θ/ and /ð/, the teacher writes them on the board and asks the students whether they already know the two sounds. The teacher makes sure that the class understands that these symbols represent the letters “th” and writes: th → /θ/; and th → /ð/ on the board. The students have to describe what they need to do with their lips, mouth, etc. when they pronounce /θ/ and /ð/. The teacher collects the ideas and models an exaggerated version of the “th”-sound (Kelly, 2000). The students are then instructed to place their index finger against their lips. They should try to touch their finger with their tongue and breathe out to produce /θ/. Then they should try to add voice to produce /ð/.

In the presentation, the students are shown a video in which David Sconda's face is shown modelling the sound in isolation and also in example words (Sconda, 2008). Sconda exaggerates his pronunciation and facial expressions in an entertaining way, so the video is quite engaging for students. The students are asked to speak along to the video and exaggerate their mouth movements as well. Then, the teacher puts the phonetic posters F2 and F3 on the blackboard (see Appendix, p. 253) and uses the example words `thank´, `think´, `maths´ and `thin´ as well as `that´, `than´, `them´ and `paths´ in a repetition drill during which the students are asked to feel their voice boxes. Similarly, the phonetic poster F1 for the sound /z/ is put on the board and a similar exercise takes place with the example words `close´, `does´, `zip´ and `paths´ (see Appendix, p. 253).

The students are then given a worksheet in which they are asked to sort words from given sentences in the correct column with reference to the voicing status of the "th"-sound (see worksheet F1 in the Appendix, p. 251). The teacher then marks the worksheet and in case of difficulties should correct and practise any problematic words.

To practise the fricative pronunciations, some amusing tongue twisters containing /z/, /s/, /θ/ and /ð/ are presented by the teacher and on an OHP to the students which they then practise in the classroom (see Transparency F2 in the Appendix, p. 252). Finally, the students get together in groups and come up with a tongue twister on their own which includes the four fricative sounds.

4.5 Approximant: /w/

The English approximant /w/ does not equate to its German counterpart. Thus, German learners of English tend to replace the English /w/ with the /v/ (see Section 3.5.3). Although a mispronunciation of this consonant usually does not lead to a misunderstanding, the German use of the /w/ is often ridiculed[24], e.g. "Very vell!" (Quetz, 1998)

To introduce the /w/ and to help students differentiate it from /v/, the teacher writes `wet´ and `vet´ on the blackboard and models the two words for the class. The students are asked to listen and watch closely in order to detect how the different sounds are

[24] Some Germans who say "willage" instead of "village" and it always makes Anglophones laugh because it sounds like a derivation of "willy" (Clarkson 2019).

produced. After a class discussion, the teacher tells the students that /v/ is a voiced consonant. To make the sound, you need to bite your bottom lip with your top teeth and push air out. In contrast, /w/ is the short version of the long vowel /u:/. The students are told to put a pencil in the mouth and put their lips around it. They should then take the pencil out but keep their lips rounded and produce the sound by pushing air out. To learn the production of the approximant, the students should produce the two words `vet´ and `wet´ silently while watching their mouth in the mirror. Then they are asked to discuss the movements of their mouth, tongue and teeth with their peers. After that, one student produces either `wet´ or `vet´ silently while the other student guesses which of the two words was intended, and vice versa. Similar to the videos used for the vowels (see Section 4.1), the students watch a video showing an oversized close-up of the face of a native speaker modelling the sound /w/ in isolation and in the context of reference words, e.g. `walk´ (BBC, 2010). The students are asked to repeat after the speaker. After that, the teacher writes the word `one´ (which was featured in the video) on the blackboard and makes it clear that although there is no letter [w] in `one´, it is pronounced with a /w/. The teacher then puts up a poster with the phonetic transcription of /w/ and the example words `walk´, `one´, `swing´ and `sweet´ (see Appendix, p. 259) and instructs the class in a repetition drill.

To practise identifying the correct sound, the students are given a worksheet containing minimal pair items for /v/ and /w/, e.g. `vest´ vs. `west´ (see Worksheet W1 in the Appendix, p.254). The speech examples are presented from an audio file (Bowler, 2005) and the students are asked to listen for the differences between the items. The audio file is played again and this time the students are asked to speak along. After that, the teacher picks some random words from the worksheet and the students have to discern which one it is. This exercise is then repeated in group work.

To practise the /w/ sound, the students play a game called “Streetmap of Letterton” (adapted from Bowler, 2005) which is similar to “Battleship” but instead of finding ships, the students have to find houses in a grid of streets. To find the correct house, the students must correctly produce and distinguish between /v/ and /w/, as the street names include minimal pairs such as `West Road´ and `Vest Road´. Finally, the students are presented with several tongue twisters, such as “Vera’s wonderful wedding videos” (see Transparency W3 in the Appendix, p. 255). The students are asked to read them aloud and to come up with a tongue twister on their own.

This chapter has introduced the pronunciation training programme and described in detail the tasks, exercises, procedures and resources the teachers used in the classroom to deliver the intervention. Where appropriate, materials were referenced in the Appendix. To promote consistency of programme delivery across different teachers, a comprehensive teachers' manual (see Appendix, Section 10.4, p. 231ff.) was provided to participating teachers.

5. Methodological considerations: the acoustic phonetic basis of the current investigation

The previous chapter presented the pronunciation intervention. In order to analyse the impact of the investigation, auditory and acoustic analyses were carried out. This section aims to give a brief synopsis of the relevant acoustic foundations with regard to the sounds treated in the intervention.

5.1 The glottal sound source

Every voiced and periodic sound is produced by the rapid and repeated opening and closing of the vocal folds (see Figure 5-1) (Pickett, 1999). This periodic vibration is called phonation (Clark, 2007).

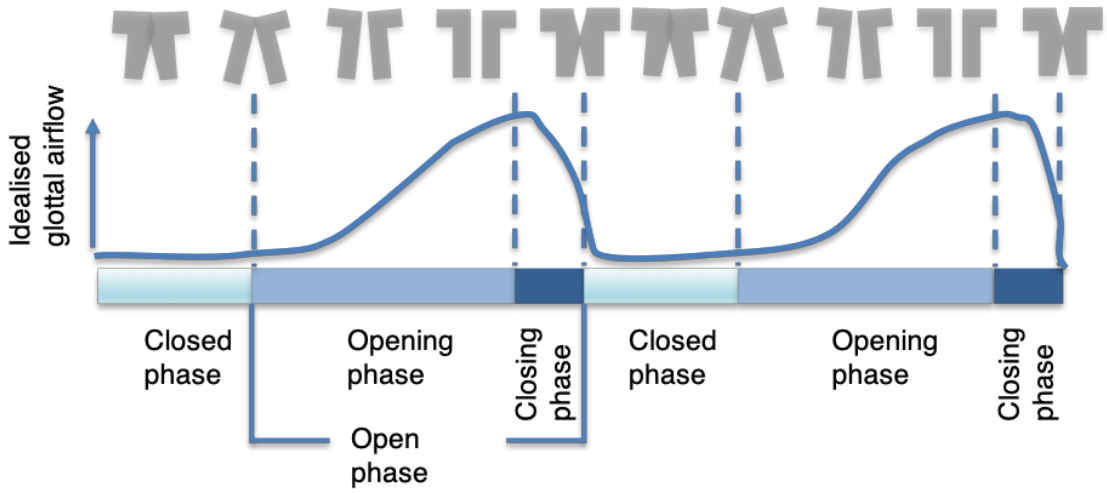


Figure 5-1: Two vocal fold vibration cycles and idealised glottal airflow waveform (adapted from Feilding, 2013)

When the vocal cords are brought together the airflow from the lungs is blocked and the subglottal air pressure increases until it forces the glottis to open. As the air flows through the opening the air pressure drops and the vocal folds move inwards and finally close. The Bernoulli effect supports the rapidness of the closure. When the subglottal pressure builds up again the cycle is repeated (Pickett, 1999). The opening and closing of the vocal cords causes puffs of air to flow through the glottal opening. The frequency of these pulses determines the fundamental frequency (f_0) of voicing and generates the perceived pitch of the produced sound (Haskins Laboratories, 2008; Pickett, 1999). Through the pressure fluctuation of the vocal fold vibration a complex periodic wave is

produced (see Figure 5-2) (Scarborough, 2005). The spectrum of the wave shows energy at the fundamental frequency of the glottal vibration and at its harmonics - which are the multiples of the fundamental frequency. The amplitude of the harmonics decreases gradually. The repetition rate of the glottal pulses corresponds to the spacing of the components of the glottal wave (Pickett, 1999) and depends on the air pressure generated by the lungs and the tension of the laryngeal muscles. The opening and closing movement determines the shape of the spectrum (Haskins Laboratories, 2008).

For all (voiced) speech sounds the glottal waveform is the source of the acoustic energy (Scarborough, 2005) and the spectrum of the glottal sound is reflected in every vowel spectrum (Pickett, 1999).

5.1.1 The Source-Filter Theory of Vowels

In 1960 Fant introduces the source-filter theory to describe speech production (Fant, 1960). To produce vowel sounds complex periodic waves created by the vocal fold vibration (source) are shaped by the vocal tract (filter) and transmitted to the outside air (Johnson, 2013; Pickett, 1999). Air in the vocal tract has particular resonance frequencies that are influenced by its shape and at which the contained air naturally seems to vibrate. These resonators can act as acoustic filters (Scarborough, 2005). Components of the sound source produced by the glottis at and near those resonant frequencies are amplified and the other frequencies are damped (Pickett, 1999). The specific resonances of the vocal tract are called formants (F) and are peaks in the vocal tract's filter function (see Figure 5-2 b/c) (Scarborough, 2005). In order of their frequencies the formants of a speech sound are numbered in ascending order; first formant (F_1), second formant (F_2), third formant (F_3) and so on (Pickett, 1999). As the overall shape of the vocal tract is altered in order to produce certain sounds the formant frequencies change accordingly.

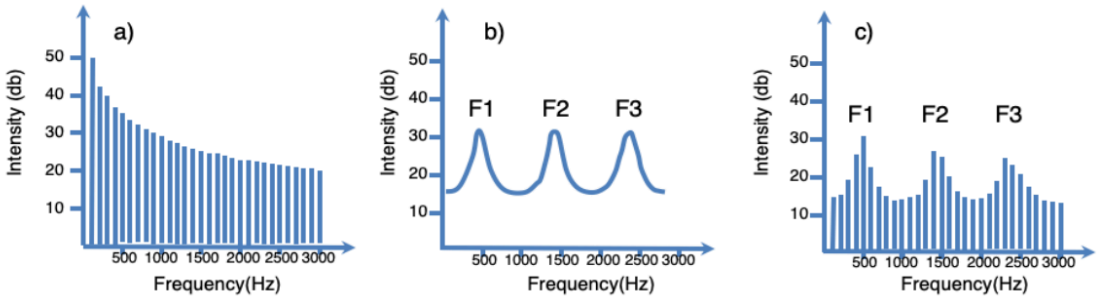
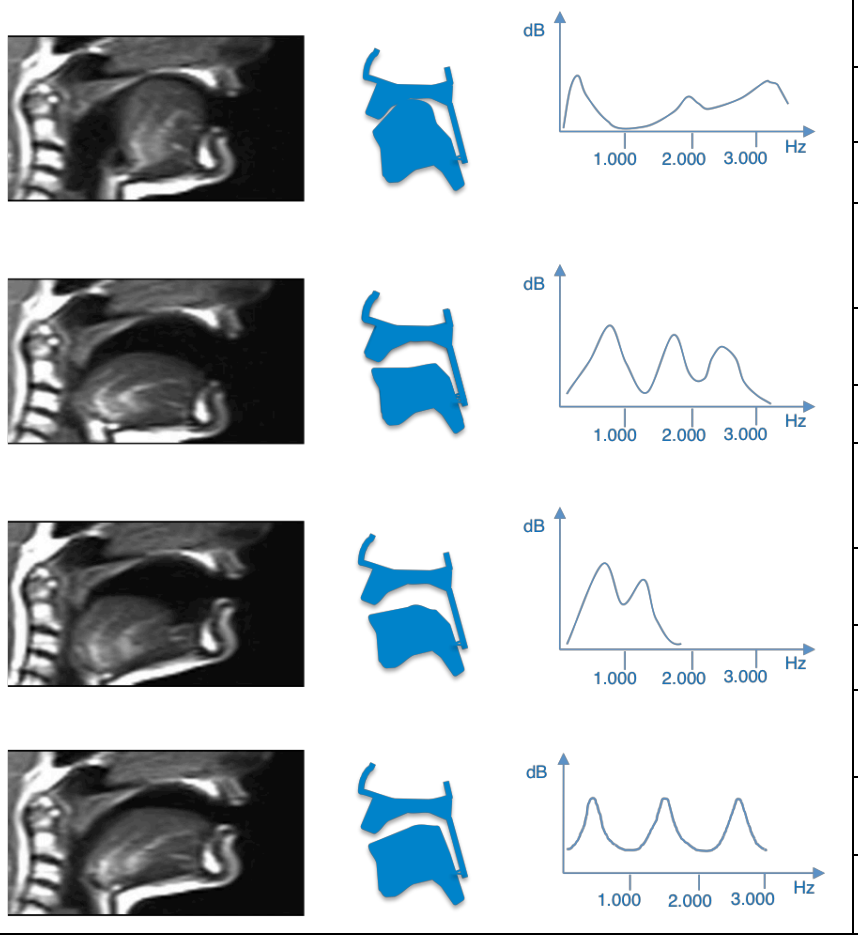


Figure 5-2: Source-Filter Theory a) glottal spectrum, b) Vocal tract filter response, c) vowel spectrum (adapted from Herrmann, 2010)

The source-filter theory helps to explain the details of vowel spectra through the glottal sound source spectrum that is filtered by the vocal tract (Pickett, 1999) and this filter determines the frequency characteristics of a particular sound (see Figure 5-3) (Scarborough, 2005).



/i/ close, front, long		
F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
280	2620	3380
/æ/ open front, short		
F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
800	1760	2500
/ɔ:/ half-open, back, long		
F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
480	760	2620
/ə/ mid central		
F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
500	1500	2500

Figure 5-3: MRI vowel images, position of vocal organs and adult male formant frequencies collected by Wells (adapted from Ladefoged, 1996, 2006; Pickett, 1999, p; UCL, 2014; Ultrax, 2012[25])

[25] Copyright clearance of the MRI pictures was obtained from Adam Baker on 29-01-2015

Figure 5-3 presents the first three formant frequencies of a typical adult male with an oral tract length of 17.5 cm. However, the length of the vocal tract depends on the speaker. The longer the length of the vocal tract the lower is the frequency. Thus, adults who generally have a longer vocal will have lower formant frequencies. As women and children tend to have shorter vocal tracts, they also have higher formant frequencies (Pickett, 1999). The length of the vocal tract is also correlated with the formant location and spacing of F3 and above (Pickett, 1999).

5.2 Acoustic analysis of Vowels

In the previous chapter the test items used in the intervention were presented. The following section discusses the acoustic background of the vowel and diphthong analyses particularly regarding the vowels and diphthongs trained in the intervention. After that, section 5.4 looks at the acoustic features of the consonants, which formed part of the intervention.

5.2.1 Acoustic analysis of vowels: The F1 x F2 plane

The most important acoustic vowel properties are the formant frequencies, which can be seen and analysed with the help of sound spectrograms (Ladefoged, 2003). Vowels are usually not spoken in isolation but appear in CV, VC or CVC sequences for example and change as a function of time as the speech organs move from one articulatory position to another. Thus, the formant frequencies show continuous movements that can be seen as spectral change in a spectrogram (e.g. Figure 7-2) and are called formant transitions (Harrington, 2013; Hayward, 2000). Although, there is no standard method for identifying the place of the vowel target (Harrington, 2012), in the acoustic analysis of speech, the steady state of the vowel is usually measured at the midpoint (Hayward, 2000; Ladefoged, 2003), because it is least influenced by transitional effects. This method is applied in this thesis and therefore the following discussion on the acoustic vowel analysis refers to the measurements taken at the midpoint of the vowel.

The most widely used characteristic of vowel formants is the correspondence of the first and second formant frequencies (Hayward, 2000; Ladefoged, 2003) as the frequency location of the first two formants (F1 and F2) depend on the shape of the vocal tract and are influenced by the articulatory movement of the lips, tongue, pharynx,

and jaw (Pickett, 1999). Figure 5-4 shows that if the steady state of F_1 is displayed in a chart on the ordinate and the steady state of F_2 on the abscissa with increasing values from right to left, a plot similar to the vowels in the IPA vowel quadrilateral (see 3.5.1) is created (Ladefoged, 2003). From this plot the vowel quality can be deduced (Ladefoged, 2006). F_1 x F_2 plots are one of the standard ways to display vowel qualities (Hayward, 2000; Ladefoged, 2007).

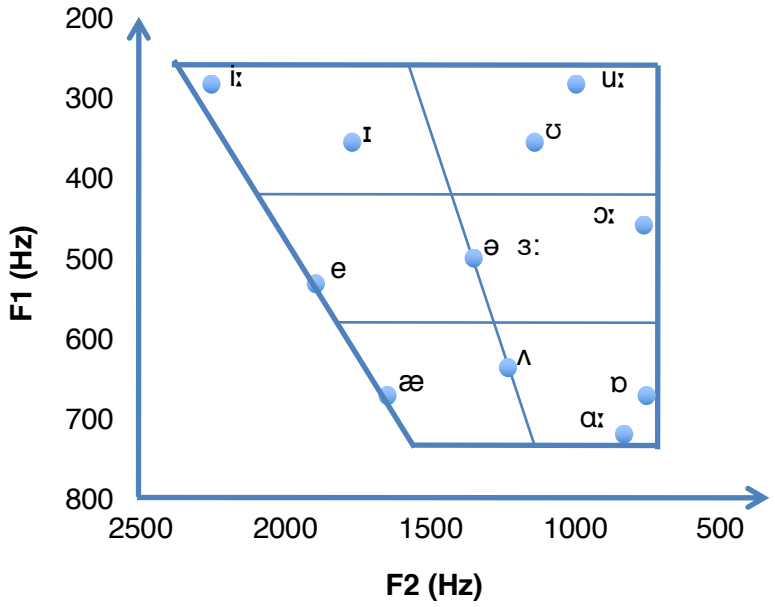


Figure 5-4: F2 x F1 plane for selected English vowels (adapted from Zhang, 2006)

Alternatively to plotting F2 versus F1, the F2 dimension can be replaced by the difference between F2-F1 which emphasises the auditory concept of `frontness´ and `backness´ of the vowels (Clark, 2007.; Ladefoged, 2006) and also reduces the differences between speakers (Hayward, 2000). On this basis F2 x F2-F1 plots were used in this thesis to display vowel qualities.

5.2.1.1 Nasalised Vowels

Vowel quality can also be influenced by nasalation. Nasals also use voicing as their sound source but in contrast to vowels, the air does not leave the vocal tract through the mouth but exits through the nasal cavity (Ladefoged, 2007). Therefore, the filtering of the oral tract is more complicated (Hayward, 2000). To produce the nasal, the velum has to move downward in order to open the velar port. This gesture begins about a 100 ms before the offset of the vowel, which causes a nasalisation of vowels for about 100 ms preceding the full onset of the nasal (Pickett, 1999). During the nasalisation of the

vowel, the vocal tract is open to produce the vowel configuration but also the velar port is also partially open adding a shunt to the overall system. This leads to extra resonances and anti-resonances and it has a frequency-tuning effect on the transmission of the glottal sound through the oral cavity (Pickett, 1999). Therefore, nasalised vowels tend to have lower formant frequencies than their non-nasalised counterparts (Johnson, 2005) and sometimes the first formant even tends to disappear (Ladefoged, 2003).

In the pronunciation intervention, two of the vowel test items (`thank´ and `aunt´) include nasalised vowels. To account for the nasalised part of the vowel, the duration and formant measurements include the nasalised portions of the vowel as well.

5.2.1.2 Other factors influencing vowel quality

The distribution of the vowels in the F1 x F2 plane does not only depend on the vowel quality itself but is also influenced by several other factors such as duration, hyperarticulation, coarticulation[26] and phonetic contrast.

In the F1 x F2 space, the peripheral positions are generally occupied by more tense vowels (e.g. /i/) while the lax vowels take up more neutral positions (e.g. /ə/). In order to produce the tense vowels, additional time and effort is needed to produce the required rather extreme articulatory gestures. In case of shorter word durations, there will be less time to produce tense vowel targets, which may result in a more schwa-like vowel reduction. Thus, short durations might result in a reduced vowel space that will shift towards the centre (Harrington, 2013).

Vowels can also be influenced by coarticulation in which the shifts in vowel quality can be attributed to the effects of the preceding and following segments (Harrington, 2013). Stevens and House (1963) demonstrated that the consonantal context leads to a centralisation of the vowels. In spite of this, Moon and Lindblom (1994) found that the vowels shift towards the direction of the neighbouring segment loci and do not necessarily centralise (cf. Harrington, 2013). Yet, the effect of coarticulation is surpassed by the displacement of vowel targets due to speaker variation (Harrington, 2013).

[26] Coarticulation describes the articulation of two or more speech sounds that influence each other (Oxford University Press, 2014)

Lindblom (1996) introduces the model of hyper- and hypoarticulation (H&H) in which speakers economizes their articulatory effort enough to be sufficiently intelligible to the listener. Thus, if it is necessary to be clearly understood, the speakers hyperarticulate their speech. This is especially common when vowels are produced in contexts from which they are difficult to predict and might be confused with others due to phoneme substitution. Hyperarticulated speech is generally associated with an expansion of the vowel space and a decrease in coarticulatory overlap. On the contrary, in everyday situations with clear contexts, speech is usually economised and hypoarticulated. In these cases segmental reduction takes place, vowels are produced with a more lax quality and the vowel spaces decrease (Harrington, 2013). Wright (2003) looked at isolated L1 words, which were ranked according to their frequency. He showed that low frequency words occupied extended vowel spaces in comparison to high frequency words (Harrington, 2012; Wright, 2003). This result supports the Lindblom's H&H model as low frequency words demand hyperarticulation in order to be clearly understood.

Emotions can be expressed via vocal cues and listeners are rather good in inferring their meaning. Emotions are primarily controlled by the limbic system and they affect the respiration, phonation, several temporal phenomena, such as tempo and pausing, and also the location of the formants (Banse & Scherer, 1996; Scherer, 1995). On this account, research carried out on the impact of stress and fear has shown that they can also lead to vowel reduction (Harrington, 2012).

5.3 Acoustic Analysis of Diphthongs

Diphthongs are vowel sounds that form a single syllable and include a change in vowel quality (Ladefoged, 2007). Therefore they can also be analysed with the help of F2/ F2-F1 vowel charts (see 5.2.1 above). However, as the vowel quality of the diphthong changes, measurements cannot simply be taken at the vowel targets. In order to keep track of the change in vowel quality the diphthongal movement of the formants can be plotted at several time intervals. These time points can be connected with an arrow pointing from the beginning to the end of the diphthong (Ladefoged, 2007). As the beginning and end of the diphthongs are immediately affected by coarticulation effects from the surrounding consonants (Jacewicz, 2009) in this thesis the formant measurements of the diphthongs were taken at four equidistant temporal locations

corresponding to the 20%, 40%, 60%, and 80% portions of the diphthong duration. Presenting the diphthongs in a formant chart leaves out information about the rate of change from the first vowel target to the second, especially as the second diphthong target is likely to be reduced (Harrington, 2013; Hillenbrand, James, Getty, Clark, & Wheeler, 1995). In addition, the spectral change is not influenced by coarticulation and is a systematic property of a specific diphthong (Fox & Jacewicz, 2009).

5.4 Acoustic features of Consonants

After discussing how to analyse vowels and diphthongs in 5.2 and 5.3, the acoustic background of consonants are described in the following section. Again, not all consonant features will be presented but there is a focus on the sounds in the test items which were trained in the intervention.

In contrast to vowels, consonants differ in their vocal tract shaping and in their sound source (Pickett, 1999). They obstruct the flow of air through the vocal tract (Roach, 2009) and there are three degrees of stricture caused by articulatory movements: complete closure, close approximation and open approximation (Carr, 2008) which can cause an absence of sound, an aperiodic sound, or a weaker voiced sound (Pickett, 1999).

Consonants can be classified according to the following articulatory features (Pickett, 1999):

- Manner of articulation
- Voicing
- Place of articulation

Table 5-1: The English consonants (adapted from Pickett, 1999,)[27]

Articulatory Features		Name of Feature						Distinctive Features
		Nonobstruent		Obstruent				
		Sonorant		Interrupted		Continuant		
Manner of Articulation		Glide	Nasal	Stop		Fricative		Articulator
			Ported					Velar Port
				Voiced	**Voiceless** Open Glottis	**Voiced**	**Voiceless** Open Glottis	Glottis
Place of articulation	Front Bilabial	**w**	m	**b***	p		ʍ	Lips
	Labiodental					v	f	
	Middle Dental					**ð***	**θ***	Tongue Blade (Coronal)
	Alevolar	l	n	**d***	t	**z***	s	
	Palatal	r				**ʒ***	ʃ	
	Back Velar	j	ŋ	**g***	k			Tongue Body (Dorsal)

5.4.1 Acoustic features of the approximant /w/

With regard to their acoustic features approximants are similar to vowels or diphthongs and are therefore also called semi-vowels (Harrington, 2013). However, the approximant /w/ which is part of the intervention typically differs in three ways from the corresponding vowel /u/.

First, due to the front bilabial place of articulation of /w/, there is a greater constriction of the vocal tract than for /u/ which leads to a small rise in pressure. To attain the /w/ constriction position the lips and tongue have to move rapidly and the formant frequencies of F1 and F2 are reduced during constriction. When the constriction is released and the articulators move back to the position of a following vowel, the frequency values start to rise again causing a u-shaped formant transition in the spectrogram which is typical of the approximants (Harrington, 2013; Pickett, 1999).

Secondly, depending on the following vowel the articulation of /w/ varies slightly and the constriction of the lips might be accompanied by a slight back constriction of the tongue (Ladefoged, 2006; Pickett, 1999).

[27] The consonants analysed in the intervention are highlighted with an asterisk and printed in bold type

Thirdly, the articulatory movement to and from the approximant constriction takes about 75 ms. In some cases where the vowel formants are located far from the consonants, the F2 transition might take a bit longer while the F1 transition stays about the same. The timing of the transition is an important feature of the approximant and it is faster than the movements between two diphthongal vowels (Pickett, 1999).

In the intervention /w/ is trained in initial and intervocalic position. For both positions the patterns are very similar. However, there might be a briefer constriction phase for the initial consonant as the sound-source production of the glottis might not start at the very beginning of the approximant constriction (Pickett, 1999).

As approximants share characteristics with diphthongs, in this study the pronunciation of the approximant /w/ will be evaluated in a similar way to the diphthongs (see 5.3). In addition to F1/ F2-F1 charts, spectral change will be taken into account to look at the direction and the timing of the approximant production.

5.4.2 Acoustic features of plosives (stops)

The following section looks at the acoustic features of plosives. Following the hypothesis that German learners of English might transfer the feature of German final devoicing to the English language, the intervention and the following section focuses on the voiced and voiceless plosives /b, p, d, t, g, k/ in final position.

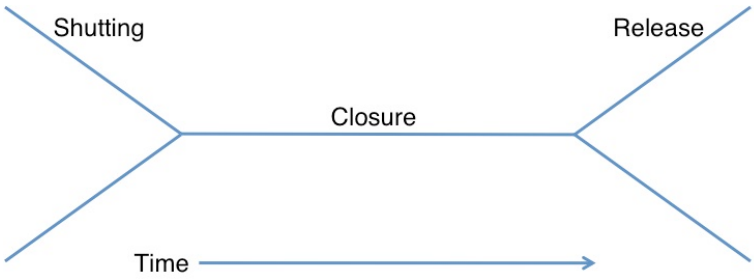


Figure 5-5: The three stages of plosive production (adapted from Johnson, 2005)

To produce a plosive, first the articulators have to move towards each other during a closure phase. During the occlusion there is complete silence for voiceless plosives or a low frequency sound of the lowest harmonics as long as voicing is maintained for voiced fricatives (Pickett, 1999). After the closure the articulators separate again as the plosive is released (see Figure 5-5). The air pressure in the

mouth rises during the closure stage, which results in a burst upon opening. These release bursts also occur in word final position (Hayward, 2000). However, these final bursts are on the whole weaker than their word initial counterparts.

The approximant /w/ that was introduced in the preceding section and the plosive /b/ have the same place of articulation, and are both formed by the constriction of the lips. However, in contrast to /w/ the plosives are produced with a dynamic movement (see 5.4.1) that includes the complete obstruction of the vocal tract (Johnson, 2005). During the complete lip closure, F1 would theoretically reach a frequency of zero and therefore right before the complete obstruction F1 reaches a lower frequency for /b/ than for /w/ (Pickett, 1999). To produce the low F1 typical for plosives, the articulators have to move more rapidly from and towards the neighbouring vowels. Because the plosive movement closes the vocal tract completely (see Figure 7-5), three acoustic features of a voiced stop are produced (Pickett, 1999):

- a very steep decrease toward closure and thus a rapid F1 transition;
- weak or absent low frequency sound during the closure;
- a burst of air pressure release.

The voicing feature of plosives depends on differences in vocal fold adjustment (Pickett, 1999). Not only is the position of the vocal folds responsible for voicing itself; it also produces other acoustic differences. The muscles of the larynx usually hold the vocal folds in an open position for voiceless plosives or in a closed position ready for voicing (Pickett, 1999). The open position of the vocal folds of voiceless plosives presents no obstacle for the flow of air from the lungs to the mouth and subglottal and mouth pressure are virtually equal in the closed phase of the plosive. Upon the release a strong flow of air travels through the small lip opening causing an intense and turbulent burst of the consonant (Pickett, 1999).

During the closure phase of a voiced plosive, vocal fold pulsing continues for some time, until the pressure difference between the pressure in the subglottal and mouth cavity becomes equal. When the lips open again for a voiced plosive release the mouth pressure goes down until there is sufficient pressure difference to the subglottal pressure for voicing action to resume again (Pickett, 1999). As the flow of air from the larynx is obstructed by the closed position of the vocal cords to produce voicing, the burst of voiced plosives is considerably weaker and has a shorter duration than the voiceless plosive burst (Pickett, 1999).

In addition to the voicing and difference of the burst, the length of the preceding vowel can distinguish voiced and voiceless plosives. Voiced plosives in pre-boundary position are preceded by considerably longer vowels than their voiceless counterparts. This voicing effect is very large and can take about 50 to 100 ms. It can be even larger for final position utterances (Pickett, 1999).

To assess the transfer of final devoicing from the German to the English language, the duration of the preceding vowel will be the main focus of the analysis.

5.4.3 Acoustic features of fricatives

The fricatives /θ, ð/ do not exist in the German language. Thus German learners of English tend to have difficulties in pronouncing them correctly. Very often the `th´ is substituted by /z, s/ or /f/ (see Section 3.5.3). Furthermore, it does not seem to cause any problems for German learners of English to pronounce the two fricatives /s/ and /z/. However, as the distribution of these two speech sounds is different in German and English and depends on specific phonetic contexts (Eckert & Barry, 2005), /s/ and /z/ were also included in the pronunciation intervention. Hence the following section concentrates on the acoustic features of /θ, ð, s, z/.

Fricatives can be easily identified by their typical hissing sound (Pickett, 1999) that is caused by a very narrow constriction of the vocal tract (Pickett, 1999). Due to the constriction the air molecules move irregularly and a turbulent noise (see Figure 7-6) is produced by the glottis (Johnson, 2005). Similar to white noise this friction sound covers a broad range of frequencies and shows random fluctuation in amplitude. However, fricatives are not only constricted by the vocal tract but the turbulent airstream hits a further obstruction shaped by the articulators to form the fricative consonant (Johnson, 2005). The length of the cavity in front of the constriction determines the spectral shape of this noise. As the front cavity gets smaller from the production of the fricatives /h/ to /ʃ/, /s/, /θ/, and /f/, the strongest resonances move upwards in frequency. This thesis looks at the fricatives /s/, /θ/ and /ð/. The alveolar fricative /s/ has the strongest resonances in a region around 4 kHz and the resonances of the interdental fricative /θ/ are around 5kHz (Pickett 2001). Despite the presented idealised acoustic properties, researchers have noted that fricatives are hard to identify out of context and that there are great difficulties in the analysis of fricatives (Cox, 2008; Johnson, 2005; Ladefoged, 2003 ; Smith, 2013; Wrench, 1995). This might be due to the fact that

different coarticulatory contexts and speaker variation have a strong influence on fricative frequency (Johnson, 2005). Additionally, the idealised frequency resonances mentioned above are not reliable because the number of formants can vary and fricatives also display anti-formants (Wrench, 1995). Especially, with regards to the dental fricatives, which are produced with a very small front cavity, there might be a diffuse spectrum with no major resonances and overall low energy for /θ/ (Harrington 2013). It is for these reasons, that the centre of gravity is frequently used to analyse fricative spectra (Forrest, Weismer, Milenkovic, & Dougall, 1988; Johnson, 2005; Wrench, 1995). This approach, which is also applied in this thesis, models the fricative spectrum as a single normal distribution, which reflects the formant of the dominant front cavity. The centre of gravity depicts the mean of this distribution (Wrench, 1995).

Voiced and voiceless fricatives are produced at the same place of articulation and accordingly their frequency spectra are influenced by the front cavity of the constriction. However, voiced fricatives show weaker intensities, as the vocal folds are held closer to produce the voicing and the available airflow is poorer in contrast to the voiceless fricatives where the vocal folds are held apart (Pickett, 1999). The vocal cords vibrate to produce the voiced fricative and thereby they modulate the airflow supplied to the fricative constriction. Correspondingly, the turbulence amplitude also shows a periodic modulation (Pickett, 1999) and thus the fricatives display ongoing f_0 frequencies.

Similar to the plosives there is also a strong duration effect of the constriction period and the neighbouring vowels. The vowel is lengthened by the voicing of the following consonant to 120 ms in utterance final position. In the non-final but pre-boundary position the lengthening is about 30 ms. In contrast, the consonant constriction is shortened by voicing (80ms) in utterance final position and 35 ms in non-final position.

In order to evaluate the fricatives, (see Section 7.5.2.4) the centre of gravity is used in this thesis to determine the place of articulation. A listener distinguishes the voiced from the voiceless fricatives by the duration of the neighbouring vowel (Eckert & Barry, 2005) and the length of the fricative (Cox, 2008). Thus, to assess the voicing feature both vowel and fricative duration are taken into account.

5.4.4 Acoustic features of affricates

The fricative /ʒ/ occurs in the German language only in loan words from French such as *Genre* or *Blamage* and the affricate /dʒ/ only exists in German foreign words like *Dschungel* and *Maharadscha.* In German, both voiced obstruents cannot occur in word final position due to the final devoicing rule (Eckert & Barry, 2005), and there is strong evidence that German speakers of English substitute the voiceless variant /tʃ/ over /dʒ/ not only in word final but also in initial and mid position (Eckert & Barry, 2005).

The affricates /tʃ/ and /dʒ/ are also seen as an additional manner of fricative articulation but in contrast to fricatives, affricates are preceded by an occlusion instead of more open articulation (Pickett, 1999). Therefore, they share properties with both plosives and fricatives. With reference to the acoustic analysis, the affricates will be treated similar to the fricatives and plosives and therefore durations of the preceding or following vowels and the affricate durations as well as centre of gravity measurements will be taken into account to analyse the affricate /dʒ/.

5.5 Word frequency effects on speech production

This section looks at the impact of word frequency on speech production. Educators, language teaching methodologists, psycholinguists, lexicographers, statisticians of language and corpus linguists have long since discovered the crucial role played by word frequency in natural languages (Sobkowiak, 2009). Word frequencies illustrate how often the vocabulary of a particular language is used. The more common a word the more important it is to know as the most frequent words often belong to the basic vocabulary of a language (Kilgarriff, 1997).

In 1965, Oldfield and Wingfield first investigated the word-frequency effect in speech production. They demonstrated in a picture-naming task that it took the participants longer to name pictures with low-frequency names (LF) (e.g. syringe) than pictures with high-frequency names (e.g. basket). Oldfield and Wingfield attributed this effect to the word naming itself and not as a result of object recognition (Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965).

Traditional psycholinguistic models have suggested that all speech output results from a segment by segment assembly (Varley, Whiteside, Windsor, & Fisher, 2006). However, contemporary research proposes a dual route approach that operates in

phonetic encoding and which depends upon word type and situational context (Whiteside & Varley, 1998b). Whereas the `indirect route´ uses sub-syllabic units, the `direct route´ operates via stored phonetic entities (Varley & Whiteside, 1998; Varley et al., 2006). These entities might result from the frequent encoding of high frequency units. A similar approach can be found in other domains of skilled motor control where frequent practice results in the formulation of global movement schemata (Schmidt, 1988). The 60 muscle groups of the vocal tract need to be coordinated to produce about 10 to 15 phonemes per second. This multifaceted process therefore results in high degrees of freedom. To generate a precise phoneme production, this complex computation makes use of verbo-motor patterns simplifying the links between muscle commands. Thus, the degrees of freedom are reduced and the articulatory movements become faster and more stable (Keller, 1987; Whiteside & Varley, 1998a). On this basis, speech produced via a more direct phonetic encoding route is more cohesive through higher degrees of coarticulation. In contrast, the speech output produced by a more indirect route shows more variability due to the verbo-motor patterns being calculated anew each time which results in lower degrees of coarticulation and greater degrees of freedom (Whiteside & Varley, 1998a). The direct route is computationally more efficient and is likely to operate for high frequency syllables and words. On the contrary, the indirect route is more likely to be used for less frequent word units as well as for newly learned L1 and L2 vocabulary. In addition, the indirect route might be used in cases of more conscious speech production, such as test situations or lectures (Whiteside & Varley, 1998a) Empirical investigations of the dual-route model of speech control have found shorter duration rates for high-frequency forms in comparison to matched low-frequency cognates (Jescheniak & Levelt, 1994; Varley et al., 2006). The increased duration rates were attributed to consolidation of the stored verbo-motor patterns.

L2 learners might produce more variable and inconsistent vowel formant patterns and increased utterance durations than those of L1 speakers. This might be due to the use of a more indirect phonetic encoding route. Along the same line of argument training effects should decrease duration rates and result in lower degrees of coarticulation.

5.6 Summary

Chapter 5 introduced the acoustic phonetic basis of the current investigation with regard to the selected sound core of the intervention programme. On this basis, the source-filter theory of vowels (see Section 5.1.1) was presented followed by the specific acoustic features needed for the analyses (see Sections 5.2, 5.3 and 5.4) of the vowels, diphthongs, plosives, fricatives, affricates and approximants. Finally, other factors influencing speech production, such as duration, coarticulation and frequency effects were presented. The next Chapter 6 presents the research focus of the hypotheses of the thesis based on the sound core of the intervention programme in conjunction with the auditory and acoustic phonetic measures which will be used to determine the efficacy of the pronunciation intervention.

6. Research Focus and Hypotheses

Drawing upon the content presented in the previous Chapters, the devised pronunciation training was implemented in the year 2011 over the course of five months within the language education of 105 English L2 learners at the ages of 10 to 12. The students belonged to two fifth grade classes at a comprehensive school, and to two fifth grades at a grammar school in Frankfurt/Main, Germany. To examine the pronunciation intervention, quasi-experimental classroom research was conducted using a control group design (see Section 7.1). In this context, each of the participating schools provided two additional fifth grade classes as control subjects (n = 95). To get within-subject measurements, the pronunciation data were collected at three time points; prior to the intervention (pre-test), after the intervention (post-test) and five months later (follow-up). Where feasible, between-subjects measures were collected via student questionnaires in order to look at individual variations such as differences due to the participants' sex, age, or social background. Although complete datasets were collected at all points for 181 out of the 200 initial participants only 16 students (see Section 7.2.4) were included in the final analysis due to the time constraints of a PhD project. This final set of participants was selected to maximise comparability between individual participants (see Table 7-2). From each of the eight classes a boy and a girl about the same age were selected. In addition, all of the final 16 participants had German as their L1. The assessment of the pronunciation data included two levels of analysis. Firstly, all stimuli were assessed auditorily using a three-point scale to evaluate the students' performance. Secondly, all data were analysed using the acoustic computer analysis programme Praat. The acoustic parameters in the analysis included fundamental and formant frequencies, vowel and consonant duration, vector and trajectory length, spectral rate of change and the centre of gravity to assess changes in pronunciation (see Chapter 5 and Section 7.5). The purpose of this study was to determine the effects of a pronunciation intervention programme in the classroom of young learners of English. In this context, several research hypotheses addressing this topic were set-up.

All subjects participating in this study will still be going through a range of maturation processes and also receive ongoing English language input. Hence, the first hypothesis H1 states the L2 pronunciation performance of both, the intervention and control group will increase over time (DeCoster, 2001; Mackey & Gass, 2005). In this thesis pronunciation performance is defined by the mean accuracy and acoustic measures.

Research Hypothesis #1 (H1):

- For both groups of subjects, the mean L2 pronunciation performance score will increase from pre-test to post-test and from post-test to follow-up test.

It is the purpose of this thesis to show that pronunciation training based on second language theories will significantly increase pronunciation abilities and outscore maturation processes. Accordingly, the second hypothesis H2 supports this claim.

Research Hypothesis #2 (H2):

- The group receiving the pronunciation intervention will perform significantly better with regards to their L2 pronunciation abilities on the post-test and follow-up-test compared to the control group

The third hypothesis H3 accounts for the fact that performance usually shows the highest increase directly after treatment (Shadish et al., 2001).

Research Hypothesis #3 (H3):

- The group receiving the pronunciation intervention will perform better on the post-test compared to the follow–up test

A number of studies (Best, 1995; Flege et al., 2003; Iverson et al., 2003; Kuhl, 2000) shows that diverse L2 sounds pose different degrees of difficulty for a language learner. Therefore hypothesis four H4 looks at the performance level for each sound. It remains to be seen if the treatment will foster all target sounds in the same way or whether some acquisitional patterns will emerge.

Research Hypothesis #4 (H4):

- The mean pronunciation performance scores will differ with reference to specific target sounds

To account for the fact that the pronunciation performance of the treatment group does not only improve due to the training effects of the intervention but rather due to the intervention itself, test-items that are explicitly trained in the treatment are contrasted with untrained test-items. Within this context the fifth research hypothesis H5 looks at the training effects.

Research Hypothesis #5 (H5):

- A) In the pre-test there will be no difference in the mean performance scores between the trained and the untrained test items.
- B) The intervention group will show higher mean performance scores on the trained test-items compared to the untrained test-items in the post- and follow-up tests.
- C) The intervention group will show higher mean performance scores in the post- and follow-up tests on the untrained items in comparison to the control group.
- D) The control group does not show any difference in the mean performance scores between the trained and untrained test-items.

Frequency effects are known to influence speakers' performances (Bishop & Keating, 2012; Cholin, 2008), as more frequent use consequently leads to increased abilities. Thus, the study controls for high- and low-frequency test-items in both the intervention and the control group. Research hypothesis H6 deals with the fact that higher frequency items will outscore the lower frequency items.

Research Hypothesis #6 (H6):

- For all groups of subjects, the mean performance score on higher frequency items will be greater than the mean performance score on lower frequency items.

Before proceeding to evaluate the hypotheses at hand, it is first necessary to have a closer look at the data collection and analyses employed in this study. Therefore, the following chapter 7 presents the methods section used in this thesis.

7. Methods

Based on current second language acquisition theories, established and proven pronunciation tasks, and outcomes of questionnaires completed by 245 English language teachers in Hesse (see Section 2.4), a core set of significant pronunciation problem areas of English L2 learners was identified and a pronunciation training programme was devised (see Chapter 4), which was implemented over five months within the language education of about 105 English L2 learners at the ages of 10 to 12 in Frankfurt/Main. It is the main aim of this thesis to determine whether, and to what extent the pronunciation intervention can improve L2 pronunciation. Hence, this section describes and discusses the methods used to evaluate the pronunciation intervention. First, the study design is presented in section 7.1, then the participants are introduced in section 7.2. The set-up of speech material used to test the subjects' pronunciation is illustrated in section 7.3. Finally, the data collection and data analysis are described in sections 7.4 and 7.5.

7.1 Study design

To examine the pronunciation intervention, quasi-experimental classroom research was conducted using a control group design (DeCoster, 2001; Gass & Mackey, 2007). To get within-subject measurements, the data were collected at three time points; prior to the intervention (pre-test), after the intervention (post-test) and six month later (follow-up). Where feasible, between-subjects measures are considered in order to show individual variations such as differences due to the intervention, the participants' sex and social background. In addition, the speech material itself is analysed for frequency and training effects.

7.1.1 Classroom Research

Quasi-experimental classroom research offers several advantages and drawbacks. The main reason for choosing quasi-experimental classroom research lies in its validity for real-life education since laboratory studies limit the possibilities to transfer the findings to the actual classroom (Hulstijn, 1997). However, it is very difficult to isolate the variables of a study in a classroom setting, as it is barely impossible to control for the many confounding factors in a classroom, the diverse nature of students, previous

knowledge and educational contexts. For this reason, classroom research is never completely objective (Finkbeiner, 1996). To eliminate alternative explanations as best as possible, and to link the results to the intervention, this research project adheres to the standards of quality for classroom research (see Wellenreuther (2000) for further information). Evidence of this is provided at the relevant points within the thesis.

7.2 Participants

This section presents the participating parties in this research project and outlines the criteria applied to the recruitment of the research subjects. As the research project focuses on classroom research in Germany, the `Ministry of Education´ provided ethical clearance (see Appendix, p. 206) and school consent was obtained[28].

7.2.1 Ethical clearance

Prior to commencing the study, ethical clearance was sought from the `Hessian Ministry of Education´. In addition, the school committees of the respective schools agreed to partake in the study. On December 7th, 2009, the Hessian Ministry of Education granted the approval of the research project according to § 84 of the decree of scientific research in schools of the Hessian School Law in concordance with the following standard practice conditions:

- Every subject concerned by the study has to be informed that participating is voluntary and that there are no consequences from not taking part. A written consent form is needed from the head of schools, teachers and students who are taking part in the research. According to data protection regulations a parent consent form is not needed in this study.
- Every subject concerned by the study has to be informed about the aims and contents of the research. Moreover, the nature of participation and data analysis has to be made clear prior to the study.
- The scientific research has to be done anonymously. In case biographical characteristics are collected, they have to be separated from other collected data and material already during the data analysis and they have to be saved in different places. At the latest they are to be deleted after the data analysis. As video and audio are counted as biographical data, specific restrictions of the participation in and analysis of the recordings have to be taken into account.

[28] As the PhD is supervised in Sheffield, UK, the Chair of the ethics review procedures at the Department of Human Communication Sciences in Sheffield informally approved the German ethics approval (see Appendix, p 206)

- The data collected is only to be used for the requested aims. The transmission of data to third parties is not allowed.
- The research done in school has to be conducted in such a way that there is a minimal interference with the regular school education.
- The study has to be carried out according to the Hessian Data Protection Act, in particular to § 37, § 10, § 13, and § 33.

7.2.2 Participating schools

Several criteria were used for the selection of schools. These included the possibility to cover a wide range of social and learning backgrounds, to provide a large enough amount of participating students; to be in the close vicinity of the university conducting the research project; and most importantly, a strong interest in taking part in the pronunciation study. Finally, two schools in Frankfurt am Main were chosen for this study: An integrated comprehensive school[29] in the district of Niederusel and a grammar school in the inner city area Westend.

The comprehensive school has about 1200 students and offers eight classes in each of the grades from five to ten. Half of them are classes integrating students with special needs and are therefore supplied with an additional social education worker (Schneider, 2014).

The second school is a grammar school in the inner city of Frankfurt am Main. In total, 1050 students attend the school, which offers four classes in each of the grades from five to thirteen. Music as well as content- and language-integrated learning (CLIL)[30] are important parts of the school profile. Moreover, the school offers preparation courses for diverse modern language certificates, such as the Cambridge Business English Certificates (BEC) and Certificate of Proficiency in English (CPE) and an International Baccalaureate Diploma (IB) (Grammar school 1, 2014).

Both schools agreed to provide a room for the data collection and the required media equipment for the study, such as projectors, laptops, and access to student computers.

[29] Germany has a tripartite school system comprising of three different school forms, i.e. Basic level (Hauptschule), middle level (Realschule), and upper level (Gymnasium). An integrated comprehensive school means that students of all three levels are taught together.

[30] Content and Language Integrated Learning (CLIL) involves teaching a curricular subject through the medium of a language other than that normally used. The subject can be entirely unrelated to language learning, such as history lessons being taught in English in a school in Spain (European Commission Multilingualism, 2008).

7.2.3 Participating teachers and the implementation of the intervention

The `Hawthorne effect´ describes that the mere presence of a researcher can change the classroom routine and therefore alter the students' behaviour in many subtle ways potentially resulting in more attentive behaviour (Gass & Mackey, 2007; Wellenreuther, 2000). Moreover, to cause minimal interference with the regular school education as demanded by the Hessian Ministry of Education (see Section 7.2.1) and to keep the educational cost at a minimum, the pronunciation intervention was designed to be a small but integral part of the English lessons conducted over the school term. Hence, English teachers were asked to carry out the classroom experiment. Finding teachers who were willing and able to teach the pronunciation intervention in grade five (student age: about 11 years) also played an essential part in the selection of the participating schools. Subsequently to presenting the research project to English school departments, two teachers at the grammar school and two teachers at the comprehensive school who were able to teach the intervention and available in the required timeframe were chosen to carry out the pronunciation training in their English lessons. All of them were fully educated teachers of English with an excellent command of the English language (at level C2 of the Common European Framework of Reference).

In classroom research, it is rather difficult to monitor how far the participating teachers adhere to the researcher's guidelines in the delivery of the chosen content of the experiment (Gass & Mackey, 2007). Therefore, the teachers were trained before the intervention on how to use the teachers' manual that included the pronunciation tasks and close guidance was provided throughout the project. To ensure the comparability of the pronunciation teaching, the teachers' manual not only included the pronunciation tasks and materials but also provided an exact plan listing the detailed procedures on how to teach the contents of the intervention. Although the teachers could choose freely when to teach the pronunciation intervention, content wise, they were asked to stick precisely to the manual and to keep a diary on the plan's execution (see Chapter 4 and Section 10.4 for teachers' manual).

As the study looks at the influence of the intervention, the other conditions in the classroom need to be as similar as possible in the eight groups. With reference to the extent of the pronunciation training, the duration of the intervention corresponded by and large to the regular pronunciation work done in the classrooms of the control

groups. As deduced from the teacher questionnaires, teachers include about nine hours of pronunciation work in their classroom over the course of a school term. This time frame was hence applied as the extent of the pronunciation intervention.

7.2.4 Participating classes and students

Referring to the effect of age on learning second language pronunciation (see chapter 2), pupils between the ages of 10 to 12 were identified as the most suitable subjects for the pronunciation study. Therefore, grade five was chosen to be the most fitting age group.

In classroom research a complete randomisation of the subjects is usually not possible as the students remain in intact classes throughout the treatment so that the intervention is as little intrusive as possible (Gass & Mackey, 2007; Mackey & Gass, 2005). Hence, in this particular line of research, the assignment to control and intervention group is usually done via matching of intact classes (Mackey & Gass, 2005). However, it is impossible in a classroom experiment to have completely comparable subjects and it might also be possible that the outcome is only true for a specific group (Wellenreuther, 2000).

As four teachers (see Section 7.2.3) at two schools were selected to carry out the pronunciation intervention, their four English classes were appointed as treatment groups. The assignment of the control groups took place according to possible access and the highest achievable comparability between the intervention and control classes. Therefore, two fifth grade classes at each of the two schools carrying out the intervention were chosen (see Table 7-1). These control groups contained students of a similar age and language as well as social background. To control for the differences in the distribution of the pronunciation training, all four classes were taught by four different teachers at the grammar school. In contrast, at the comprehensive school each of the two teachers taught a control and an intervention group. Moreover, in one case the intervention group also included special needs[31] students.

[31] The special needs students were excluded from the study. Nevertheless, classes with special needs students are explicitly mentioned as they have an additional educational social worker and a different learning atmosphere.

Table 7-1: Matching of the experimental groups

	Comprehensive School A					*Grammar school B*				
class	5c	5d	5e	5f	Total	5a	5b	5c	5d	Total
students (n=)	18	24	18	20	80	36	30	27	27	120
male	12	15	9	15	51	15	15	11	13	54
female	6	9	9	5	29	21	15	16	14	66
intervention group		x	x			x			x	
control group	x			x			x	x		
same teacher	x		x							
special needs students	x		x							

To safeguard anonymity and to ensure an objective evaluation of the data, the participants' names were encrypted. The code consisted of the first two letters of the name, the first two letters of the surname and the first two numerals of their birth date (cf. Fehling, 2008); e.g.:

- Franz Böhme 01.01.1999 → FR BÖ 01

As demanded by the guidelines of the Hessian Ministry of Education all participants involved in the study were informed that participation was voluntary and that there were no consequences from not taking part. Due to the set-up of the experiment, the following types of pupils were excluded from the test population:

- pupils with speech, hearing, reading or mental impairments
- pupils whose L1 is English
- pupils who did not complete all elements of the data collection

The initial sample consisted of 200 students. However, three of them did not complete the student questionnaire and 16 were not available for all three data collection periods and were therefore excluded from the study. Complete sets data were collected for 181 participants. However, the acoustic analysis was very time-consuming. Therefore, it was only possible to analyse the data of 16 students[32]. This final set of participants was selected according to maximise the comparison (see Table 7-2). From each of the eight classes a boy and a girl about the same age were selected. Half of the students belonged to the comprehensive school A and the other to the grammar school B. In addition, all of the final 16 participants had German as their L1.

[32] Comparable studies look at 11.61 subjects (Sakai & Moorman, 2018)

Table 7-2: Final set of the participants taking part in the intervention study

	Comprehensive School/ School A					Grammar school/ school B				
class	5c	5d	5e	5f	Total	5a	5b	5c	5d	Total
students (n=)	2	2	2	2	8	2	2	2	2	8
male	1	1	1	1	4	1	1	1	1	2
female	1	1	1	1	4	1	1	1	1	2
intervention group		x	x			x			x	
control group	x			x			x	x		
same teacher	x		x							
special needs students	x		x							

Table 7-3 and Table 7-3 provide an overview of the selected set of participants.

Table 7-3: Participants intervention study by group, sex, age, class and school

Participant	Intervention	Sex	Age	Class	School
1	Control	female	10.48	5c	A
2	Control	male	10.62	5c	A
4	Control	female	10.65	5f	A
3	Control	male	10.53	5f	A
5	Control	female	11.19	5b	B
6	Control	male	11.41	5b	B
8	Control	female	11.17	5c	B
7	Control	male	10.7	5c	B
9	Intervention	female	11.4	5d	A
10	Intervention	male	10.98	5d	A
12	Intervention	female	10.72	5e	A
11	Intervention	male	11	5e	A
13	Intervention	female	11.25	5a	B
14	Intervention	male	10.91	5a	B
15	Intervention	female	11.38	5d	B
16	Intervention	male	10.8	5d	B

7.3 Speech material

The following section describes the speech material used for the data analyses and assessment of the intervention programme. As discussed in Section 3.5, 14 British English sounds were selected for the intervention study:

- Monophtongs: /ɑ:/, /ɔ:/, /æ/, /ə/;
- Diphthongs: /eɪ/, /ɪə/;
- Consonants:
  - Plosives (in final position): /b/, /d/, /g/;
  - Fricatives: /θ/, /ð/, /z/;
  - Affricate: /dʒ/;
  - Approximant: /w/.

In speech, sounds usually do not appear in isolation but in meaningful contexts. Therefore, in this study, each of the selected sound is contained in a target word. As all the participants were language learners at a beginner's level, they did not have a lot of English vocabulary at their command. To have enough suitable target words, first a vocabulary analysis took place (see the following section). This list was then evaluated with reference to word frequency (see Section 7.3.2). To find out whether a change in pronunciation can be directly linked to items that are explicitly trained in the intervention the final word list contained trained and untrained speech items (see Section 7.3.3).

7.3.1 Vocabulary analysis

The participants in the study learn English vocabulary based on the books used in their English lessons. The grammar school students used *Cornelsen*'s `G21 A1´ (Schwarz, 2006), and the comprehensive school students worked with *Diesterweg*'s `Notting Hill Gate 1´ (Edelhoff, 2007a). The vocabulary the pupils learned and used at the time of the intervention corresponded by and large with the vocabulary lists taken from the two books - especially as both books also included vocabulary learnt at grammar school. The word lists from the two respective textbooks were compared to compile a list of matching vocabulary. From this list, words were chosen that contained the target sounds. Wherever possible only monosyllabic words were included in the final core of speech stimuli to control for coarticulation and stress effects. However, in case of /ə/, there was an insufficient amount of words. Therefore, the two disyllabic items (`melon´, `mother´ were included as stimuli (see Table 7-4 for an overview of the speech stimuli).

7.3.2 Frequency analysis

In order to control for frequency effects (see Section 5.5) the stimuli needed to be compiled from lower and higher frequency words. So, the matched vocabulary list was analysed according to the frequency of the lemmatised[33] words using the grammatically tagged British National Corpus (BNC) which is based on a count of a 100 million word collection of samples of written (90%) and spoken (10%) language representing a wide cross-section of British English from the later part of the 20th century[34]. Words which

[33] Lemma: The headword represents the inflectional variants of the word (Leech, 2001; Kilgarriff, 1997) e.g. *like* is a headword and *likes*, *liked*, and *liking* are its inflected forms (Chujo, 2004).

[34] The written part of the BNC (90%) includes extracts from regional and national newspapers, specialist periodicals and journals for all ages and interests, academic books and popular fiction, published and

occur with an overall frequency of 100 or more per million can be called as high frequency words (Chujo, 2004). Words between 100 and a 10 per million can be seen as middle frequency words and words with a frequency of less than a 10 per million can be seen as low frequency words (Leech, 2001).

Although the written and spoken part of the BNC includes texts for all ages and interests; it does not specifically focus on the most frequent words used by 10 to 12 years old pupils. Therefore, some low frequency words that belong to the immediate living environment of the pupils, such as `bag´ which is used in the classroom everyday, were considered as higher frequency words for the participating subjects. As the vocabulary list of the language learners was quite limited, not high versus low but **higher** (HF) versus **lower** frequency (LF) items were compared.

7.3.3 Speech stimuli

Table 7-4: Overview analysed peech stimuli

	Sound	Testitem		Frequency per million	HF/LF	Trained	No.
Vowels	/ɑ:/	arm	/ɑ:m/	202	HF	T	1
		aunt	/ɑ:nt/	33	LF	T	2
		dark	/dɑ:k/	104	HF	NT	3
		scarf	/skɑ:f/	5	LF	NT	4
	/ɔ:/	all	/ɔ:l/	215	HF	T	5
		draw	/drɔ:/	15	LF	T	6
		small	/smɔ:l/	518	HF	NT	7
		chalk	/tʃɔ:k/	9	LF	NT	8
	/ə/	from[35]	/frəm/	4134	HF	T	9
		clever	/klevə/	25	LF	T	10
		mother	/mʌðə/	295	HF	NT	11
		melon	/melən/	3	LF	NT	12
	/æ/	thank*[36]	/θæŋk/	131*	HF	T	13
		clap	/klæp/	0	LF	T	14
		bag*	/bæg/	75*	HF	NT	15
		maths	/mæθs/	10	LF	NT	16
Diphthongs	/ɪə/	year	/jɪə/	1639	HF	T	17
		dear	/dɪə/	41	LF	T	18
		hear	/hiə/	367	HF	NT	19
		beer	/bɪə/	38	LF	NT	20
	/eɪ/	page*	/peɪdʒ/	151*	HF	T	21
		grey	/greɪ/	48	LF	T	22
		wait	/weɪt/	213	HF	NT	23
		cage	/keɪdʒ/	13	LF	NT	24

unpublished letters and memoranda, school and university essays, among many other kinds of text. The spoken part (10%) consists of orthographic transcriptions of unscripted informal conversations (recorded by volunteers selected from different age, region and social classes in a demographically balanced way) and spoken language collected in different contexts, ranging from formal business or government meetings to radio shows and phone-ins (Oxford University Computing Services, 2012).

[35] In the context of this study the reduced form of /ə/ is used (Wells 2008).

[36] * high frequency for the age group. Moreover, this item is frequently used as an example in the textbooks

Plosives	/b/ final	job	/dʒɒb/	326	HF	T	25
		web	/web/	46	LF	T	26
		club	/klʌb/	202	HF	NT	27
		verb	/vɜb/	14	LF	NT	28
	/d/ final	bad	/bæd/	264	HF	T	29
		bread	/brɛd/	38	LF	T	30
		child	/tʃaɪld/	710	HF	NT	31
		gold	/gəʊld/	4	LF	NT	32
	/g/ final	big	/bɪg/	338	HF	T	33
		egg	/ɛg/	62	LF	T	34
		bag*	/bæg/	75*	HF	NT	35
		leg	/lɛg/	118	LF	NT	36
Fricatives	/z/	close	/kləʊz/	154	HF	T	37
		jeans	/dʒiːnz/	13	LF	T	38
		goes	/gəʊz/	148	HF	NT	39
		zoo	/zuː/	9	LF	NT	40
	/ð/	that	/ðæt/	7308	HF	T	41
		them	/ðem/	1733[37]	LF	T	42
		the	/ðə/	61847	HF	NT	43
		smooth	/smuːð/	30	LF	NT	44
	/θ/	thing	/θɪŋ/	776	HF	T	45
		thin	/θɪn/	56	LF	T	46
		month	/mʌnθ/	398	HF	NT	47
		teeth	/tiːθ/	47	LF	NT	48
Affricates	/dʒ/	job	/dʒɒb/	326	HF	T	49
		jeans	/dʒiːnz/	13	LF	T	50
		John	/dʒɒn/	328	HF	NT	51
		cage	/keɪdʒ/	13	LF	NT	52
Approximants	/w/	one	/wʌn/	1962	HF	T	53
		swing	/swɪŋ/	34	LF	T	54
		wait	/weɪt/	213	HF	NT	55
		witch	/wɪtʃ/	9	LF	NT	56

The present study examines English sounds in target words spoken by an intervention and control group of German learners of English. As discussed in the previous section, the items were matched for higher (HF) and lower frequency (LF) and consist of stimuli that were trained (T) in the intervention as well as known but untrained (NT) words to assess the subjects' abilities to transfer the learned pronunciation to other contexts. In total, 28 higher frequency and 28 lower frequency words containing the selected sounds were chosen as the target speech stimuli (see Table 7-4).

To have a comparable pre-, post-, and follow-up test, the same set of 56 stimuli was used for each test (cf. Mackey & Gass, 2005, p. 149). To get a fair view of the subjects' pronunciation ability, each test item was presented three times at each test period. In total, each data set consisted of 168 (3 x 56) words that were randomised beforehand[38]. The items were presented in the same sequence for all participants.

[37] `them´ was selected as a lower frequency item due to a typo in the frequency number. However, it still has a lower frequency than `the´ and `that´.

[38] Randomness was achieved by atmospheric noise (Oxford University Computing Services, 2012).

However, the order of the stimuli was randomised in a different order for the pre-, while- and post-test to avoid familiarisation and coarticulation effects.

7.4 Data collection

Having discussed the set-up of the speech material in the previous part, this section now presents the different procedures of the data collection.

7.4.1 Pilot testing

All procedures, materials and methods of this research project were piloted beforehand and necessary revisions were made before they were used with the final participants. Two English teachers in Frankfurt, Hesse, agreed to pilot and evaluate the teacher questionnaire and the teachers' manual. Consequently, several tasks of the manual and the questionnaire were excluded or altered. With reference to the pronunciation intervention two ten years old girls from a comprehensive school in Wiesbaden, Hesse, volunteered to test and evaluate the tasks, the technical equipment and test items used for the data collection and the student questionnaire. As a result, particularly the vocabulary used and the phrasing of the intervention tasks and the wording and length of the questionnaire were revised. The produced pilot data were then used to test and evaluate the data analyses.

7.4.2 Student questionnaires

To find out about the biographical and language background of the students and their attitude towards pronunciation, the 200 pupils were asked to complete an online questionnaire created with `Limesurvey´ software (version 1.85, (Limesurvey, 2012)). The questionnaire mainly consisted of closed-item questions and gathered information on the following topics (see appendix chapter 10.3.3 for original student questionnaire):

- class/ school;
- age;
- sex;
- nationality;
- amount and level (e.g. native speaker, I always talk in Hindi with my mum) of languages spoken,
- number of years of formal English instruction;

- time (if any) spent in an English speaking country;
- familiar pronunciation tasks (e.g. listening and sorting);
- pronunciation difficulties (e.g. `th´);
- attitude towards pronunciation teaching.

The students were asked to complete the online survey at the respective school workstations and guidance on how to fill in questionnaire was provided at all times. To provide anonymity, the data were encrypted.

Due to the constraints of the PhD project, only 16 students were chosen for the final analyses (see Section 7.2.4). The questionnaire data was intended for between measures data. Due to the small final sample of the test subjects, the questionnaire was not used in this thesis.

7.4.3 Elicitation technique

Studies on L2 foreign accent use different kinds of data elicitation techniques. Mostly, the subjects are asked to read the sentences or words (e.g. Asher & Garcia, 1969; Bongaerts et al., 1997; Moyer, 1999), others ask the students to recount personal experience, or do picture naming tasks to produce samples of free speech (e.g. Oyama, 1976; Thompson, 1991). Another possibility is the delayed repetition technique (Piske, MacKay, & Flege, 2001b, p. 193). All of them offer several advantages and drawbacks. Oyama (1976) and Thompson (1991) report that possibly due to differences in reading ability, read speech is often judged to be more foreign accented than speech that is produced freely. However, in free speech subjects might not produce the desired word or try to avoid difficult L2 sounds (Moyer, 2007; Piske et al., 2001b). Additionally, since morphological, syntactic, and pragmatic abilities are needed, the result might be confounded due to any of these (Moyer, 2007). It is also hard to find enough suitable pictures or topics to elicit the required words. The delayed repetition technique takes considerably longer than reading, and particularly for young learners might result in the production of incorrect items. Moreover, the subjects might adopt flaws in the modelling of the word stimuli. Being aware of the fact that reading word stimuli also tests reading ability and not just pronunciation, this technique was nevertheless adopted in thesis for the following reasons: It was not possible to find enough pictures or speech tasks to elicit the required words with the target sounds for these particular beginner groups. Furthermore, the time for data collection was limited due to school restrictions and the researcher could have only provided a non-native speech model. Due to the data

elicitation technique used here, students with reading disabilities were excluded from the study. Moyer (2007) demands that the task design needs to include a spectrum of different elicitation techniques. Unfortunately, this could not be applied in this thesis due to the twofold evaluation of the data that includes not only an auditory but also an acoustic analysis. As this latter analysis is not based on impressionistic measures, only one data elicitation technique was used.

In segmental phonetics the test items are often retrieved embedded in a carrier phrase, such as "*Please, say ….. for me*" to account for similar stress and duration patterns. However, as the target sounds occur in the test-items in word-initial, mid-word and word final position there might be coarticulation between the stimuli and the carrier phrase and it might be impossible to identify the segment boundaries (Podesva & Sharma, 2013). Moreover, words are usually hyperarticulated when they are produced in isolation. This leads to a reduction of coarticulation and to an expanded vowel space (see Section 5.2.1.2). It is for these reasons that no carrier phrase was used in the study.

7.4.4 Recording of the stimuli

To monitor any change due to pronunciation intervention, all participants were recorded before and directly after the intervention. In addition, a follow-up recording was taken five months after the post-test. The recordings were taken in a separate room at the respective participating schools. To take part in the research/ recording, the students were asked to leave their class. The entire task for each recording session took about ten minutes for each participant and the subject was seated facing a laptop monitor. The recordings were controlled by the Alvin2 software (Hillenbrand, James, 2012) that displayed the word stimuli to be read by the student. In total 168 words were presented on the laptop screen across the three sessions (see Figure 7-1). Each word production by the participant was saved as a separate sound file. The microphone was held by the student and a green bar in the Alvin2 program showed the student that his or her voice was at the right level for the recording.

Figure 7-1: Recoding of the stimuli

The speech samples were recorded and digitised at a 44.1 kHz sampling rate by an Olympus LS 11 digital recorder (Dresing, 2011a). This device was used as a converter and it was connected to a Samsung R519-Aura Darlio Notebook to synchronise the recording with the software program Alvin2. In some cases two recordings had to be made at the same time due to some unforeseen school events and the availability of the test subjects. In these cases a trained university student of English in her fourth year of study recorded the data. As only one Olympus LS 11 recorder was available, the stimuli were recorded with an Olympus LS 10 recorder using the same sampling rate. The data were then saved as a long sound file and later cut and edited with the speech analysis program Praat (Boersma, Weenik, 2012). The recordings took place at school during normal school days. Although the schools provided a separate room for the recording, there was some background sound due to noise of radiators, students running and talking in the hallways and in the schoolyard.

7.5 Data analysis

The previous sections of the current chapter examined the study design of this research project and looked at the participating parties as well as the speech material used to evaluate the pronunciation and the data collection procedures. Based on this description, this section now goes on to describe the two levels of data analyses. First, all stimuli were assessed auditorily using a three-point scale to evaluate the students' performances (see Section 7.5.1). Second, the data were analysed acoustically (see Section 7.5.2). The parameters in the analysis included fundamental and formant frequencies, duration, amplitude, spectral peak and spectral changes to gauge the pronunciation.

After the data collection, each word stimulus was saved as a separate sound file containing encrypted information about the participant, the test items itself, the repetition and the recording session. To ensure an objective and blind evaluation, all the data were then sorted by stimuli. Among all the 16 participants, for each of the 56 items there are 144 similar stimuli (16 children x 3 repetitions x 3 data collections points = 144).

7.5.1 Auditory Analyses

For the auditory analysis the 144 repetitions of each test items were randomly presented in one rating session so that the rater was able to familiarise herself with the pronunciation of a specific sound in a test item and to ensure an objective and blind evaluation. Therefore 56 sessions for each stimuli (see Table 7-4) took place and in total 8352 (58 x 144) items were rated. The data were presented via headphone at a comfortable level. The learners' sound production was auditorily analysed by the researcher herself[39]. Although usually native speakers are used to judge non-native speech as they are seen as the experts of the language (Moyer, 2007), this procedure was not applied in this thesis due to two reasons. First, in a PhD it is important that the researcher does the evaluation herself, second, the study deals with English pronunciation instruction in German schools. In this context, German English teachers need to be able to evaluate the students' pronunciation as well.

To achieve an objective assessment of the data analysis, two steps were applied. First, while evaluating the data, the researcher listened to the English model pronunciation of the items from Tim Bowyer (2006) on the free online "*Howjsay: Talking Dictionary of English Pronunciation*". Bowyer is a native English teacher who graduated at the London Institute of Education and is the founder of the `*Fonetiks´* family of websites (Bowyer, 2006). Second, the researcher provided judgement on the learner production of the sounds with the help of a 3-point rating scale ranging from zero (the sound was not produced), to one (the sound was not produced correctly) to two (no foreign accent) to ensure a similar rating of the presented test items. The rater was able to listen to each word as often as she wished and she was able to correct the judgement if needed.

[39] The researcher is a native speaker of German but speaks English at the C2 level of the Common European Framework of Reference (CEFR) and has received linguistic training in the field of research.

The ratings for each of the 56 stimuli were transcribed in a spreadsheet and for the three repetitions for each test periods (pre, post, follow-up) the sums were calculated and the 168 data points were transferred to SPSS. At this point information about the student such as age, sex, language, class and school background as well as the classification according to control and intervention group was added. Additionally, the item-bound information such as higher and lower frequency and trained and untrained was included. This set-up allows to sort and statistically analyse the data with reference to the demanded parameters and specific sounds. The researcher's values were used in the subsequent analyses.

7.5.1.1 Interrater judgement auditory analysis

To ensure a reliable analysis of the data, an interrater analysis was carried out on approximantly 10% (N = 864) of the data, which were selected semi-randomly. It included pre-, post- and follow-up-test data, and contained different sound categories as well as higher and lower frequency and trained and untrained items (see Table 7-5).

Table 7-5: Auditory test items judged by interrater

	Sound Category	Testitem		Frequency	Trained/ Not Trained
1	Vowel	melon	/melə̲n/	LF	NT
2	Diphthong	page	/pe̲ɪ̲dʒ/	HF	T
3	Plosive	egg	/ɛg̲/	LF	T
4	Fricative	month	/mʌnθ̲/	HF	NT
5	Affricate	cage	/keɪd̲ʒ̲/	LF	NT
6	Approximant	one	/w̲ʌn/	HF	T

In contrast to the researcher who is a non-native speaker of English, the rater is a monolingual native speaker of English. The rater took part in a training session to gain familiarity with the rating scale and the procedure. In order to reduce rater biases, the rater was blind to the nature of data she was coding (e.g. pre-test, control group, high frequency, trained).

A Spearman's rank-order correlation was run to assess the agreement between the judgements of the two raters. The outcomes showed that there was a moderate positive correlation between the two raters, r_s (862) = .520, p < .001.

As Cohen's κ is the more commonly used test for interrater reliability, it was also run to determine the rater's and interrater's judgements of the auditory data. The results show that there was a fair agreement between the two raters, κ = .276, 95% CI [.231, .321], p < .001.

Table 7-6: Interrater reliability auditory analysis (Cohen's κ)

Interrater * Rater Crosstabulation					
Count					
		Rater			Total
		0	1	2	
Rater 2	0	**50**	2	2	54
	1	26	**76**	7	109
	2	17	304	**380**	701
Total		93	382	389	864

The second rater seemed to be more generous with regards to a positive judgement (see Table 7-6).

Additional to the auditory evaluation of the data, an acoustic analysis of the same test items was carried out and is described in the following section (see Section 7.5.2).

7.5.2 Acoustic analyses

The acoustic analyses were carried out with the Praat software (Boersma & Weenik, 2012). For each of the 8064 audio data files (56 different test items x 16 participants x 3 repetitions at the 3 data collection points), text grids were generated in Praat. The segmentation of the sounds was done manually in Praat using the spectrogram and waveform of a speech sound. Praat scripts were customised and used to retrieve the required information and to generate the output. As discussed in Section 3.5, 14 sounds were selected for the intervention study and recorded within the context of 56 word stimuli (see Table 7-4). Due to their specific acoustic features the 14 sounds belong to six sound categories: vowels, diphthongs, plosives, fricatives, affricates and approximants. As each of these categories contain defining acoustic parameters (see Sections 5.2 to 5.4.4), distinct acoustic measurements have to be performed for each category (see Table 7-7 to Table 7-10). The mean values from the three repetitions at each of the three data collection periods were used for the final statistical analyses and data summaries.

7.5.2.1 Acoustic analysis of the vowels

Full formant frequency structure and voicing are inherent phonetic features of vowels in spectrograms, the vowel onsets and offsets were located using the full formant structure, characterised by a sequence of salient formant frequencies in a spectrogram and voicing as a cue for segmentation (see Figure 7-2). In addition, listening was used to determine the boundaries. To increase the accuracy throughout the analyses,

segmentation boundaries were usually placed at the point where the waveform crosses the amplitude axis (zero crossing). In case of transition phases, the boundaries were placed at the nearest zero crossing from the temporal midpoint of this area (Machač & Skarnitzl, 2009).

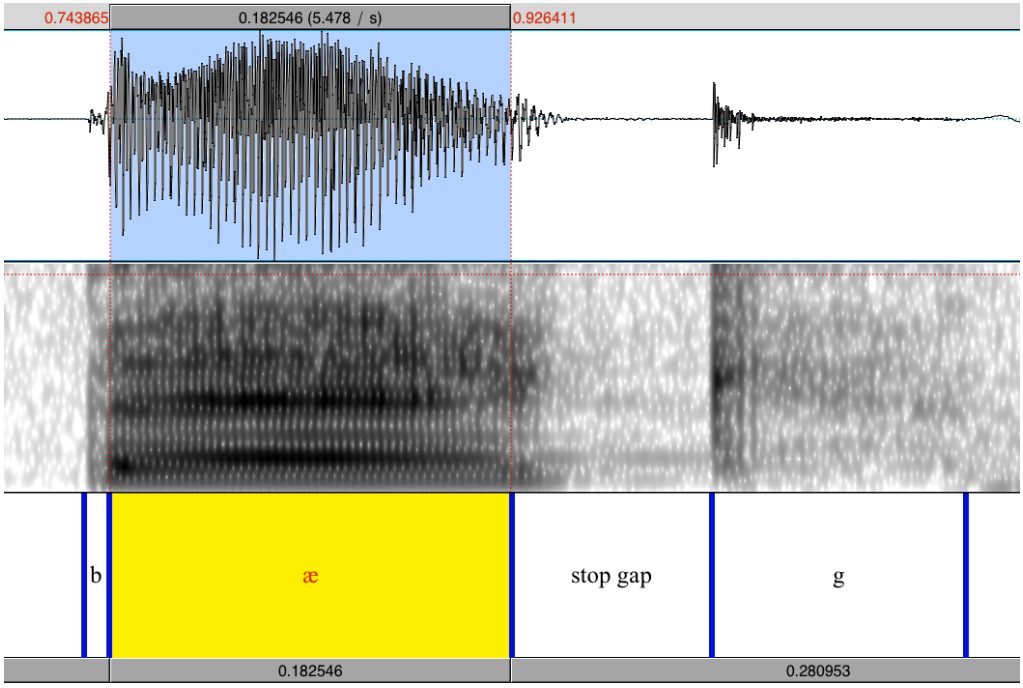


Figure 7-2: Sound segmentation vowel (bag)

The acoustic vowel measurements (see Table 7-7) included vowel duration (in ms) and F1 and F2 frequencies (in Hz) that were taken at the onset (0%) and the temporal midpoint (50%) of the vowel. When displayed in a vowel plot, F1 and F2 measurements can be used to determine vowel quality (see Section 5.2.1).

Table 7-7: Word context and properties of the vowel sounds /ɑ:/, /ɔ:/, /ə/, /æ/

	Sound	Testitem		HF/LF	trained	Acoustic measurements of the vowel
Vowels	/ɑ:/	arm	/ɑ:m/	HF	T	vowel duration F1 (0%,50%), F2 (0%, 50%)
		aunt	/ɑ:nt/	LF	T	
		dark	/dɑ:k/	HF	NT	
		scarf	/skɑ:f/	LF	NT	
	/ɔ:/	all	/ɔ:l/	HF	T	
		draw	/drɔ:/	LF	T	
		small	/smɔ:l/	HF	NT	
		chalk	/tʃɔ:k/	LF	NT	
	/ə/	from	/frəm/	HF	T	
		clever	/klevə/	LF	T	
		mother	/mʌðə/	HF	NT	
		melon	/melən/	LF	NT	
	/æ/	thank	/θæŋk/	HF	T	
		clap	/klæp/	LF	T	
		bag	/bæg/	HF	NT	
		maths	/mæθs/	LF	NT	

7.5.2.2 Acoustic analysis of the diphthongs and approximants (semi-vowels)

For diphthongs, articulation moves from one vowel to the next. Analogous to the segmentation of the vowels, the diphthong boundaries were manually placed at the zero crossing at the onset and offset of voicing and full formant structure. As the formant changes define diphthong quality (see Section 5.3) the diphthong measurements were taken at 20%, 40%, 60% and 80% of the diphthong to account for spectral change (see Table 7-8).

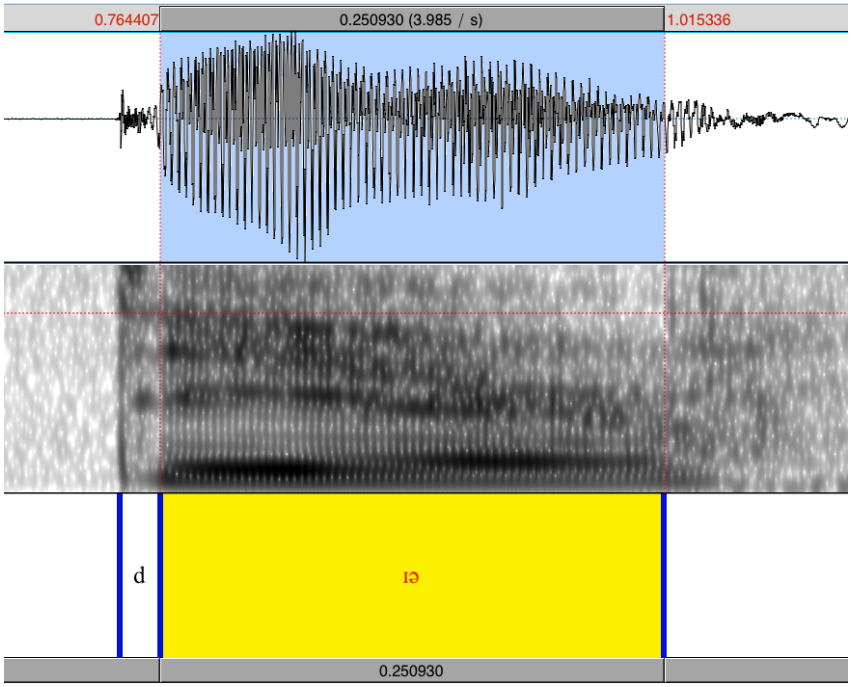


Figure 7-3: Segmentation diphthong (dear)

As the approximant leads into the following vowel, there is no definitive way of separating the approximant from the ensuing vowel in a waveform or spectrogram analysis. Thus, the measurements were taken at the onset of the approximant and the offset of voicing of the following vowel (see Figure 7-3). Similar to the diphthongs the place of articulation changes between the approximant and the vowel and thus the equivalent analyses used for diphthongs can be applied for the semi-vowels.

Table 7-8: Word context and properties of the diphthongs /ɪə/, /eɪ/ and the approximant /w/

	Sound	Testitem		HF/LF	Trained	Acoustic measurements of the diphthong/approximant
Diphthongs	/ɪə/	year	/jɪə/	HF	T	diphthong/ approximant-vowel duration F_1 (20%,40%, 60%, 80%) F_2 (20%,40%, 60%, 80%) Spectral rate of change
		dear	/dɪə/	LF	T	
		hear	/hiə/	HF	NT	
		beer	/bɪə/	LF	NT	
	/eɪ/	page*[40]	/peɪdʒ/	HF	T	
		grey	/greɪ/	LF	T	
		wait	/weɪt/	HF	NT	
		cage	/keɪdʒ/	LF	NT	
Approximants	/w/	one	/wʌn/	HF	T	
		swing	/swɪŋ/	LF	T	
		wait	/weɪt/	HF	NT	
		witch	/wɪtʃ/	LF	NT	

7.5.2.2.1 Euclidean distances

In cluster analyses, Euclidean distances are commonly used to measure the distances between points in an `n´-dimensional space (see Figure 7-4) using Pythagora's Theorem ($a^2+b^2= c^2$) (Halibisky, 1999).

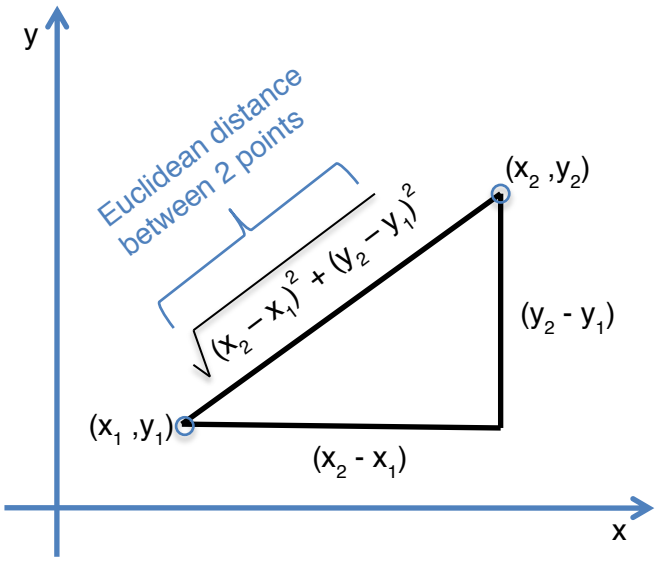


Figure 7-4: Euclidean distances (adapted from Halibisky, 1999)

Similarly, Euclidean distances are used in phonetic research to assess the distinctiveness of similar vowel qualities in a F1/ F2 plane in which the Euclidean distances quantifying the distance between each data point and the centroid of the entire cluster (Herrmann, 2010). Therefore, this distance can be used to measure the changes of vowel quality over time (Fox & Jacewicz, 2009).

[40] * high frequency for the age group

As the formant midpoint values are seen as the steady state of the vowel and the data is plotted in F1 x F2-F1 planes (see Section 4.2), in this thesis the Euclidean distances are calculated between the centroid and the vowel midpoint values of F1 and F2-F1 for each of the four vowels treated in the intervention (/ɑː/, /ɔː/, /ə/, and /æ/) in order to assess the changes in vowel quality over the three data collection points (t_0, t_1 and t_2) and between the intervention and control group.

$$Euclidean\ Distance = \sqrt{\big(vowel\ (F2-F1) - centroid(F2-F1)\big)^2 + (vowelF1 - centroidF1)^2}$$

7.5.2.2.2 Vector length (VL)

The Euclidean distance can be used to determine the length of a vector in a F1 / F2-F1 space (see Section 7.5.2.2.1). In the same way, it can be used to determine the length of formant movement by calculating the difference between the starting and end point of a diphthong. As the surrounding consonants have an immediate effect on the vowel formants, the measurements of the first and last 20% of the diphthong were discarded in this thesis (see Section 4.2). So, in order to assess the change in the formant movements over time, formant measurements were taken at four (20%, 40%, 60% and 80%) temporal points of the diphthong. The vector length (VL) is calculated as the Euclidean distance (in Hz) between the 20% and 80% temporal points of the diphthong in the F1 / F2-F1 plane (Fox & Jacewicz, 2009):

$$VL = \sqrt{(F1_1 - F1_4)^2 + ((F2-F1)_1 - (F2-F1)_4)}$$

7.5.2.2.3 Trajectory length (TL)

The vector length measures the overall magnitude of the formant movement between the beginning and end of the measurement. Diphthongs usually appear in the shape of a curve in the F1/ F2-F1 plane. However, the VL does not give any information about the curves in the formant tracks. Therefore, in order to look at the formant changes more closely over the course of the diphthong duration, the trajectory length can be calculated for each of the three measured sections of the diphthong; from 20% to 40%, from 40% to 60% and from 60% to 80%. The length of each section (VSL_N) is measured as follows:

$$VSL_n = \sqrt{(F1_n - F1_{n+1})^2 + ((F2-F1)_n - (F2-F1)_{n+1})}$$

The sum of all the three trajectory lengths of the three vowel sections can then described as the overall formant trajectory length (Fox & Jacewicz, 2009):

$$TL = \sum_{n=1}^{3} VSL_n$$

7.5.2.2.4 Spectral **r**ate **o**f **c**hange (roc)

The trajectory length provides detailed account of the formant change, but does not give any information on the amount of frequency change over time. Yet, the variation of spectral change over the diphthong's duration (di_dur) accounts for the differences in the diphthong's dynamic structure. Therefore, the spectral rate of change of the TL is calculated over the measured 60% portion (20% to 80%) of the diphthong.

$$TL\_roc = \frac{TL}{0.60\ x\ di\_dur}$$

To compare the sections of the diphthong between groups, the rate of change for each the three sections (VSL_roc) between the 20% and 40%, 40% to 60% and 60% to 80% temporal points can be measured (Fox & Jacewicz, 2009):

$$VSL\_roc_n = \frac{VSL_n}{0.20\ x\ di\_dur}$$

7.5.2.3 Acoustic analysis of the plosives

One of the main distinctions between the voiced and voiceless plosive in post-vocalic position is the duration of the preceding vowel, which is segmented as described above (see Section 7.5.2.1). Additionally, there is no noticeable fundamental frequency nor formant structure present in the occlusion phase of the voiceless plosive (Machač & Skarnitzl, 2009). Thus, to detect voicing, f0 is measured at 75% and 100% of the vowel preceding the post-vocalic plosive.

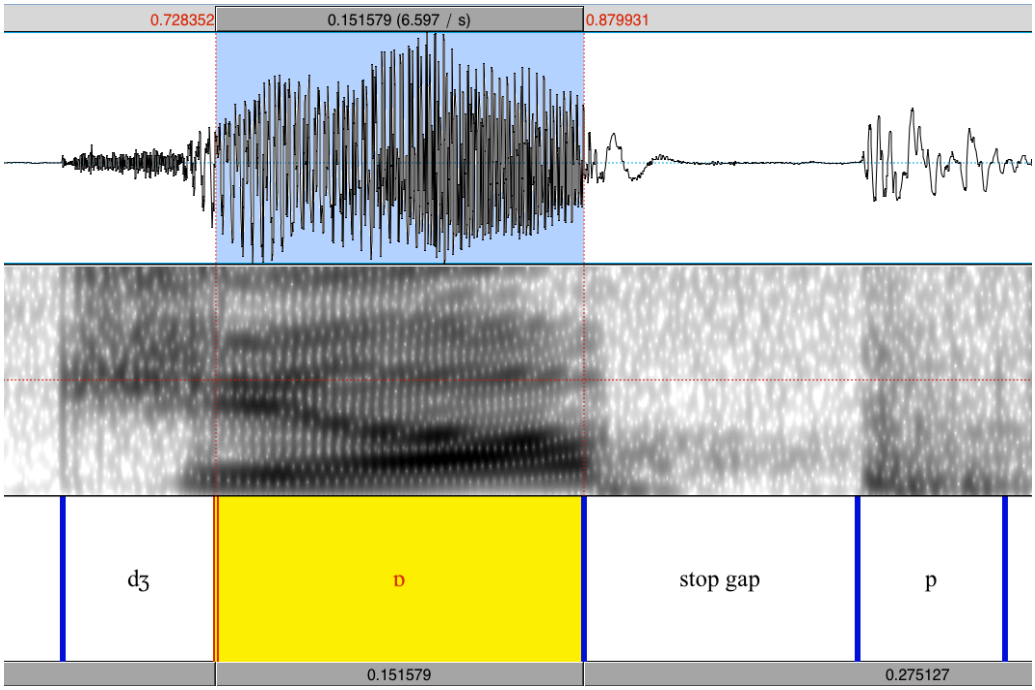


Figure 7-5: Segmentation plosive (job – pronounced as /dʒɒp/)

The absence of a formant structure throughout the plosive was used for its segmentation. As the release of a voiced plosive is not always visible, the stop gap and plosive duration could not always be measured. Therefore the duration of the preceding vowel and the absence or presence of fundamental frequency which shows voicing (see Section 5.4.2) are used as the main cues for the voicing distinction in this thesis (see Figure 7-5 and Table 7-9).

Table 7-9: Word context and properties of the plosives /b, d, g/ in final position

	Sound	Testitem		HF/ LF	Trained	Measurements of the preceding vowel	Measurements of the plosive
Plosive	/b/ final	job	/dʒɒb/	HF	T	vowel duration postvocalic f_0 (0%,25%)	plosive duration
		web	/web/	LF	T		
		club	/klʌb/	HF	NT		
		verb	/vɜb/	LF	NT		
	/d/ final	bad	/bæd/	HF	T		
		bread	/brɛd/	LF	T		
		child	/tʃaɪld/	HF	NT		
		gold	/gəʊld/	LF	NT		
	/g/ final	big	/bɪg/	HF	T		
		egg	/ɛg/	LF	T		
		bag	/bæg/	HF	NT		
		leg	/lɛg/	LF	NT		

7.5.2.4 Acoustic analysis of the fricatives and affricates

Fricatives and affricates can be segmented from vowels with the help of the onset and offset of a full formant structure. Once more, the vowel segmentation was carried out as described above in Section 7.5.2.1 (see Figure 7-6).

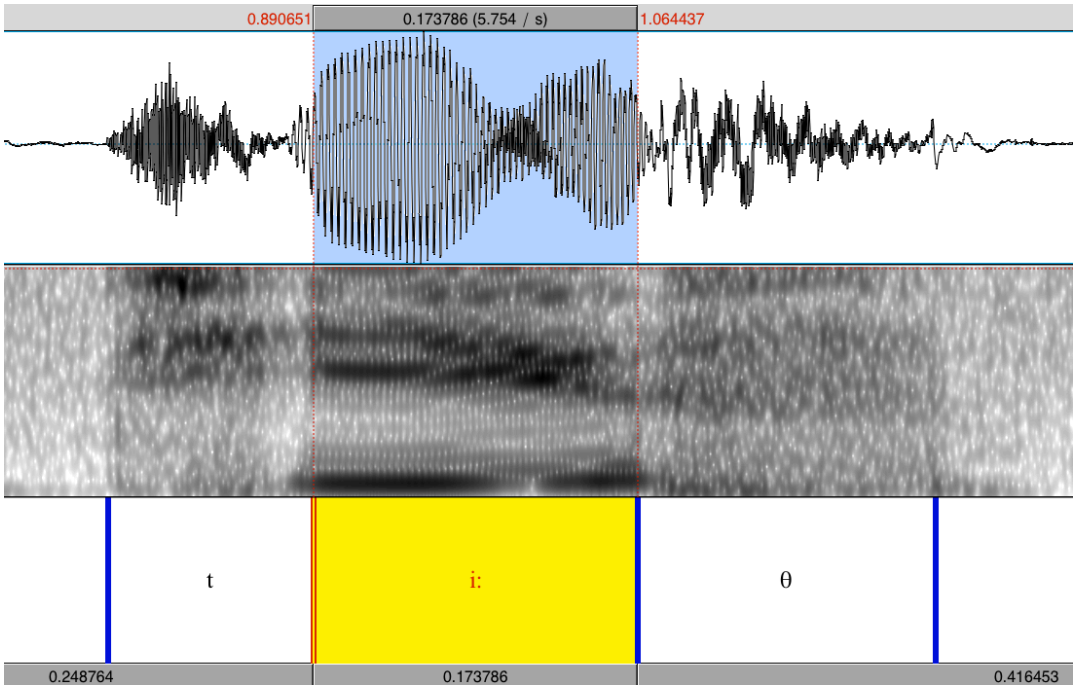


Figure 7-6: Segmentation fricative (teeth)

Yet again, to establish voicing, the duration of the vowel is taken into account as well as fricative length, as shorter fricative durations indicate voicing. Fricatives and affricates display an aperiodic character of the waveform and relative intensity differences can be seen in the spectrogram (see Figure 7-6). These features can be exploited as segmentation guidelines. The manner of articulation can be deduced from the centre of gravity (see Section 5.4.3). This approach models the fricative spectrum as a single normal distribution, which reflects the formant of the dominant front cavity that defines the fricative. The centre of gravity depicts the mean of this distribution (Wrench, 1995). Table 7-10 provides an overview over the word context, properties and acoustic measurements of the fricatives /z/, /ð/ and /θ/ and affricate /w/. The source-filter theory helps to explain the details of vowel spectra through the glottal sound source spectrum that is filtered by the vocal tract (Pickett, 1999) and this filter determines the frequency characteristics of a particular sound (see Figure 5-3) (Scarborough, 2005).

Table 7-10: Word context and properties of the fricatives /z, ð, θ/ and the affricate /w/

	Sound	Testitem		HF/LF	trained	Measurements of the adjacent vowel	Measurements of the fricative/affricate
Fricatives	/z/	close	/kləʊz/	HF	T	vowel duration prevocalic: f_0 (75%,100%) postvocalic: f_0 (0%, 25%)	fricative/ affricate duration centre of gravity
		jeans	/dʒiːnz/	LF	T		
		goes	/gəʊz/	HF	NT		
		zoo	/zuː/	LF	NT		
	/ð/	that	/ðæt/	HF	T		
		them	/ðem/	LF	T		
		the	/ðə/	HF	NT		
		smooth	/smuːð/	LF	NT		
	/θ/	thing	/θɪŋ/	HF	T		
		thin	/θɪn/	LF	T		
		month	/mʌnθ/	HF	NT		
		teeth	/tiːθ/	LF	NT		
Affricates	/dʒ/	job	/dʒɒb/	HF	T		
		jeans	/dʒiːnz/	LF	T		
		John	/dʒɒn/	HF	NT		
		cage	/keɪdʒ/	LF	NT		

7.5.2.5 Interrater judgement acoustic analysis

To ensure a reliable analysis of the data, a second rater analysed about 10% of the acoustic data, which were selected semi-randomly. It included pre-, post- and follow-up-test data and contained different sound categories as well as higher and lower frequency and trained and untrained items (see Table 7-11).

Table 7-11: Acoustic test items and measurements judged by rater and interrater

Measurements Interrater Analysis	Category	Test Item		Frequency higher vs lower	Training trained vs untrained
Vowel duration F1 midpoint (vowel)	Vowel	all	/ɔːl/	HF	T
Diphthong duration F1 midpoint (diphthong)	Diphthong	dear	/dɪə/	LF	T
Vowel duration F1 midpoint (vowel)	Plosive	club	/klʌb/	HF	NT
Vowel duration F1 midpoint (vowel)	Fricative	zoo	/zuː/	LF	NT
Vowel duration F1 (midpoint vowel)	Affricate	job	/dʒɒb/	HF	T
Approximant/ vowel duration F1 (midpoint approximant/ vowel)	Approximant	swing	/swɪŋ/	LF	T

In contrast to the rater who is a non-native speaker of English, the second rater is a monolingual native speaker of English with a BA in English language. Moreover, he has a good command of the software program Praat and wrote his BA thesis in the field of English phonetics. The rater took part in a training session to familiarise himself with the Praat setup, and with the segmentation guidelines applied by the rater (see

Machač & Skarnitzl, 2009). In order to reduce rater biases, he was blind to the characteristics of the data they were coding (e.g. pre-test, control group, higher frequency, not trained).

Although different sound categories are included in the data, the vowels in all test items play a crucial role for the acoustic data analyses. This does not only hold true in case of the vowels, diphthongs and semi-vowels (approximants), but also for the assessment of the voicing in plosives, affricates and fricatives (see Section 5.4). Therefore, vowel durations and F1 frequency values at the midpoint of the vowel for all interrater test items (see Table 7-11) were selected as suitable data points for the interrater analysis (see Table 7-11).

Table 7-12: Duration (in ms) and F1 (Hz)

Category	Measure	Rater		Interrater		Pearson Correlation	Sig 2-tailed
		Mean	SD	Mean	SD		
Vowel all	Vowel duration	285.9	75.64	343.3	85.87	.799	.000
	F1 midpoint (vowel)	702.78	118.47	698.64	116.99	.945	.000
Diphthong dear	Diphthong duration	351.1	67.86	377.5	69.82	.947	.000
	F1 midpoint (diphthong)	573.05	108.81	578.26	112.10	.982	.000
Plosive club	Vowel duration	142.4	50.70	152.8	51.84	.801	.000
	F1 midpoint (vowel)	850.44	173.32	846.25	172.99	.989	.000
Fricative zoo	Vowel duration	276.6	71.04	305.1	97.89	.707	.000
	F1 midpoint (vowel)	435.64	74.34	433.20	76.09	.919	.000
Affricate job	Vowel duration	191.5	51.29	210.4	55.93	.001	.990
	F1 (midpoint vowel)	744.35	97.02	741.49	101.75	.008	.929
Approxi-mant swing	Approximant/vowel duration	126.3	31.45	172.1	50.62	.644	.000
	F1 (midpoint approximant/ vowel)	542.62	86.54	536.50	82.76	.835	.000
Total	Durational measures	230.1	101.10	261.3	111.69	.874	.000
	F1midpoint meaures	642.16	179.98	639.76	179.75	.935	.000

A Pearson's correlation was run to assess the relationship between the vowel duration ratings between the rater and the second rater and also to evaluate the measurements for F1 at the vowel midpoint taken by the rater and the second rater. The outcomes are presented in Table 7-12. There was a strong positive correlation between the total vowel durations measured by the rater and the interrater, r (841) = .874, $p < .001$ and there was also a strong positive correlation between the total measurements for the first formant at vowel midpoint taken by the rater and the interrater, r (841) = .935, $p < .001$ indicating a very good agreement between the two raters. However, looking at the distinct categories, it is evident that there was no significant correlation between the first rater and second rater for the duration and formant frequencies of the test item `job´. An inspection of the annotated Praat data

revealed that the interrater often included parts of the affricate /dʒ/ in the segmentation of the vowel and therefore made systematic inaccuracies in the measurements. It is for this reason that the data for `job´ is included in the study despite the missing agreement between the first and second rater.

8. Results

This chapter presents the outcomes of a pronunciation intervention study in which 16 German students of English took part over the course of five months. As discussed in Chapter 7, half of the students received the intervention programme whereas the other half served as a control group. The participants came from eight different classes in two school types to account for differences in the education background. They were evenly divided according to sex, and their mean age at the entry point of the study was 11 years (SD = 0.32 years) (see Table 7-3 for an overview).

An independent-samples t-test was run to determine if there were age differences in the intervention and the control group which each contained eight participants. There were no outliers in the data, as assessed by the inspection of a boxplot. Engagement scores for the intervention and control group were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there was homogeneity of variances, as assessed by Levene's test for equality of variances (p = .152). The age was slightly higher in the intervention group (M = 11.06, SD = 0.26) than in the control group (M = 10.84, SD = 0.36). However, there was no statistically significant difference in the mean age between the intervention and the control group, M = 0.21, 95% CI [- 0.12 to 0.55], t(14) = 1.367, p = .193. This shows that the two groups were age-matched.

The pronunciation data were collected at three time points: A pre-test (t_0) was carried out prior to the intervention and a post-test (t_1) was done directly after the intervention. To control for long-term effects follow-up data were again collected five months after the study ($t_{2,}$). Besides time, the repeated measures also included measures that described features of the test items such as higher and lower frequency, as well as trained and untrained (see Table 7-4). The between subject measure was group (control and intervention). The same dataset was evaluated with two kinds of analyses. First, auditory analyses were applied and the results are presented in Section 8.1. and then discussed in Section 8.2. of this chapter. Second, Section 8.3 presents the results of the acoustic analyses. The chapter closes with a discussion of the results of the acoustic analyses (see Section 8.4).

8.1 Auditory Analyses

To evaluate the students' pronunciation performances, at each test period the students were asked to read the 56 (x 3 repetitions) test items that included the 14 targeted sounds (see

Table 7-4). All stimuli were assessed auditorily by a rater using a three-point (0,1,2) scale to evaluate the participants' performances. The sums of the means of each of the three repetitions with a possible maximum score of six were then used for the statistical data analysis (see Section 7.5.1). Table 8-1 presents the means and the standard errors of the auditory analysis for each data collection point by group, frequency and training status.

Table 8-1: Means and SE of the auditory analysis

Condition		N	Pre-test (t_0)		Post-test (t_1)		Follow-up-test (t_2)	
			Mean	SE	Mean	SE	Mean	SE
group	intervention (I)	448	3.88	.114	3.95	.111	4.15	.111
	control (C)	448	3.59	.114	4.01	.111	3.95	.111
frequency	higher frequency (HF)	448	3.89	.089	4.11	.076	4.16	.103
	lower frequency (LF)	448	3.58	.106	3.84	.098	3.93	.080
training	trained (IT)	448	3.85	.070	4.07	.066	4.10	.098
	not trained (NT)	448	3.62	.104	3.89	.099	4.00	.068

One of the aims of the intervention study is to find out whether different sounds pose different pronunciation problems to second language learners. Therefore, the following tables (Table 8-2 to Table 8-4) show the means and standard errors of the auditory analysis for each of the 14 sounds treated in the intervention by group status at each data collection point.

Table 8-2: Means and SE of the vowels and approximant of the auditory analysis

time	group	/ɑ:/		/ɔ:/		/ə/		/æ/		/w/	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	3.66	.23	3.31	.33	4.91	.25	4.22	.24	3.97	.24
	C	3.88	.23	3.59	.33	4.97	.25	4.41	.24	3.41	.24
t_1	I	3.56	.34	3.25	.30	4.72	.23	4.16	.28	4.44	.19
	C	4.28	.34	3.81	.30	5.22	.23	4.72	.28	4.44	.19
t_2	I	4.13	.28	3.94	.31	4.75	.23	4.09	.23	4.53	.27
	C	4.19	.28	3.78	.31	5.13	.23	4.41	.23	4.41	.27

Table 8-3: Means and SE of the diphthongs and plosives of the auditory analysis

time	group	/ɪə/		/eɪ/		/b/		/d/		/g/	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	3.66	.22	4.41	.25	3.38	.14	3.78	.21	4.06	.39
	C	3.41	.22	4.66	.25	3.47	.14	3.34	.21	3.44	.39
t_1	I	3.81	.28	4.09	.40	3.25	.23	3.69	.37	3.72	.27
	C	4.03	.28	4.94	.40	4.03	.23	3.69	.37	3.78	.27
t_2	I	4.03	.26	4.53	.39	3.41	.24	4.03	.28	4.09	.40
	C	3.47	.26	4.72	.39	3.97	.24	4.13	.28	3.59	.40

Table 8-4: Means and standard error of the fricatives and affricates of the auditory analysis

time	group	/z/		/ð/		/θ/		/dʒ/	
		Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	3.41	.13	3.84	.20	3.88	.33	3.97	.27
	C	3.06	.13	2.88	.20	2.53	.33	3.19	.27
t_1	I	3.50	.10	4.56	.16	4.19	.40	4.31	.32
	C	3.22	.10	3.22	.16	2.91	.40	3.84	.32
t_2	I	3.44	.12	4.38	.25	4.38	.29	4.34	.21
	C	3.16	.12	3.50	.25	3.13	.29	3.75	.21

In order to show whether a training effect takes place due to the pronunciation programme, half of the test items were explicitly trained in the intervention. Therefore, the following tables (Table 8-5 to Table 8-7) show the means and standard errors of the auditory analysis for each of the 14 sounds treated in the intervention by group and training status (**T**) at each data collection point.

Table 8-5: Means and standard error of the vowels and approximant of the auditory analysis by training status

time	group	T[41]	/ɑ:/		/ɔ:/		/ə/		/æ/		/w/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	IT	3.63	.25	3.44	.46	4.94	.30	4.13	.27	3.63	.23
		NT	3.69	.28	3.19	.26	4.88	.36	4.31	.31	4.31	.33
	C	IT	3.75	.25	3.63	.46	4.75	.30	4.81	.27	3.31	.23
		NT	4.00	.28	3.56	.26	5.19	.36	4.00	.31	3.50	.33
t_1	I	IT	3.56	.33	3.38	.42	4.63	.25	4.44	.33	4.19	.31
		NT	3.56	.39	3.13	.27	4.81	.36	3.88	.30	4.69	.36
	C	IT	4.19	.33	3.75	.42	5.31	.25	4.56	.33	4.69	.31
		NT	4.38	.39	3.88	.27	5.13	.36	4.88	.30	4.19	.36
t_2	I	IT	4.25	.32	4.00	.36	4.56	.32	4.31	.27	4.00	.22
		NT	4.00	.41	3.88	.38	4.94	.32	3.88	.33	5.06	.37
	C	IT	4.25	.32	3.63	.36	5.06	.32	4.50	.27	4.69	.22
		NT	4.13	.41	3.94	.38	5.19	.32	4.31	.33	4.13	.37

Table 8-6: Means and standard error of the diphthongs and plosives of the auditory analysis by training status

time	group	T	/ɪə/		/eɪ/		/b/		/d/		g/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	IT	3.56	.19	5.25	.26	3.50	.24	3.94	.24	4.06	.39
		NT	3.75	.30	3.56	.49	3.25	.12	3.63	.24	4.06	.41
	C	IT	3.38	.19	5.25	.26	3.88	.24	3.25	.24	3.50	.39
		NT	3.44	.30	4.06	.49	3.06	.12	3.44	.24	3.38	.41
t_1	I	IT	3.81	.34	4.31	.37	3.56	.28	3.75	.37	3.69	.37
		NT	3.81	.30	3.88	.57	2.94	.25	3.63	.44	3.75	.27
	C	IT	4.06	.34	5.25	.37	4.19	.28	3.75	.37	3.81	.37
		NT	4.00	.30	4.63	.57	3.88	.25	3.63	.44	3.75	.27
t_2	I	IT	3.69	.24	4.81	.42	3.63	.33	4.19	.29	4.06	.41
		NT	4.38	.36	4.25	.47	3.19	.17	3.88	.37	4.13	.42
	C	IT	3.50	.24	5.00	.42	4.31	.33	3.69	.29	3.69	.41
		NT	3.44	.36	4.44	.47	3.63	.17	4.56	.37	3.50	.42

[41] **T** = Training, IT = trained in the intervention, NT= not trained in the intervention

Table 8-7: Means and standard error of the fricatives and affricate of the auditory analysis by training status

time	group	T	/z/		/ð/		/θ/		/dʒ/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	IT	3.25	.10	4.00	.25	3.81	.24	4.38	.27
		NT	3.56	.19	3.69	.22	3.94	.64	3.56	.36
	C	IT	2.94	.10	3.13	.25	3.63	.24	3.19	.27
		NT	3.19	.19	2.63	.22	1.44	.64	3.19	.36
t_1	I	IT	3.19	.10	4.75	.19	4.44	.30	4.69	.30
		NT	3.81	.19	4.38	.29	3.94	.62	3.94	.50
	C	IT	2.94	.10	3.50	.19	3.94	.30	3.75	.30
		NT	3.50	.19	2.94	.29	1.88	.62	3.94	.50
t_2	I	IT	3.13	.07	4.63	.32	4.19	.33	4.63	.28
		NT	3.75	.22	4.13	.28	4.56	.45	4.06	.29
	C	IT	2.94	.07	3.88	.32	3.69	.33	3.81	.28
		NT	3.38	.22	3.13	.28	2.56	.45	3.69	.29

There is evidence that word frequency affects speech production (see Section 5.5). It is for this reason that Table 8-8, Table 8-9 and Table 8-10 show the means and standard errors of the auditory analysis for each of the 14 sounds treated in the intervention by group and frequency status (**F**) at each data collection point.

Table 8-8: Means and standard error of the vowels and approximant of the auditory analysis by training status

time	group	F[42]	/ɑ:/		/ɔ:/		/ə/		/æ/		/w/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	HF	3.63	.33	3.56	.30	4.69	.27	4.56	.30	4.13	.27
		LF	3.69	.33	3.06	.44	5.13	.32	3.88	.34	3.81	.28
	C	HF	4.38	.33	4.25	.30	5.06	.27	4.63	.30	3.56	.27
		LF	3.38	.33	2.94	.44	4.88	.32	4.19	.34	3.25	.28
t_1	I	HF	3.44	.40	3.81	.31	4.44	.27	4.50	.31	4.69	.22
		LF	3.69	.36	2.69	.50	5.00	.30	3.81	.33	4.19	.31
	C	HF	4.38	.40	4.81	.31	5.25	.27	5.19	.31	4.38	.22
		LF	4.19	.36	2.81	.50	5.19	.30	4.25	.33	4.50	.31
t_2	I	HF	4.06	.34	4.13	.38	4.50	.32	4.44	.32	4.94	.27
		LF	4.19	.33	3.75	.47	5.00	.31	3.75	.28	4.13	.38
	C	HF	4.25	.34	4.38	.38	4.94	.32	4.50	.32	4.56	.27
		LF	4.13	.33	3.19	.47	5.31	.31	4.31	.28	4.25	.38

[42] **F** = Frequency, HF = higher frequency test items, LF = lower frequency test items

Table 8-9: Means and standard error of the diphthongs and plosives of the auditory analysis by frequency status

time	group	F	/ɪə/		/eɪ/		/b/		/d/		/g/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	HF	3.69	.29	4.63	.24	3.63	.30	3.75	.33	4.00	.39
		LF	3.63	.22	4.19	.37	3.13	.13	3.81	.21	4.13	.42
	C	HF	3.63	.29	4.63	.24	3.69	.30	3.75	.33	3.44	.39
		LF	3.19	.22	4.69	.37	3.25	.13	2.94	.21	3.44	.42
t_1	I	HF	3.88	.36	4.25	.34	3.50	.28	3.69	.36	3.63	.25
		LF	3.75	.35	3.94	.53	3.00	.30	3.69	.43	3.81	.35
	C	HF	4.38	.36	5.06	.34	4.19	.28	3.88	.36	3.63	.25
		LF	3.69	.35	4.81	.53	3.88	.30	3.50	.43	3.94	.35
t_2	I	HF	4.25	.35	4.69	.40	3.38	.27	4.19	.29	4.25	.44
		LF	3.81	.28	4.38	.45	3.44	.29	3.88	.32	3.94	.39
	C	HF	3.69	.35	4.44	.40	4.25	.27	4.19	.29	3.69	.44
		LF	3.25	.28	5.00	.45	3.69	.29	4.06	.32	3.50	.39

Table 8-10: Means and standard error of the fricatives and affricate of the auditory analysis by frequency status

time	group	F	/z/		/ð/		/θ/		/dʒ/	
			Mean	SE	Mean	SE	Mean	SE	Mean	SE
t_0	I	HF	3.25	.14	4.31	.28	3.69	.37	4.19	.30
		LF	3.56	.23	3.38	.31	4.06	.36	3.75	.41
	C	HF	2.94	.14	3.31	.28	2.56	.37	3.50	.30
		LF	3.19	.23	2.44	.31	2.50	.36	2.88	.41
t_1	I	HF	3.13	.10	5.13	.27	4.00	.39	4.81	.27
		LF	3.88	.20	4.00	.21	4.38	.48	3.81	.52
	C	HF	2.94	.10	3.75	.27	2.56	.39	3.94	.27
		LF	3.50	.20	2.69	.21	3.25	.48	3.75	.52
t_2	I	HF	3.06	.09	5.06	.34	4.38	.35	4.69	.27
		LF	3.81	.22	3.69	.28	4.38	.36	4.00	.30
	C	HF	2.94	.09	3.75	.34	3.19	.35	3.88	.27
		LF	3.38	.22	3.25	.28	3.06	.36	3.63	.30

The auditory data were analysed using SPSS (IBM Corp., 2013). All effects are reported as significant at $p < .05$. A mixed-model ANOVA[43] was carried out to determine whether there were any main effects and interaction effects for the repeated measures time, frequency and training between the intervention and control group (see Table 8-11). The Greenhouse-Geisser correction was used to detect violations of sphericity. Kesselmann et al. (1980) consider Mauchly's test of sphericity a poor method to detect violations of sphericity as it often fails to detect departures from sphericity in small samples (which are used in this study). Therefore, Maxwell & Delaney (2003) suggest ignoring the result of Mauchly's test and use the Greenhouse-Geisser correction. This method is applied in this study and the following table shows the significant between and within subjects effects for the repeated measures after the Greenhouse-Geisser correction.

[43] There were no outliers in the data, as assessed by studentized residuals (no residuals ≥ ±3 SD. Levene' s test showed that for 154 out of the 168 mean testitems homogeneity of variance ($p > .05$) could be established. Shapiro-Wilk's test showed that most ratings were normally distributed. As the mixed ANOVA is somewhat robust to deviations from normality, sphericity and homogeneity of variance, violations are no exclusion criteria for this test (Lund Research Ltd, 2013),

Table 8-11: Significant within subjects and between subjects effects

Measures	df	Error[44]	F	Sig.	Partial Eta Squared
group (intervention/ control)	1.00	14	1.28	.277[45]	.084
time (pre-/ post-/ follow-up-test)	1.70	23.86	8.39	**.003**	.375
time x group	1.70	23.86	2.77	.090[46]	.165
training (trained/ untrained)	1.00	14	9.13	**.009**	.395
frequency (higher/ lower frequency)	1.00	14	9.65	**.008**	.408
sound	5.60	78.46	9.17	**.000**	.396
sound x group	5.60	78.46	3.53	**.005**	.202
time x training	1.98	27-78	2.61	.092[43]	.157
time x frequency x group	1.82	25.44	2.79	.085[43]	.166
training x frequency	1.00	14.00	3.93	.067[43]	.219
training x sound	5.20	72.93	4.32	**.001**	.236
training x sound x group	5.20	72.83	2.67	**.027**	.160
frequency x sound	6.43	90.04	5.47	**.000**	.281
training x frequency x sound	5.33	74.68	7.12	**.000**	.337

The mixed model ANOVA revealed (see Table 8-11) that there were statistically significant main effects for time, training, frequency and sound. Moreover, there were significant interaction effects for sound x group, training x sound, training x sound x group, frequency x sound as well as training x frequency x sound. In addition, the interactions between time x group, time x training, time x frequency x group and training x frequency approached significance. The following sections look at the main and interaction effects in more detail.

8.1.1 Main effect: Group

There was no main effect between the two groups over all data collection points (F (1, 14) = 1.28, p = .277) with the intervention group (M = 3.994, SE = .091) receiving only slightly higher ratings compared to the control group (M = 3.849, SE = .091). However, the mixed model ANOVA revealed (see Table 8-11) a trend between the interaction between time x group (F (1,7, 23.86) = 2.77, p = .090) and significant interactions between group x sound (F (5.60, 78.46)= 3.53, p = .005) and group x training x sound (F (5.2, 72.83) = 2.67, p = .027).

8.1.2 Main effect: Time

There was a significant main effect of the pronunciation ratings at the different time points during the intervention (F (1.7, 23.86) = 8.39, p < .003) with the pronunciation ratings

[44] The error is fractioned due to the fact that the error is calculated according to the Greenhouse-Geisser correction
[45] The between factor group does not show significance but is included for reference.
[46] These subject measures are included as they approach significance.

increasing from pre-intervention (M[47] = 3.74, SE = .08) to post-intervention (M = 3.98, SE = .078) to follow-up intervention (M = 4.05, SE = .078), in that order (see Figure 8-1). Post hoc analysis with a Bonferroni adjustment revealed that pronunciation ratings statistically significantly increased from pre-intervention to post-intervention (*M*[48] = 0.24, 95% CI [0.074, 0.406], p = .004) and from pre-intervention to follow-up-intervention (*M* = 0.311, 95% CI [0.077, 0.546], p = .009), but not from post-intervention to follow-up-intervention (*M* = 0.071, 95% CI [- 0.170, 0.312], p = .089).

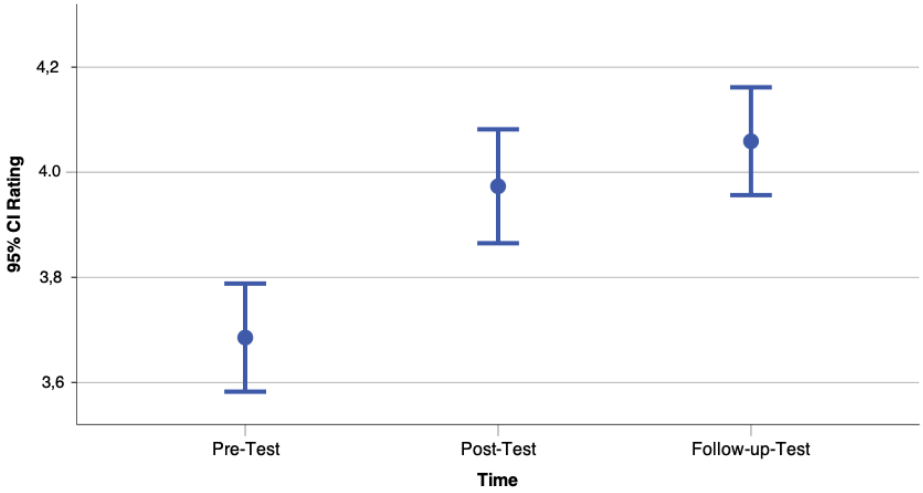


Figure 8-1: Mean ratings over time

8.1.3 Main effect: Training

There was also a statistically significant effect for trained (M = 4.01, SE = .064) vs. untrained test items (M = 3.84, SE = .076; F (1.000, 14) = 9.13, p < .009). Post hoc tests with a Bonferroni adjustment revealed that trained items had higher pronunciation ratings than untrained items (*M* = 0.172, 95% CI [0.050, 0.294], p = .009).

8.1.4 Main effect: Frequency

Frequency also showed a statistically significant main effect (F (1.000, 14.00) = 9.65, p < .008). Post hoc tests with a Bonferroni adjustment showed that higher pronunciation ratings were given for higher frequency items (M = 4.06, SE = .076) than for lower frequency items (M = 3.786, SE = .079; *M* = 0.272, 95% CI [0.084, 0.459], p < .008).

[47] M is used for `mean´, and (Lund Research Ltd, 2013)
[48] In this thesis *M* is used for `mean difference´ in order to differentiate it from the M for `mean´.

8.1.5 Main effect: Sounds

There was also a statistically significant main effect of sound (F (5.60, 78.46) = 9.17, p < .000). The mean ratings are presented in Figure 8-2.

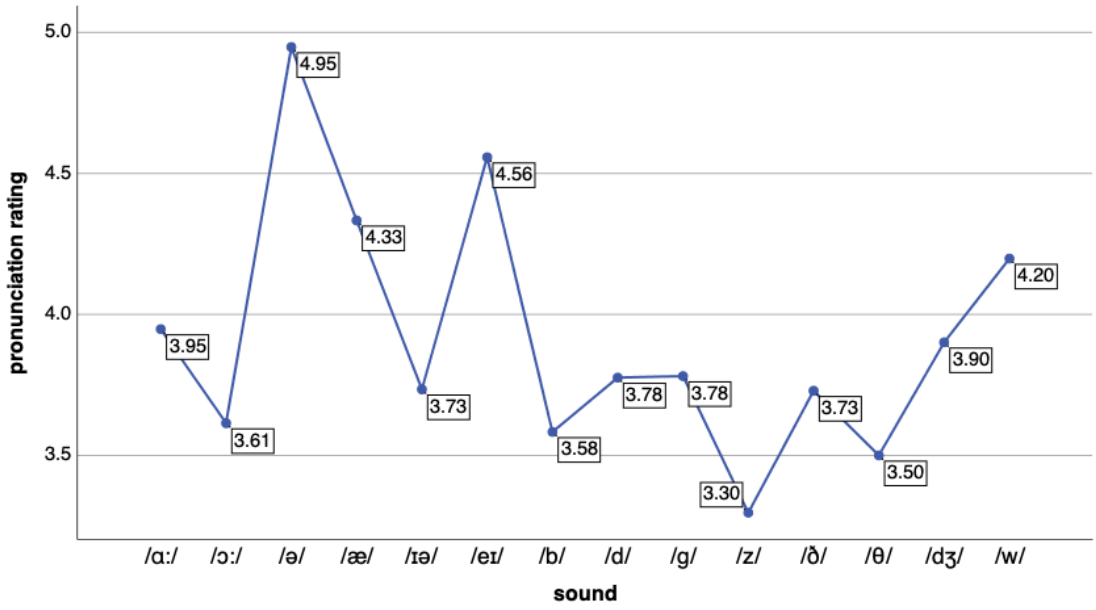


Figure 8-2[49]: Mean sound ratings

Post hoc tests with Bonferroni adjustments revealed the following differences between the **vowel /ə/** (M = 4.95, SE = .141), which shows the highest rating and the:

- **vowel /ɑ:/ (**M = 3.95, SE = .178), (*M* = 1.00, SE = .194, p < .014);
- **vowel /ɔ:/** (M = 3.61, SE = .197), (*M* = 1.33, SE = .133, p < .000);
- **vowel /æ/** (M = 4.33, SE = .132), (*M* = 0.61, SE = .135, p < .042);
- **diphthong /ɪə/** (M = 3.73, SE = .130), (*M* = 1.21, SE = .169, p < .000);
- **plosive /b/** (M = 3.58, SE = .119), (*M* = 1.36, SE = .163, p < .000);
- **plosive /d/** (M = 3.78, SE = .142), (*M* = 1.17, SE = .197, p < .003);
- **plosive /g/** (M = 3.78, SE = .220), (*M* = 1.17, SE = .230, p < .016);
- **fricative /z/** (M = 3.30, SE = .072), (*M* = 1.65, SE, .124 p < .000);
- **fricative /ð/** (M = 3.73, SE = .101), (*M* = 1.22, SE = .181, p < .001);
- **fricative /θ/** (M = 3.50, SE = .188), (*M* = 1.45, SE = .215, p < .001);

[49] Despite the fact that the variables are not continuous, line charts are used in this section for interpretation and legibility purposes. This is especially relevant for the following interaction graphs, as non-parallel lines and line crossings are indicators for significant interactions (Field 2009).

- **affricate /dʒ/** (M = 3.90, SE = .146), (*M* = 1.05, SE = .199, p < .011); and
- **approximant /w/** (M = 4.20, SE = .130), (*M* = 0.75, SE = .164, p < .040).

Post hoc tests with Bonferroni adjustment also showed a difference between the **approximant /w/** (M = 4.20, SE = .130) and the:

- **fricative /z/** (M = 3.30, SE = .072), (*M* = 0.90, SE = .104, p < .000).

In addition to the main effects for time, training, frequency and sound presented above, the mixed model ANOVA revealed a trend (see Table 8-11) between the interaction between time x group (F (1,7, 23.86) = 2.77, p = .090). It also revealed statistically significant interactions between sound x group F (5.604, 78.46) = 3.537, p < .005), training x sound (F (5.20, 72.93) = 4.32, p = .001), training sound and group (F (5.20, 72.83) = 2.67, p = .027), frequency x sound (F (6.43, 90.04) = 5.47, p = .000) and training x frequency x sound (F (5.33, 74.68) = 7.12, p = .000). The following sections looks at the interaction effects in more detail.

8.1.6 Interaction between group x time

Despite the fact that there was no main effect for group, the ANOVA showed a trend between the interaction group and time (F (1,7, 23.86) = 2.77, p = .090)). A closer inspection of Table 8-1 and Figure 8-3 reveal that the mean ratings for the intervention group at the pre-test (M = 3.89, SE = .114) were slightly higher than for the control group (M = 3.56, SE = .114). A t-test run on the pre-test data revealed no significant difference between the two groups (p = .082). Contrariwise, the mean data showed higher ratings for the control group at the post-test (M = 4.01, SE = .111) in comparison to the intervention group (M = 3.95, SE = 111). The ratings in the control group decreased between the post-test and follow-up-test (M = 3.95, SE = .111) whereas the data shows a strong increase between the post-test (M = 3.95, SE = 111) and the follow-up-test (M = 4.15, SE .111; see Figure 8-3) in the intervention group. A t-test showed a significant difference between the two groups in the follow-up test (p = .048).

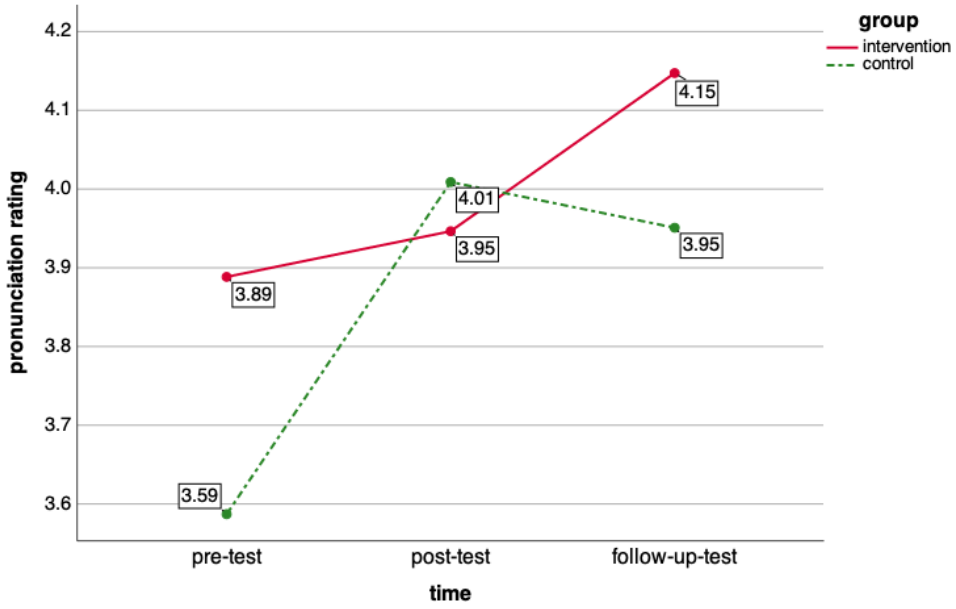


Figure 8-3: Mean pronunciation ratings for group x time

8.1.7 Interaction between sound x group

There was a statistically significant interaction effect (see Table 8-11) between sound and group over all three data collection points (F (5.604, 78.46) = 3.537, p < .005). To examine this interaction, a simple effects analysis was carried out to compare the pronunciation ratings of each sound between the intervention and the control group (see Figure 8-4).

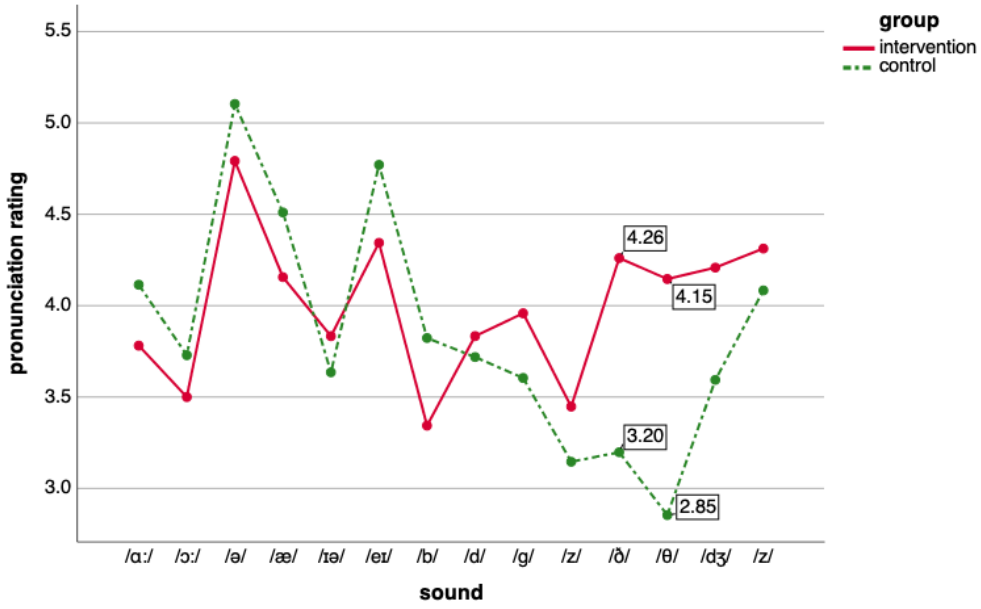


Figure 8-4: Interaction between sound and group

The analysis revealed that there were significant differences[50] in the ratings for the **fricative /ð/** between the intervention group (M = 4.26, SE = .142) and the control group (M = 3.20, SE = .455; *M* = 1.06, SE = .201, p < 0.00). The same holds true for the ratings given for the **fricative /θ /** in the intervention group (M = 4.15, SE = .266). These were also significantly higher than for the control group (M = 2.85, SE = .266; *M* = 1.29, *SE* = .376, *p* = .004). Additionally, the differences in the ratings for the **plosive /b/** (*M* = .479, SE = .238, p = .064) and the **affricate /dʒ/** (*M* = .615, SE = 291, p = .053) approached significance.

8.1.8 Interaction between sound and training

There was also a significant interaction effect (see Table 8-11) between the pronunciation ratings of the different sounds in the trained and untrained test items for both groups over all three data collection points (F (5.20, 72.93) = 4.32, p = .001).

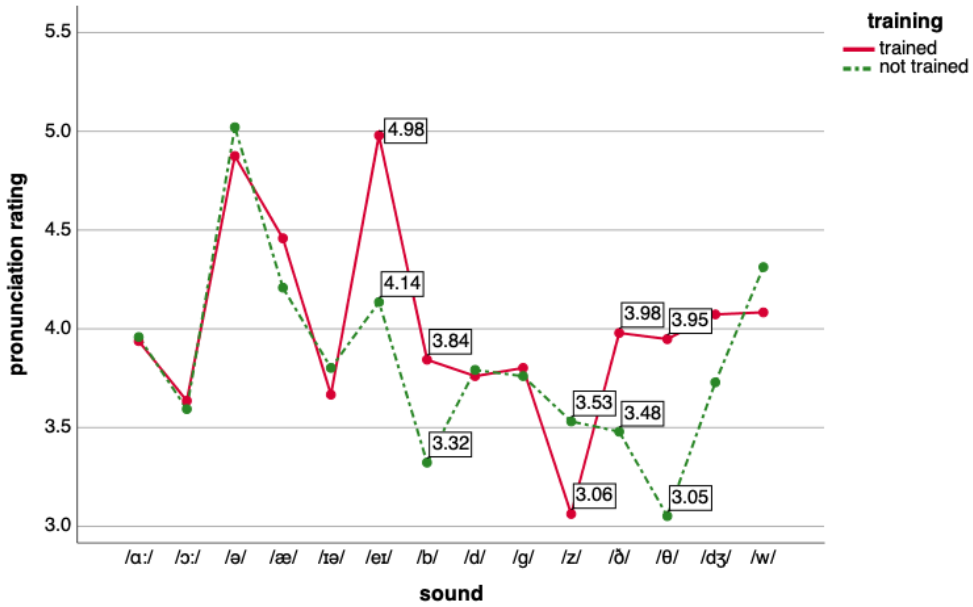


Figure 8-5: Interaction between sound and training

The simple effects analysis (see Figure 8-5) showed that the **diphthong /eɪ/** received higher mean ratings for the trained (M = 4.98, SE = .189) compared to the untrained (M = 4.14, SE = .3.12) test items (*M* = .844, SE = .302, p = .014). The **final plosive /b/** also showed significantly higher ratings for trained test items (M = 3.84, SE = .169) than for untrained test items (M = 3.32, SE = .085; *M* = .521, *SE* = .123, *p* = .001). Similarly, the **fricative /ð/** received higher ratings for the trained test items (M = 3.98, SE = .108) compared to the untrained test

[50] Numbers are given for the significant comparisons in the diagrams.

items (M = 3.48, SE = .144; *M* = .500, SE = .156, p = .006). The same holds true for the ratings of the **fricative /θ/** that were significantly higher for trained items (M = 3.95, SE = .149) compared to the ratings for the untrained test items (M = 3.05, SE = .329; *M* = .896, SE = .345, p = .021).

Contrary to the assessment of the aforementioned sounds, the ratings for the **fricative /z/** were lower for the trained test items (M = 3.06, SE = .058) compared to the untrained test items (M = 3.53, SE = .120; *M* = - .469, SE = .120, p. = .002).

The results presented above included the ratings for both groups. The following section 8.1.8.1 shows the interaction for sound and training according to intervention and control group status.

8.1.8.1 Interaction between sound, training and group

There were significant interaction effects between sound x training x group (F (5.20, 72.83) = 2.67, p = .027).

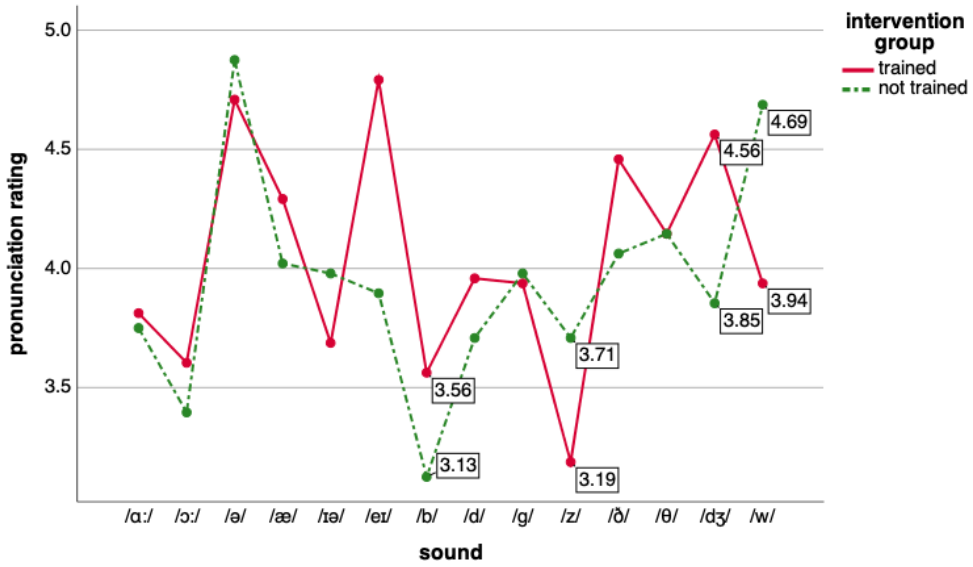


Figure 8-6: Interaction between training x sound x group (intervention group)

The simple effects analysis (see Figure 8-6) showed that the **final plosive /b/** received significantly higher ratings for the trained test items (M = 3.56, SE = .240) compared to the untrained test items (M = 3.13, SE = .120) in the intervention group (*M* = .438, SE = .174, p = 0.025). The ratings for the **affricate /dʒ/** also displayed significantly higher ratings for trained test items (M = 4.56, SE = .198) than for untrained test items (M = 3.85, SE = .289; *M* = .708, *SE* = .276, p = .022). Moreover, the simple effects analysis also approached

significance for the **diphthong /eɪ/** with higher mean ratings for the trained (M = 4.79, SE = .268) compared to the untrained (M = 3.90, SE = .441) test items (*M* = .896, SE = .426, p = .054).

On the contrary, the ratings for the **fricative /z/** were lower for the trained test items (M = 3.19, SE = .081) in contrast to the untrained test items (M = 3.71, SE = .170; *M* = - .521, SE = .170, p. = .008). Similarly, the **approximant /w/** received lower ratings for the trained items (M = 3.94, SE = .159) in comparison to the untrained items (M = 4.69, SE = .308; *M* = -.750, SE = .324, p = .036).

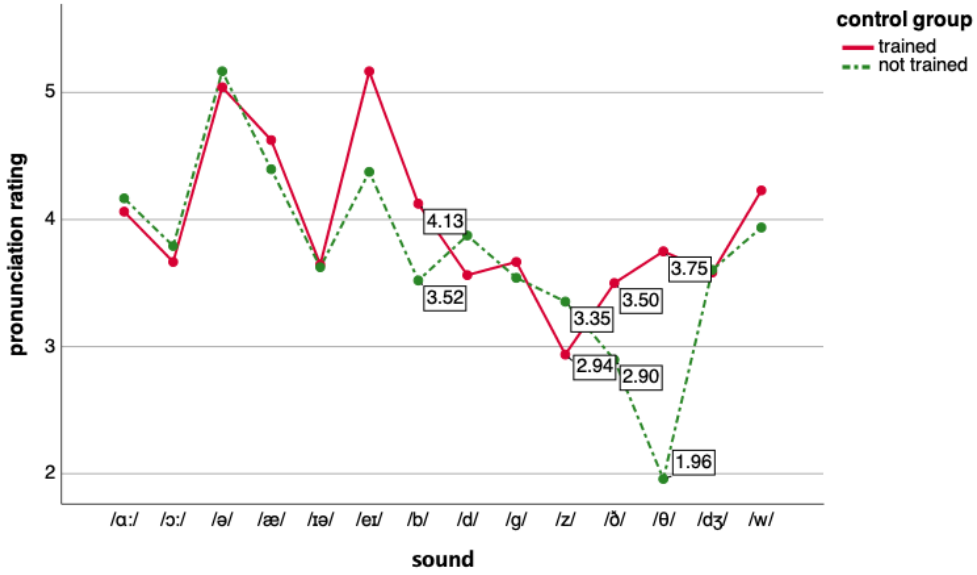


Figure 8-7: Interaction between training x sound x group (control group)

With regard to the control group, the simple effects analysis (see Figure 8-7) showed that the **plosive b in final position** received significantly higher ratings for the trained test items (M = 4.13, SE = .240) compared to the untrained test items (M = 3.52, SE = .120; *M* = .604, SE = .174, p = .004). The **fricative /ð/** also showed higher ratings for the trained test items (M = 3.50, SE = .152) in comparison to the untrained test items (M = 2.90, SE = .204; *M* = 1.31, SE = .220, p = .016). Similarly, the **fricative /θ/** received higher ratings for the trained test items (M = 3.75, SE = .211) compared to the untrained test items (M = 1.96, SE = .465; *M* = 1.792, SE = .488, p = .003).

However, the **fricative /z/** displayed significantly lower pronunciation ratings for the trained test items (M = 2.94, SE = .081) in comparison to the untrained test items (M = 3.35, SE = .170; *M* = -. 417, SE = .170, p = .028).

8.1.9 Interaction between sound and frequency

There were significant interaction effects (see Table 8-11) between the mean ratings of the sounds in higher and lower frequency test items (F (6.43, 90.04) = 5.47, p = .000) over all three data collection points.

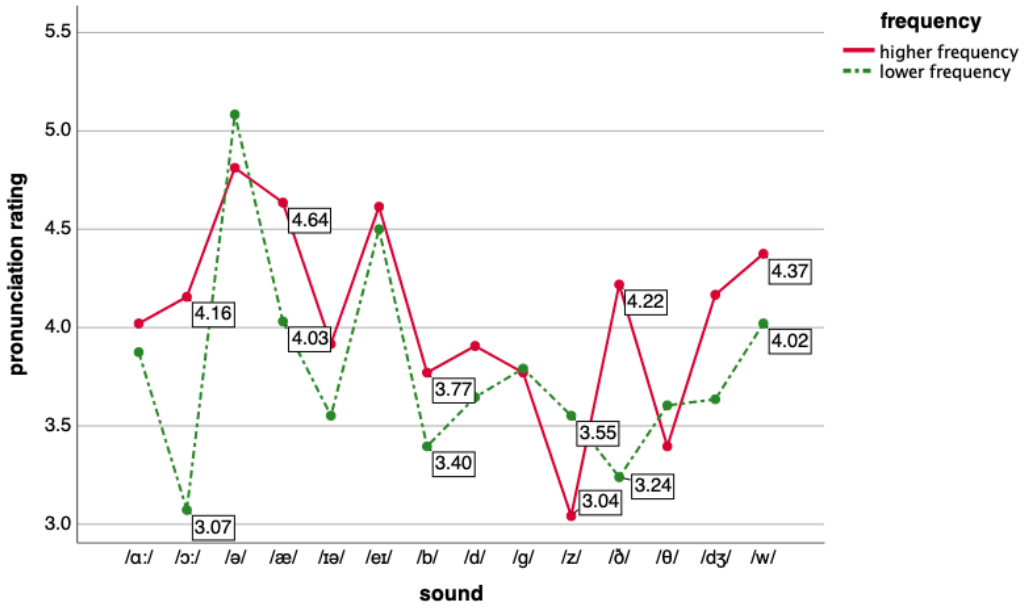


Figure 8-8: Interaction between sound x frequency

The simple effects analysis (see Figure 8-8) revealed that the **vowel /ɔ:/** showed higher ratings for the higher frequency test items (M = 4.16, SE = .200) compared to the lower frequency test items (M = 3.07, SE = .293; *M* = 1.083, SE .311, p = .004). Similarly, the **vowel /æ/** received higher ratings for the higher frequency test items (M = 4.64, SE = .177) in comparison to the lower frequency (M = 4.03, SE = .166) test items (*M* = .375, SE = .221, p = .016). Moreover, the **plosive /b/ in final position** had significantly higher pronunciation ratings for higher frequency test items (M = 3.77, SE = .154) than for lower frequency test items (M = 3.40, SE = .129, *M* = .375, SE = 155, p = .030). The **fricative /ð/** also received significantly higher ratings for higher frequency test items (M = 4.22, SE = .177) compared to lower frequency test items (M = 3.24, SE = 128; *M* = .979, SE = .234, p = .001). Equally, the **approximant /w/** had significantly higher ratings for higher frequency test items (M = 4.37, SE = .135) in comparison to lower frequency test items (M = 4.02, SE = .171; *M* = .354, SE = .164, p = 0.049).

In contrast to the sounds mentioned above, the **fricative /z/** received lower ratings for higher frequency test items (M = 3.04, SE = .069) compared to lower frequency test items (M = 3.55, SE = .130; *M* = - .510, SE = .150, p = .004).

The results presented above include the frequency ratings according to the different sounds. The following Section 8.1.9.1 shows this interaction for sound and frequency according to the training status.

8.1.9.1 Interaction between sound, training and frequency

The interaction between sound, training and frequency (see Table 8-11) showed significant effects (F (5.33, 74.68) = 7.12, p = .000).

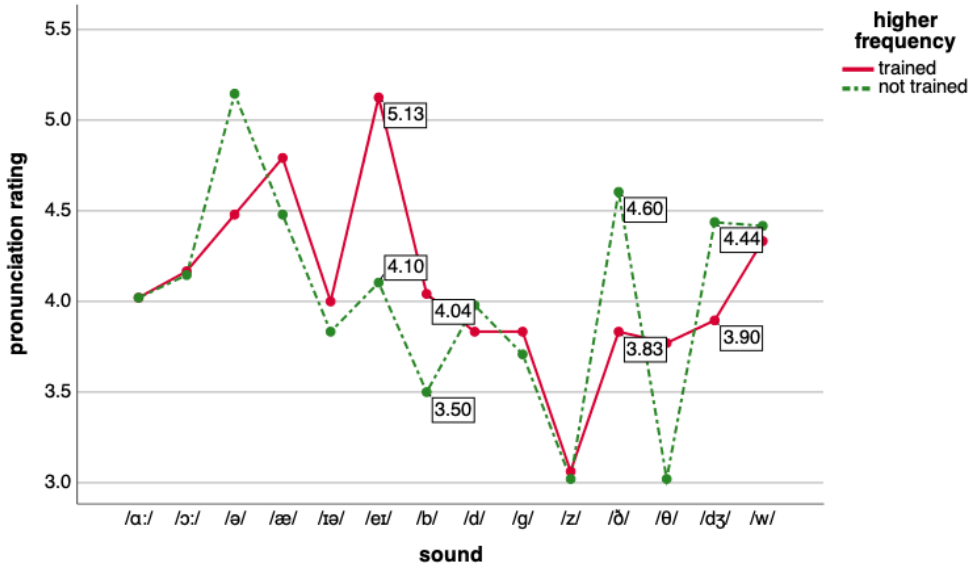


Figure 8-9: Interaction between sound x frequency (higher frequency) x training

Among the higher frequency test items (see Figure 8-9), the pronunciation of the **diphthong /eɪ/** showed significantly higher ratings for the trained test items (M = 5.13, SE = .179) compared to the untrained test items (M = 4.10, SE = .274; *M* = 1.021, SE = .295, p = .001). Similarly, the **plosive /b/ in final position** received higher ratings for the trained test items (M = 4.04, SE .218) in comparison to the untrained test items (M = 3.50. SE = .133; *M* = .542, SE = .188, p = .012) among the higher frequency test items.

Contrary to the two sounds mentioned above, the **fricative /ð/** also showed lower ratings for the trained test items (M = 3.83, SE = .181) compared to the untrained test items (M = 4.60, SE = .248; *M* = -.771, SE = .252, p = .008). Similarly, the **affricate /dʒ/** received significantly lower ratings for the trained test items (M = 3.90, SE = .155) in comparison to

the untrained test items for the higher frequency status (*M* = 4.44, SE = .226; *M* = -.542, SE = .242, p = .042).

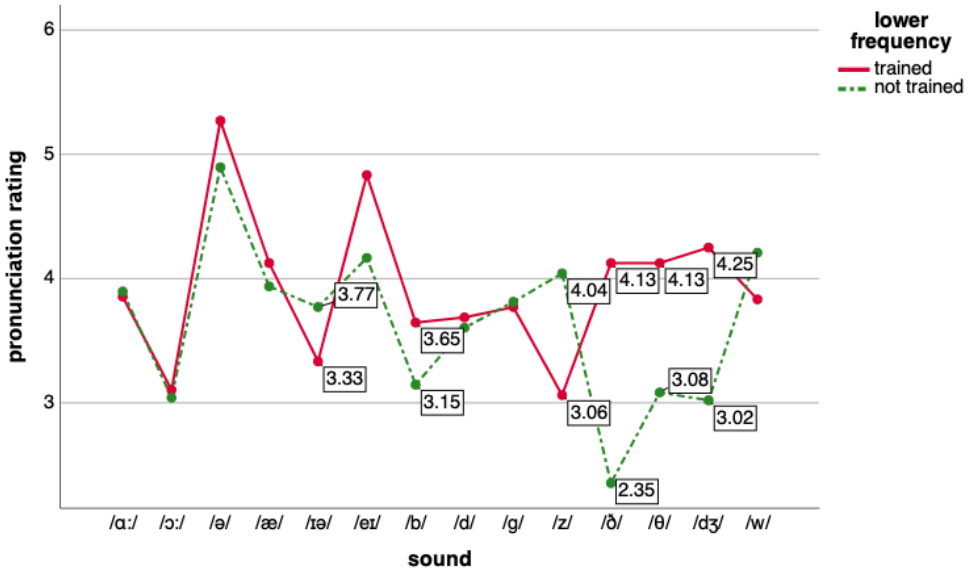


Figure 8-10: Interaction between sound x frequency (lower frequency) x training

Among the lower frequency test items (see Figure 8-10) the pronunciation of the **plosive /b/ in final position** showed significantly higher ratings for the trained test items (M = 3.65, SE = .189) compared to the untrained test items (M = 3.15, SE = .1.04; *M* = .500, SE = .163, p = .008). The **fricative /ð/** also received significantly higher ratings for the trained test items (M = 4.13, SE = .171) in comparison to the untrained test items (M = 2.35, SE = .2.03, *M* = 1.77, SE = .286, p = .000). Similarly, the **fricative /θ/** displayed higher ratings for the trained test items (M = 4.13, SE = .274) than for the untrained test items (M = 3.08, SE = .359, *M* = 1.042, SE = .374, p = .015). The **affricate /dʒ/** also received significantly higher ratings for the trained test items (M = 4.25, SE = .180) compared to the untrained test items (*M* = 3.02, SE = .426) among the lower frequency items (*M* = 1.229, SE = .451, p = .016).

Contrary to the sounds mentioned above, the **diphthong /ɪə/** got significantly lower ratings for the trained test items (M = 3.33, SE = .151) in comparison to the untrained test items (M = 3.77, SE = .194; *M* = -.438, SE = .151, p = 0.012). Equally, the **fricative /z/** displayed lower ratings for the trained test items (M = 3.06, SE = .044) compared to the untrained test items (M = 4.04, SE = .248; *M* = - 9.79, SE = .242, p = 0.01).

Thus far, the previous section has presented the outcomes of the auditory analysis which are discussed in the following section 8.2. Chapter 9 then goes on to present the conclusions that can be drawn from the results.

8.2 Discussion auditory analysis

The auditory analysis begins with an overview of the mean pronunciation ratings for intervention, frequency and training at the pre-, post- and follow-up-test periods. Table 8-1 shows that the ratings for all conditions increased with time with the exception for the mean ratings of the control group, that show a slight decrease between post- and follow-up-test. This overall improvement was expected as all subjects matured over time and also received ongoing language input in their English lessons. In order to have a closer look at the data and to find out whether the presented differences and changes are significant the following sections discuss the output of the mixed model ANOVA.

The main effect of time revealed that the mean ratings increased significantly between pre- and post-test and then also showed a slight increase between post- and follow-up-test (see Figure 8-1). Thus, it can be deduced that the highest increase in pronunciation abilities took place directly after the intervention and that a maintenance phase took place between the post- and follow-up-test. To find out whether this improvement might be a result of the pronunciation training programme, the data need a closer inspection.

Prior to the intervention study the ratings for the intervention and control group should be fairly similar, as the intervention had not yet started, and therefore there should not be a difference between the two groups. However, Table 8-1 and Figure 8-4 reveal that the mean ratings for the intervention group at the pre-test (M = 3.88, SE = .114) were already slightly higher than for the control group (M = 3.56, SE = .114) showing that there might be some bias between the intervention and control group although the groups were assigned semi-randomly (see Section 7.2.4). However, a t-test run on the pre-test data revealed that there was no significant difference between the two groups (p = .082). The data also showed that the significant increase (see Figure 8-1) in the pronunciation rating between pre- and post-test mainly takes place in the control group (see Figure 8-3) which did not receive the intervention. Moreover, the ratings in the control group decreased between the post-test (M = 4.01, SE = .111) and follow-up-test (M = 3.95, SE = .111). Therefore, the significant main effect of time between pre- and post-test cannot be fully attributed to the training programme as the mean ratings in the intervention group increased only slightly between pre-test (M = 3.88, SE = .114) and post-test (M = 3.95, SE = 111) and then showed a stronger increase between post-test and follow-up-test (M = 4.15, SE .111; see Figure 8-3). However, a t-test revealed that the difference between the two groups at the follow-up test was significant (p = 0.48) revealing higher ratings for the intervention group. The portrayed mean data does not

give any information with regards to the development of the different sounds, which are therefore presented in the following paragraph.

The main effects analysis showed a strong significant effect of sound (see 8.1.5). The sound category contains the 14 different sounds trained in the intervention (see Section 3.5). The mean pronunciation ratings (see Figure 8-2) revealed that the **vowel /ə/** (M = 4.95, SE = .141), the **diphthong /eɪ/** (M = 4.56, SE = .209) and the **approximant /w/** (M = 4.20, SE = .130) received the three highest scores. In contrast, the **vowel /ɔ:/** (M = 3.61, SE = .197), the **plosive /b/** (M = 3.58 , SE = .119), the **fricatives /θ/** (M = 3.50, SE = .188) and **/z/** (M = 3.30, SE = .072) had the lowest three mean ratings and therefore seemed harder to pronounce. The diverse mean scores of the sounds support the assumption, that different sounds pose different problems to language learners. However, this will be discussed further in the conclusion in Chapter 9. Moreover, the presented scores comprise the numbers from both groups over all three data collection points. In order to assess this matter, a further examination of the significant interaction between sound x group (see Section 8.1.6) revealed that the **fricatives /ð/** and **/θ**/ seemed to improve significantly through the intervention with significantly higher mean ratings for **/ð/** (M = 4.26, SE = .142) and **/θ/** (M = 4.15, SE = .266) in the intervention than in the control group (/ð/: M = 3.20, SE = .142 and /θ/: M = 2.85, SE = .266). The interaction graph (see Figure 8-4) also showed much higher ratings for the **affricate /dʒ/** in the intervention in comparison to the control group. However, this interaction only approached significance (p = .053). The sound x group interaction does not take the three data collection points into account. To find out when the changes presented above take place the means for each sound at each data collection point have to considered.

A look at the descriptive means (see Table 8-2, Table 8-3 and Table 8-4) shows that the scores for the **fricative /ð/** strongly increase from 3.84 (SE = .20) in the pre-test to 4.56 (SE = .16, *M* = .72) in the post-test and then slightly decrease to 4.38 (SE = .25, *M* = -.18) in the follow-up-test in the intervention group. The mean ratings in the control group increase from 2.88 (SE = .20) to 3.22 (SE = .16, *M* = .34) to 3.50 (SE = .25, *M* = .028) over all three data collection points. With regards to **/θ/,** the mean ratings increase from 3.88 (SE = .33) to 4.19 (SE = .40, *M* = .31) to 4.38 (SE = .29, *M* = .19) in the intervention group. The control group data increases from 2.53 (SE = .33) at the pre-test to 2.91 (SE = .40, *M* = .38) at the post-test to 3.13 (SE = .29, *M* = .22) at the follow-up test. The mean ratings for **/dʒ/** in the intervention group show an increase in the mean ratings from 3.97 (SE = .27), to 4.31 (SE = .32, *M* = .34) to 4.34 (SE = .21, *M* = .03) over all three data collection points. The control group data for **/dʒ/** strongly increases from 3.19 (SE = .27) in the pre-test to 3.84 (SE = .32,

M = .65) in the post-test and then slightly decrease to 3.75 (SE = .21, *M* = -.09) in the follow-up-test. For the presented fricatives **/ð/** and **/θ**/ and the affricate **/dʒ/**, which are not part of the German sound inventory the biggest increase takes place between pre-and post-test. This holds true for both groups.

To assess whether a change in the pronunciation performance is due to the intervention programme or just take place because of a training effect, half of the test items were explicitly trained in the pronunciation programme (see Section 7.3.3). Thus, this paragraph looks at the training status. The main effect of training revealed that the trained test items had significantly higher mean pronunciation ratings (M = 4.01, SE = .064) in comparison to the untrained test items (M = 3.84, SE = .076; see 8.1.3). In order to understand which sounds in particular were affected by the training, the interaction between sound x training needs to be considered. Figure 8-5 and Section 8.1.8 show that the pronunciation of the **diphthong /eɪ/**, the final **plosive /b/** and the **fricatives /ð/** and **/θ/** seemed to profit significantly from the training effect. However, the **fricative /z/** displayed significantly lower pronunciation ratings in the trained dataset.

To find out whether the presented results were actually caused by the training effect of the intervention, the significant interactions between training x sound x group (see Section 8.1.6) have to be taken into account. As the control group did not receive a special treatment for the trained test items, ideally the ratings of the trained and untrained test items should be fairly similar and indeed Figure 8-7 shows that the ratings for the **vowels /ɑ:/, /ɔ:/, /ə/** and **/æ/,** the **approximant /w/**, the **diphthong /ɪə/**, the **plosives /d/** and **/g/** and the **affricate /dʒ/** do not differ a lot between trained and untrained items. Nevertheless, despite the ratings for the **diphthong /eɪ/** that does not show a significant difference between trained and untrained status (see Figure 8-7), the results of the control group are in concordance with the main effect showing higher ratings for the **plosive /b/** and the **fricatives /ð/ and /θ/** and lower ratings for the untrained items that entail the **fricative /z/**. Although the classification of the test items into trained and untrained status was done semi-randomly (see Section 7.3) there might be confounds in the sorting. Consequently, the test items that contained **/b/, ð/, /θ/ and /z/** and showed significant training effects in the control group need to be examined. The trained items that included the plosive /b/ in final position are `job´ and `web´, and the untrained items are `club´ and `verb´. The significant training effect might be due to the lower frequency untrained item `verb´. Although the same word exists in the German language, the subjects struggled with reading it. Moreover, the stimuli `job´ and `web´ are used as loanwords in the German language and are therefore probably frequently used by the

students taking part in the intervention. The **fricative /ð/** is included in the trained items `that´ and `them´ and in the untrained items `the´ and `smooth´. Whereas the voiced fricative in the trained items `that ´ and `them´ do not produce many problems, `smooth´ seems to be a lot harder to pronounce due to the fact that the articulation of the voiceless fricative and nasal cluster followed by the vowel makes it a rather complex utterance. Moreover, it became apparent in the testing, that several children did not seem to know the word despite being part of their textbook vocabulary, and struggled with reading it. The trained items containing the voiceless **fricative /θ/** are `thing´ and `thin´ and the untrained items are `month´ and `teeth´. `Month´ might have led to production challenges as it includes a nasal fricative cluster and therefore its level of difficulty might be higher than that of the other items. The **fricative /z/** received lower ratings for the trained compared to the untrained items. The trained items are `close´ and `jeans´ and the untrained, `goes´ and `zoo´. Similar to `verb´ and `month´ (see above), the lower ratings might be due to the item `jeans´ that contains a nasal fricative cluster and it is prone to transfer from its German devoiced counterpart (/dʒiːnz/ vs. dʒiːns/). Moreover, the stimulus `goes´ belongs to one of the first words the students acquire in the foreign language classroom and is thus used very frequently. Despite the different levels of difficulties in the test items mentioned above, also random effects of the small data set might also play a role in the different ratings between the trained and untrained items in the control dataset.

The presented outcomes of the interaction between training x sound x group entail the scores of all three data collection points. Therefore, a closer look at the descriptive data might reveal more information on the above mentioned test items that contained **/b/, ð/, /θ/ and /z/** and showed significant training effects in the control group. The following scores are extracted from Table 8-5, Table 8-6 and Table 8-7 and show the descriptive mean values (M) and mean differences (*M*) for the presented sounds by trained (IT) and untrained(NT) status at the three data collection points (t_0, t_1 and t_2).

/b/(IT): $M(t_0) = 3.88$ (SE = .24) $M(t_1) = 4.19$ (SE = .28, *M* = .31) $M(t_2) = 4.31$ (SE = .33, *M* = .12)
/b/(NT): $M(t_0) = 3.06$ (SE = .12) $M(t_1) = 3.88$ (SE = .25, *M* = .82) $M(t_2) = 3.63$ (SE = .17, *M* = -.25)

/ð/(IT): $M(t_0) = 3.13$ (SE = .25) $M(t_1) = 3.50$ (SE = .19, *M* = .37) $M(t_2) = 3.88$ (SE = .32, *M* = .38)
/ð/(NT): $M(t_0) = 2.63$ (SE = .22) $M(t_1) = 2.94$ (SE = .29, *M* = .31) $M(t_2) = 3.13$ (SE = .28, *M* = .19)

/θ/(IT): $M(t_0) = 3.63$ (SE = .24) $M(t_1) = 3.94$ (SE = .30, *M* = .31) $M(t_2) = 3.69$ (SE = .33, *M* = -.25)
/θ/(NT): $M(t_0) = 1.44$ (SE = .64) $M(t_1) = 1.88$ (SE = .62, *M* = .44) $M(t_2) = 2.56$ (SE = .45, *M* = .68)

/z/(IT): $M(t_0) = 2.94$ (SE = .10) $M(t_1) = 2.94$ (SE = .10, *M* = 0) $M(t_2) = 2.94$ (SE = .07, *M* = 0)
/z/(NT): $M(t_0) = 3.19$ (SE = .19) $M(t_1) = 3.50$ (SE = .19, *M* = .31) $M(t_2) = 3.38$ (SE = .22, *M* = -.12)

The three sounds **/b/, /ð/ and /θ/** of the control group show a similar pattern for the trained test items (IT) and the untrained test items (NT) with the highest increase in the ratings scores between pre-test and post-test and only a slight increase or decrease between post-test and follow-up-test with the exception of **/θ/** in the untrained dataset. For **/θ/**(NT) the scores increase between pre- and post-test and then show the strongest increase between post-test and follow-up test. Moreover, not only for the three sounds but also for the whole dataset, the trained items already display higher mean ratings (M = 3.85, SE = .070) at the pre-test (M = 3.62, SE = .104, see Table 8-1). However, these scores should be fairly similar prior to the intervention study as the intervention had not yet started and therefore there should not be a difference between the two groups and trained and untrained items. This initial difference continues from t_0, t_1 and t_2 with overall higher mean scores for the trained dataset in the control group. The reason behind this pattern might be due to the input of the ongoing English lessons that also treat these sounds (see Section 2.2) or a confound in the assignment of the test items to trained and untrained status, although the classification was done semi-randomly (see Section 7.3.3).

The **fricative /z/** behaves differently to all the other test items with overall lower ratings for the trained test items in comparison to the untrained test items in the control group. Moreover, the ratings for /z/ in the trained dataset are fairly similar despite ongoing English lessons at all data collection points whereas the scores in the untrained dataset increase between pre- and post-test and then show a slight decrease at the follow-up test.

To find out whether the intervention programme leads to a training effect, the significant interaction between training x sound x group for the intervention group has to be considered. Figure 8-6 reveals that the **plosive /b/** and the **affricate /dʒ/** received significantly higher ratings in the trained data set in comparison to the untrained data set. On the contrary, the **fricative /z/** and the **approximant /w/** got significantly lower ratings in the trained data set compared to the untrained data set. However, it should be noted that the presented scores comprise the data from pre-test, post-test and follow-up-test. Therefore, a closer look at the descriptive data might reveal more information on the above mentioned test items that contained **/b/, /dʒ/, /w/ and /z/** and showed significant training effects in the intervention group. The following scores are extracted from Table 8-5, Table 8-6 and Table 8-7 and show the descriptive mean values (M) and mean differences (*M*) for the presented sounds by trained (IT) and untrained (NT) status at the three data collection points (t_0, t_1 and t_2).

/b/(IT): $M(t_0) = 3.50$ (SE = .24) $M(t_1) = 3.56$ (SE = .28, $M = .06$) $M(t_2) = 3.63$ (SE = .33, $M = .07$)
/b/(NT): $M(t_0) = 3.25$ (SE = .12) $M(t_1) = 2.94$ (SE = .25, M = -.31) $M(t_2) = 3.19$ (SE = .17, $M = .25$)

/dʒ/(IT): $M(t_0)$ = 4.38 (SE = .27) $M(t_1)$ = 4.69 (SE = .30, *M* = .31) $M(t_2)$ = 4.63 (SE = .28, *M* = -.06)
/dʒ/(NT): $M(t_0)$ = 3.56 (SE = .36) $M(t_1)$ = 3.94 (SE = .50, *M* = .38) $M(t_2)$ = 4.06 (SE = .29, *M* = .12)

/w/(IT): $M(t_0)$ = 3.63 (SE = .23) $M(t_1)$ = 4.19 (SE = .31, *M* = .56) $M(t_2)$ = 4.00 (SE = .22, *M* = -.19)
/w/(NT): $M(t_0)$ = 4.31 (SE = .33) $M(t_1)$ = 4.69 (SE = .36, *M* = .38) $M(t_2)$ = 5.06 (SE = .37, *M* = .37)

/z/(IT): $M(t_0)$ = 3.25 (SE = .10) $M(t_1)$ = 3.19 (SE = .10, *M* = -.06) $M(t_2)$ = 3.13 (SE = .07, *M* = -.06)
/z/(NT): $M(t_0)$ = 3.56 (SE = .19) $M(t_1)$ = 3.81 (SE = .19, *M* = .25) $M(t_2)$ = 3.75 (SE = .22, *M* = -.06)

The results for the trained test items that entail the **plosive /b/** do not show a lot of change between the three data collection points. On the contrary, the untrained dataset shows a strong decrease between pre- and post-test and again an increase at the follow-up-test. The trained and untrained test items that contain the **affricate /dʒ/** behave similarly over the three data collection points with a strong increase between pre-test and post-test and a slight increase or decrease at the follow-up-test. For both sounds the trained test items display higher mean ratings than the untrained items. However, this difference is already apparent at the pre-test which might indicate a bias in the set-up of the test items.

Despite overall higher mean ratings for the trained items the **approximant /w/** and the **fricative /z/** display overall lower mean ratings for the test items that were explicitly trained in the intervention (see Table 8-10). The results for the trained and untrained test items that entail the **approximant /w/** display both a strong increase after the intervention. Although both groups display this change, the increase in the intervention group is so large that it might be a result of the intervention. However, whereas the trained items show a decrease between post-test and follow-up test, the untrained test items also show a strong increase between t_1 and t_2 and a very high pronunciation rating at the final follow-up-test. Similar to the results of the control group (see above), the **fricative /z/** does not display a lot of change in the ratings for the trained test items at the three data collection points (see Table 8-10). However, the untrained test items display an increase between pre- and post-test and a slight decrease between post-test and follow-up-test. This pattern with overall lower ratings in the trained dataset and the fairly similar scores of the trained items is fairly similar to the one of the control group. Consequently, the test items that include the fricative /z/ need to be examined. The trained items are `close´ and `jeans´ and the untrained test items are `goes´ and `zoo´. The lower ratings might be due to the item `jeans´ that contains a nasal fricative cluster and it is prone to transfer from its German devoiced counterpart (/dʒiːnz/ vs. dʒiːns/). In addition, the untrained test item `goes´ is used very frequently among EFL students at the fifths grade and might therefore be used far more often than any other trained items.

Similar to the assumption that the explicit training of test items might result in higher pronunciation scores, this thesis also looks at word frequency that illustrates how often the vocabulary of a particular language is used (see Section 5.5). The more common a word, the more often it appears in the language used and thus higher frequency items should get higher pronunciation scores due to higher articulatory routines. The test items that were used to rate the students' pronunciation were thus not only set up according to trained and untrained status but also according to higher and lower frequency status (see Section 7.3.2). As both groups receive ongoing English lessons, frequency effects should show up in both the control and the intervention group.

The main effect of frequency (see 8.1.4) revealed that the higher frequency items got significantly higher ratings (M = 4.06, SE = .076) in comparison to the lower frequency items (M = 3.78, SE = .079). The mean ratings at the pre-test already show higher ratings for higher frequency items (M = 3.89, SE = .089) compared to lower frequency items (M = 3.58, SE = .106). This trend continues with increasing scores at the post-test and follow-up-test (see Table 8-1).

The simple effects analysis (see 8.1.9) of sound x frequency revealed that the test items containing the **vowels /ɔ:/** and **/æ/**, the **plosive /b/** in final position, the **fricative /ð/** and the **approximant /w/** received significantly higher ratings by higher frequency status. Moreover, the interaction graph (Figure 8-8) shows that the higher frequency items - with the exception of the **fricatives /θ/ and /z/** - generally received slightly higher ratings compared to the lower frequency items showing a strong effect of frequency.

However, the **fricative /z/** had even significantly lower ratings among the higher frequency dataset (see Figure 8-8). As the simple effects analysis entails data from all three data collection points, a look at the descriptive data might reveal some more information. The following scores are extracted from Table 8-10 and show the descriptive mean values (M) and mean differences (*M*) for the presented sounds by higher frequency (HF) and lower frequency (LF) status at the three data collection points (t_0, t_1 and t_2).

Intervention group:

/z/(HF): $M(t_0)$ = 3.25 (SE = .14) $M(t_1)$ = 3.13 (SE = .10, *M* = -.12) $M(t_2)$ = 3.06 (SE = .09, *M* = -.07)

/z/(LF): $M(t_0)$ = 3.56 (SE = .23) $M(t_1)$ = 3.88 (SE = .20, *M* = .32) $M(t_2)$ = 3.81 (SE = .22, *M* = -.07)

Control group:

/z/(HF): $M(t_0)$ = 2.94 (SE = .14) $M(t_1)$ = 2.94 (SE = .10, *M* = 0) $M(t_2)$ = 2.94 (SE = .09, *M* = 0)

/z/(LF): $M(t_0)$ = 3.19 (SE = .23) $M(t_1)$ = 3.50 (SE = .20, *M* = .31) $M(t_2)$ = 3.38 (SE = .22, *M* = -.12)

The mean frequency data for **/z/** show a fairly similar pattern to the mean training data with only a slight or no change in the ratings of the higher frequency items and an increase between pre- and post-test and a slight decrease between post-test and follow-up-test for both groups (see Table 8-10). Again, as discussed above with regards to the training effects, the presented scores might be due to a confound in the setup of the test items that contain the fricative /z/. Despite the ratings for **/z/** and **/θ/,** the presented results support the assumption that higher frequency items are easier to pronounce due to already mastered articulatory routines (see Section 5.5). To find out whether training adds to the frequency status, the following paragraph looks at the interaction between sound x frequency x training (see Section 8.1.9.1).

Figure 8-9 shows that there are significantly positive training effects among the higher frequency items that contain the **plosive /b/** in final position and the **diphthong /eɪ/**. This is hardly surprising as the **plosive /b/** got already significantly higher ratings in the interaction between sound x frequency (see Figure 8-8) and sound x training (see Figure 8-5). Although the ratings for the **diphthong /eɪ/** were not significantly different between higher and lower frequency items, the interaction between sound x training reveals a strong training effect for **/eɪ/** (see Figure 8-5).

However, the **fricative /ð/** and the **affricate /dʒ/** received significantly lower ratings for the trained in comparison to the untrained test items among the higher frequency dataset. This result is rather astonishing as both of the sounds show higher ratings for the trained test items in the interaction between sound x training (see Figure 8-8). It seems that the frequency effects overide the training effects. A look at the speech stimuli reveals that the **fricative /ð/** is included in the trained items `that´ and `them´ and in the untrained items `the´ and `smooth´. `That´ and `the´ are the higher frequency items whereas `them´ and `smooth´ belong to the lower frequency test items (see Table 7-4). Despite the fact that `the´ is not explicitly trained in the intervention, it is probably one of the most frequent words and thus used and trained every day in the English classroom. This might have led to the higher ratings for the untrained test item `the´ compared to `that´. The **affricate /dʒ/** is included in the trained items `job´ and `jeans´ and in the untrained items `John´ and `cage´. `Job´ and `John´ are the higher frequency items whereas `jeans´ and `cage´ belong to the lower frequency test items (see Table 7-4). The interaction shows that significantly higher ratings were given to the untrained item `John´ in comparison to trained stimulus `job´. Despite the fact that `job´ is explicitly trained in the intervention, it is prone to transfer from its German devoiced counterpart (/b/ vs /p/). Moreover, the stimulus `John´ is rather well-known to German

students and might even be high frequency in their vocabulary. The presented results indicate that the frequency effects surpass the training effects.

Among the interaction between lower frequency x sound x training there is a significant positive training effect for the **plosive /b/** in final position, the **fricatives /ð/** and **/θ/** and the **affricate /dʒ/**. The interaction graph also shows higher ratings for the trained stimuli that included the sounds **/ɔ:/**, **/ə/**, **/æ/**, **/eɪ/** and **/d/** (see Figure 8-10). These results show that the trained items that included the presented stimuli received – as expected - higher ratings in comparison to the untrained items; e.g. the **plosive /b/** is included in the trained items `job´ and `web´ and in the untrained items `club´ and `verb´. `Job´ and `club´ are the higher frequency items whereas `web´ and `verb´ belong to the lower frequency test items (see Table 7-4). The scores indicate that the trained item `web´ received higher ratings than the untrained stimulus `verb´.

Similar to the interaction of sound x frequency (see Figure 8-8) and sound x training (see Figure 8-5), the **fricative /z/** received higher ratings for the untrained items in the lower frequency dataset. The untrained lower frequency test item that includes /z/ is `zoo´ and the trained lower frequency item is `jeans´. It seems that `zoo´ was far easier to pronounce for the subjects than `jeans´. As already indicated above this might be due to rather complex utterance of `jeans´ that includes the nasal fricative cluster and is prone to final devoicing (for German learner of English). In addition, the /z/ in `zoo´ appears in word initial place and might be frequently used among the age group of the test subjects. The **diphthong /ɪə/** also received significantly lower ratings for the trained in comparison to the untrained test items among the lower frequency dataset. The untrained lower frequency test item that includes **/ɪə/** is `beer´ and the trained lower frequency item is `dear´. As both items are fairly similar and `dear´ is more common than `beer´, it is possible that the training could have had a negative effect as it might have led to an extra careful pronunciation of the diphthong (see Appendix, Section 10.4.2) for reference of the pronunciation training programme).

The presented results of the auditory analysis need to be seen in connection with the acoustic analysis that looks at the same data (see Section 8.4). The outcomes of both analyses will then be discussed in the context of this study in the conclusion (see Chapter 9).

8.3 Acoustic Analysis

As was pointed out in the previous chapters, two kinds of analyses were carried out to evaluate the pronunciation intervention study. The previous Sections presented the auditory analysis and discussed its results (see Section 8.1 and Section 8.2) and this section now follows on to look at the acoustic analysis. The same dataset was used for both analyses and as explained earlier it included 56 test items (see Table 7-4), which were each recorded three times before and after the intervention as well as five months later (pre-, post-, and follow-up test). In addition, the items were matched for frequency and training effects and they contained different sound categories (vowels, diphthongs, approximants, plosives, fricatives and affricates; see Section 7.3.3). As all of these six sound categories have inherent acoustic features (see Chapter 4), different acoustic analyses were run for each of the six categories and the findings are presented in the following sections (see Section 8.3.1 to Section 8.3.6). The acoustic data were analysed using Praat (Boersma & Weenik, 2012) and SPSS (IBM Corp., 2013). Where applicable, effects are reported as significant at $p < .05$. Finally, the last section 8.4 of this chapter will discuss the results of the acoustic analysis.

8.3.1 Acoustic analysis of the vowels

The perception of vowels is in particular determined by the variation of the first three formant frequencies across the duration of the vowel (Whiteside, Grobler, Windsor, & Varley, 2010). Additionally, durational features play an important role as they involve the temporal coordination of the tongue and lip articulators (Pickett, 2001) (see section 5.2). Therefore, vowel durations and formant frequencies were investigated for the acoustic analysis of the vowels (see Section 5.2.1).

The recorded test items were analysed in Praat. After a manual annotation, the vowel durations and formant frequencies for each of the three repetitions of the 56 test items at the pre-, post- and follow-up test were automatically measured in Praat, and the acoustic measures were transferred into SPSS (see Section 7.5.2). The following Table 8-12 presents the means and the standard errors for each of the three repetitions of the vowel durations for each data collection point by group (control/ intervention) status.

Table 8-12: Mean vowel durations and standard error for the control and intervention group by time

	Vowel	Control		Intervention	
		Mean vowel duration (ms)	SE	Mean vowel duration (ms)	SE
T_0 pre-test	**/ɑː/**	281.333	23.716	310.104	18.749
	/ɔː/	276.333	25.202	271.510	19.924
	/ə/	166.333	11.597	181.250	9.168
	/æ/	197.667	20.533	204.427	16.233
T_1 post-test	**/ɑː/**	278.500	38.134	339.375	30.148
	/ɔː/	270.167	27.858	326.354	22.023
	/ə/	155.500	14.194	192.500	11.222
	/æ/	203.583	25.103	232.604	19.846
T_2 follow-up test	**/ɑː/**	313.417	25.928	324.063	20.498
	/ɔː/	289.333	23.569	288.021	18.633
	/ə/	164.500	11.742	175.833	9.283
	/æ/	217.750	18.949	215.260	14.980

The mean values show that for both groups the longer vowels /ɑː/ and /ɔː/ retain longer duration in contrast to the shorter vowels /ə/ and /æ/. This indicates a robust effect of long versus short vowels. The duration of all vowels in the intervention group increased between the pre-, and post-test and decreased between the post- and the follow-up-test. However, this effect is not seen in the control group. In contrast, the vowel duration only shows slight differences between the pre- and post-test in the control group and there is an increase in duration between the post- and the follow-up-test. For all of the four vowels the intervention group displays longer durations after the post-test.

A mixed-model ANOVA was run on the vowel durations to determine whether there were any main effects and interactions for time, frequency, training and the four vowels (/ɑː/, /ɔː/, /ə/, /æ/) between the intervention and the control group (see Table 8-13).

Table 8-13: Significant within subjects effects for the vowel data

Effects	df[51]	Error	Mean Square	F	Sig.	Partial Eta²
training (trained/ not trained)	1.000	984.312	6877.150	6.987	.023	.388
frequency (higher/ lower frequency)	1.000	11.000	33922.797	15.121	.003	.579
vowel	2.653	29.183	669968.450	88.838	.000	.890
time x frequency x group	1.903	20.937	3220.928	4.019	.035	.268
training x frequency	1.000	11.000	41984.167	23.422	.001	.680
training x vowel	2.649	29.144	87201.518	45.845	.000	.806
frequency x vowel	2.185	24.030	6819.243	3.833	.033	.258
training x frequency x vowel	2.215	24.369	52482.808	16.248	.000	.596

The ANOVA revealed (see Table 8-13) that there were statistically significant main effects for training, frequency and the vowels. Moreover, there were interaction effects for time x frequency x group, training x frequency, training x vowel, frequency x vowel and training x frequency x vowel. The group effect did not show significant results.

[51] The degrees of freedom and the error are fractioned due to the fact that they are calculated with the Greenhouse-Geisser correction.

However, the output of the ANOVA has to be interpreted with caution due to the setup of the test items. As already indicated in the discussion of the auditory analysis (see Section 8.2), and shown in the set-up of the test items (see section 7.3.3), there are biases underlying the matching of the test items as the vowels appear in different phonetic contexts. In addition, the test items also include monosyllable as well as disyllabic words (`from´ vs. `mother´) and there are also nasalised and non-nasalised vowels (`thank´ vs `bag)´. These phonetic differences in coarticulation, stress and nasalisation strongly affect vowel duration (Pickett, 2001). Though, in order to get rid of the depicted differences and to achieve the highest possible comparability, nonsense words would have been needed. Yet, to achieve a high economic validity of the study, real-life data was used and therefore it was impossible to match the phonetic sequences equally on complexity.

Moreover, the perception of vowel quality is largely determined by the first two formant frequencies of the vowels (Harrington, 2013). It is for these reasons, that the outcomes of the ANOVA are not evaluated any further, but that the vowels are analysed according to their frequency features. As the formant frequencies are also influenced by coarticulation through the identity of the preceding and following segments, the vowel target is typically measured near the temporal midpoint of the vowel (Harrington, 2013). This approach is adopted in this thesis (see section 5.2) and Table 8-14 shows the mean and standard deviation values of the F1 and F2 frequencies at the temporal midpoint of the targeted vowels.

Table 8-14: Mean and standard deviation values of F2 and F1mid vowel formant frequencies (in Hz) by group

vowel	time	intervention						control					
		F1mid			F2mid			F1mid			F2mid		
		N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
/ɑː/	pre-test	32	905	123	32	1434	193	31[52]	840	105	31	1353	219
	post-test	32	885	115	32	1358	157	32	836	163	32	1375	192
	follow-up test	32	907	87	32	1375	139	32	780	117	32	1328	171
/ɔː/	pre-test	32	777	141	32	1266	253	31	736	98	31	1274	320
	post-test	32	748	99	32	1183	197	30	713	125	30	1228	264
	follow-up test	32	733	97	31	1189	188	31	675	114	31	1201	236
/ə /	pre-test	32	730	134	32	1447	216	32	736	100	32	1459	230
	post-test	32	725	116	32	1425	206	32	684	109	32	1484	229
	follow-up test	32	702	107	32	1404	203	32	657	90	32	1412	209
/æ/	pre-test	32	846	84	32	1848	304	32	781	104	32	1749	365
	post-test	32	830	66	32	1940	293	32	754	102	32	1931	337
	follow-up test	32	842	102	32	1979	235	32	733	117	32	1883	352

[52] The total number of correct vowel productions is n = 32. In case a subject did not produce a vowel at all, this production was excluded from the study and the mean of the remaining productions was used.

As discussed in Section 5.2.1, vowel spaces represent a two-dimensional acoustic phonetic map where “F1, which is an index of degree of openness, is plotted against the difference between F2 and F1, which serves as a general index of anterior/ posterior constriction” (Whiteside et al., 2010). Figure 8-11 presents the vowel spaces using the vowel midpoint values of the first two formant frequencies for each of the four treated vowels /ɑː/, /ɔː/, /ə/ and /æ/ by time and group status. All data < 2 SD is displayed in the ellipses. The formant patterns can be used to show the spatial articulatory precision with which the language learners in the control and intervention group produce the targeted English vowels.

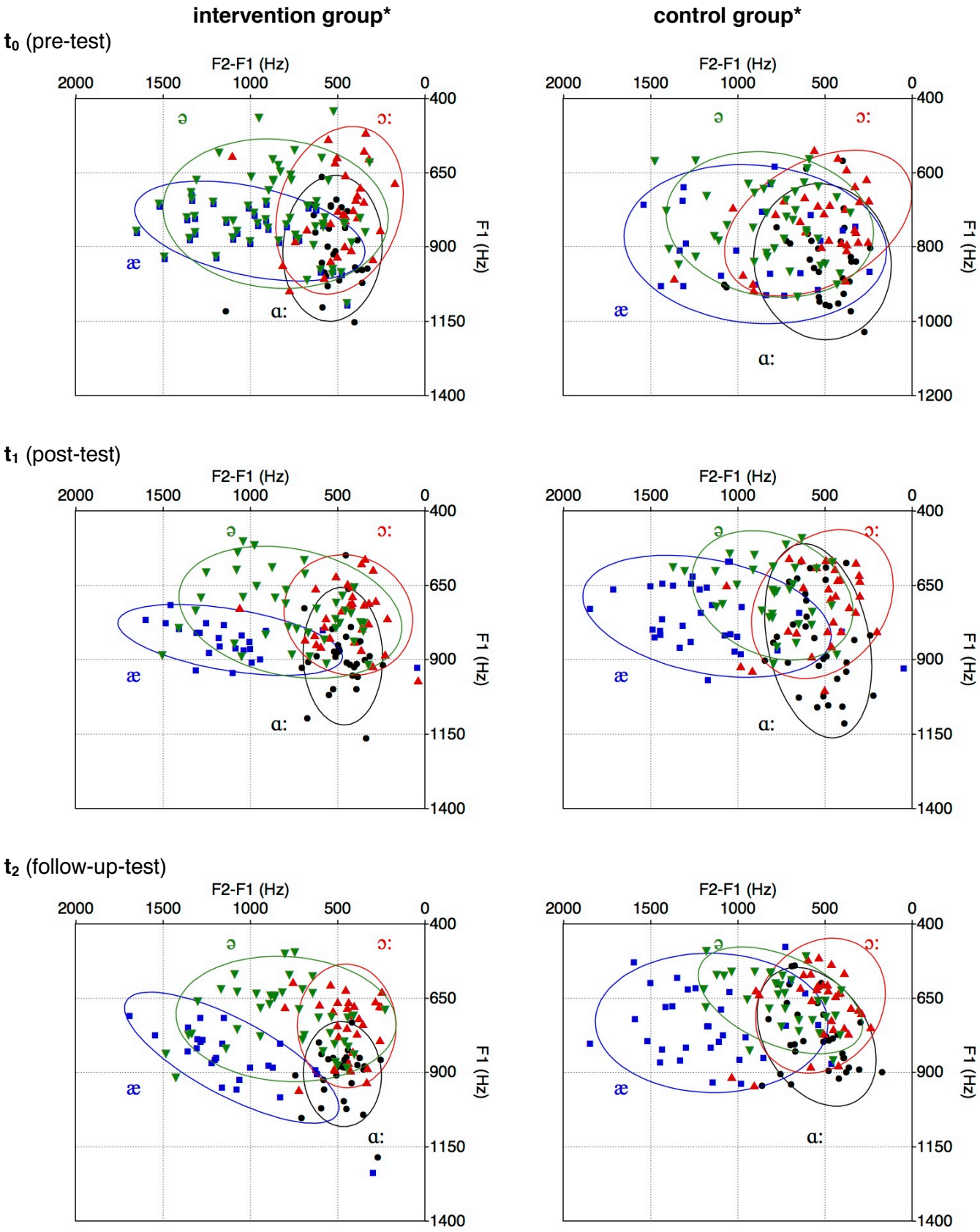


Figure 8-11: Vowel spaces by time (t_0, t_1, t_2) for the intervention and control group

Vowel: ●/ɑ:/ ▲/ɔ:/ ▼/ə/ ■/æ/, *ellipses include all data < 2SD

The vowel spaces (see Figure 8-11) show apparent differences by group and between the three data collection points.

The vowel spread clearly displays a more variable and disrupted production of the vowels for the intervention group in the pre-test, in comparison to the smaller ellipses of the post-test and the follow-up-tests. This is also supported by the decreasing standard deviation values presented in Table 8-14 for /ɑː/, /ɔː/ and /ə/ between the three test periods showing increasing stability between the three test points. However, the standard deviation for the first formant in the vowel /æ/ increases between the post- and follow-up-test. This shift is also mirrored in the upwards movement of the /æ/- ellipsis showing a more closed pronunciation of the vowel.

The inspection of the vowel ellipses for the control group indicates an anterior and posterior constriction between the pre- and the post-test for the vowels /ɑː/, /ɔː/ and /ə/. This is supported by the decreasing standard deviations for F2 (see Table 8-14) for these vowels. In contrast, the standard deviations for F1 increase between the pre- and the post-test. This is also mirrored in the more upwards constriction of the vowel ellipses of the post-test and thus a closer production of the vowels /ɑː/, /ɔː/ and /ə/. Between the post- and the follow-up-test, the standard deviation values for F1 and F2-F1 decrease, showing smaller ellipses and thus more stability in the production of /ɑː/, /ɔː/ and /ə/ at the follow-up-test. Although, the front open vowel /æ/ clearly shows less variation at the post-test in comparison to the pre-test which is supported by a smaller standard deviation for F2 (see Table 8-14), the scatter indicates a more fronted production of the vowel at the post-test. At the follow-up-test, the vowel ellipses for /æ/ shows an increased spread of the data points in the open-closed and anterior/ posterior direction which is also supported by increased standard deviation values for F1 and F2. This increasing spread and standard deviations indicate that the subjects are unsure of how to pronounce /æ/.

Comparing the ellipses of both groups, the vowel productions of the control group show more variability at the beginning of the intervention. This might indicate a bias in the matching of both groups. Although the scatter becomes less dispersed for both groups after the pre-test, the control group still shows more variation in the production of the target vowels. With the exception of the front-open vowel /æ/, the trend of the vowel spaces getting smaller, indicated by the shrinking ellipses, continues between the post- and the follow-up test for both groups indicating more stability in the

pronunciation of the vowels. Whereas the dispersion of the vowel production of /æ/ increases in the follow-up test for the control group, the intervention group shows a more closed production of /æ/ portrayed by the upward shift of the ellipsis.

In order to get more insights into the shift in vowel quality addressed above, the averaged vowel spaces and Euclidean distances need to be inspected. Figure 8-12 presents the averaged vowel spaces for both groups by time using the mean midpoint values for the vowels /ɑː/, /ɔː/, /ə/, and /æ/ and it also displays the Euclidean centroid of the cluster (see Section 7.5.2.2.1). Table 8-15 presents the Euclidean distances calculated between the centroid and the vowel midpoint values for F1 and F2-F1 for each of the four vowels by time. These distances can be used to assess the sizes and shapes of the vowel spaces. It is apparent from Figure 8-12 that there are clear differences in the averaged vowel spaces between the control and intervention group and between the pre-, post- and follow-up-test.

The vowel /ɑː/ is typically pronounced with an open back quality (see Figure 3-3). After the intervention, the vowel produced by the intervention group shifts slightly backwards and becomes less open. Although it keeps its back quality at the follow-up-test, it is again pronounced with a more open quality, which is rather on target with references to the English vowel inventory. Similar to the intervention group in the pre-test, the control group pronounces the vowel /ɑː/ with a comparable backwards constriction but with a less open quality. Throughout the three test periods there is only a slight shift towards a more anterior pronunciation of the /ɑː/ but it becomes more constricted and loses its open quality being pronounced rather neutrally.

Table 8-15: Mean values of the Euclidean distances (Hz) between the target vowel and the centroid (see Figure 8-12) in the F1 and F2-F1 vowel space

Vowel	Test time	Intervention	Control
/ɑː/	pre-test	178	185
	post-test	224	273
	follow-up-test	256	209
/ɔː/	pre-test	197	152
	post-test	250	245
	follow-up-test	230	222
/ə/	pre-test	89	53
	post-test	75	76
	follow-up-test	95	55
/æ/	pre-test	321	282
	post-test	432	419
	follow-up-test	440	440

Table 8-16: Euclidean distances (Hz) between /ɑː/ - /ɔː/ and /æ/ - /ɑ/ in the F1 and F2-F1 vowel space by group and time

	Intervention		Control	
Euclidean Distance	/ɑː/ - /ɔː/	/æ/ - /ɑː/	/ɑː/ - /ɔː/	/æ/ - /ɑː/
pre-test	134	477	107	458
post-test	142	639	125	644
follow-up-test	174	672	107	604

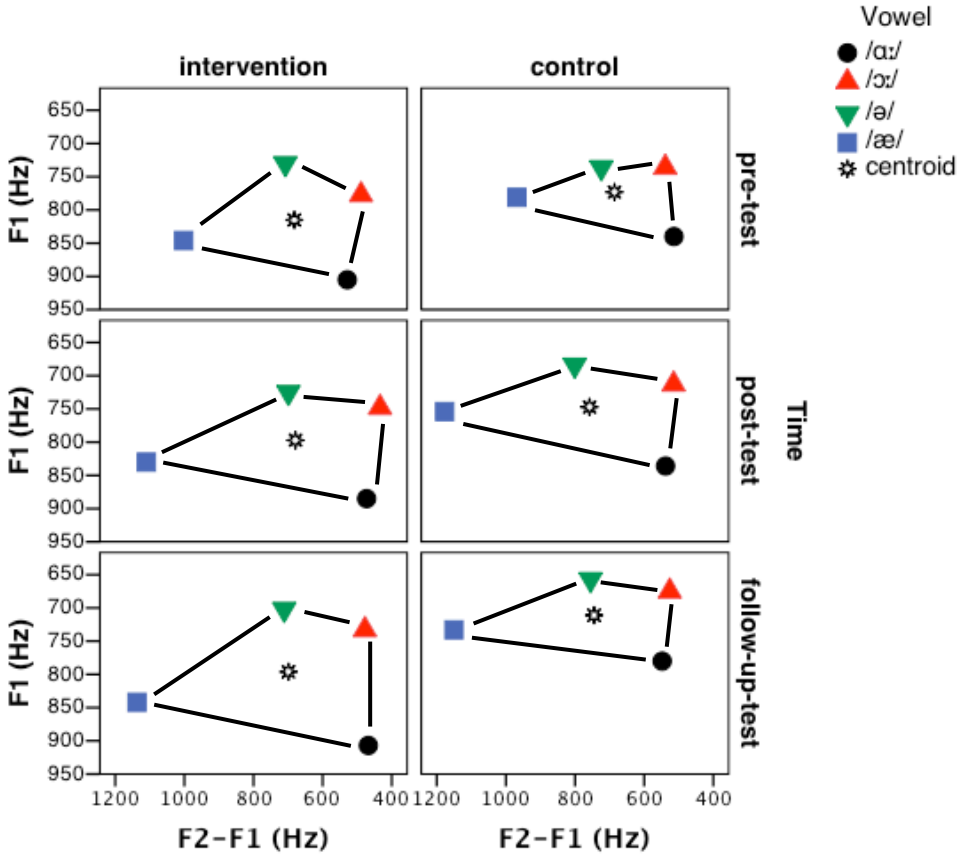


Figure 8-12: Averaged vowel spaces by time for intervention and control group

The posterior constriction increases for the back vowel /ɔː/ in the intervention group between the pre- and post-test and then slightly relaxes from post- to follow-up-

test. Additionally, it becomes slightly more closed with time. Similar patterns of constriction also take place for the control group, despite the fact that the /ɔː/ was already pronounced with a more closed quality at the pre-test leading to a more closed pronunciation of the /ɔː/ at the follow-up test in comparison to the intervention group. It seems that both groups initially pronounce the German /ɔ/ which has a more open quality than the English /ɔː/ (see Figure 3.3.), and then approximate the English target sound.

The neutral vowel /ə/ does not show much variation in the intervention group with regards to its anterior/ posterior constriction as it only becomes slightly more closed between the post- and the follow-up-test. Although having a slightly more open quality at the pre-test, the production of /ə/ of the control group shows a stronger change and becomes more open and a bit more fronted with time.

The open and the front quality of the open - front vowel /æ/ increased with time in the intervention group. At the pre-test the /æ/ of the control group displays a rather neutral quality. Between the pre- and the post-test there is a big shift towards a more fronted pronunciation of the vowel /æ/ and then a slightly more neutralised quality for the follow-up-test. With regard to the open-closed dimension, the /æ/ becomes more closed with the test points.

Summarising the changes four the four targeted vowels, it becomes apparent from the inspection of the averaged vowel spaces, that the vowel spaces of the intervention and control groups become bigger between the pre- and the post-test. This is also supported by the increase in Euclidean distances (see Table 8-15). Aside from the vowel /æ/ the vowel spaces for /ɑː/, /ə/ and /æ/ further increase in the intervention group at the follow-up test. In contrast, for the control group the Euclidean distances and also the vowel spaces decrease for all vowels but for /æ/ after the post-test.

A closer look at the front-open vowel /æ/ and the back-open vowel /ɔː/ in the vowel spaces of the intervention group reveals that the front vowel shifts even more to the front, and the back vowel shift also even backwards with time. This is supported by the increasing Euclidean distances (see Table 8-16) between /æ/ and /ɔː/ (477 Hz, 639Hz, 672 Hz). Moreover, the increase in the distance between open-back vowel /ɑː/ and back vowel /ɔː/ (134Hz, 142Hz, 174 Hz) shows the spread of the vowel space for the intervention group.

An inspection of the vowel spaces of the control group indicates that vowels shift apart in the anterior and posterior direction, especially after the post-test. However, with regard to the open closed dimension, all vowels seem to be pronounced with a more closed quality after the follow-up-test. This is also supported by the upwards shifting centroid and a smaller Euclidean distance between the open-back vowel /ɑː/ and neutral –back vowel /ɔː/ between the post- and follow-up test (644 Hz, 604 Hz).

8.3.2 Acoustic analysis of the diphthongs /ɪə/ and /eɪ/

Diphthongs are characterised by the time-varying shift of frequencies from one vowel to another (Jacewicz, 2009). To assess the position and dynamic movements of the diphthongs in the F1 and F2-F1 plane, the acoustic measurements included diphthong duration and formant values of F1 and F2-F1 taken at the 20%, 40%, 60%, and 80% temporal points of the vowel. The first and last 20% portions of the vowels were eliminated to reduce the immediate coarticulation effects of surrounding consonants. In addition, these measurements were used to calculate vector length (VL), trajectory lengths (TL) and the spectral **r**ate **o**f **c**hange for the TL (TL_**roc**) (see Section 7.5.2.2).

The first section of this chapter examines diphthong duration, then it goes on to have a closer look at the formant movements, and VL, TL and TL_roc.

8.3.2.1 Diphthong duration

Table 8-17 presents the mean and standard error values for the diphthong durations of /ɪə/ and /eɪ/ for the control and intervention groups by time. The results (see Table 8-17) show a systematic difference in duration for the two different diphthongs and between the intervention and control group.

Table 8-17: Mean diphthong durations and standard error (both in ms) for the control and intervention group by time

	Diphthong	Intervention		Control	
		Mean diphthong duration (ms)	SE	Mean diphthong duration (ms)	SE
T_0 pre-test	**/ɪə/**	395	23	332	20
	/eɪ/	301	11	252	10
T_1 post-test	**/ɪə/**	376	20	320	17
	/eɪ/	320	21	252	17
T_2 follow-up-test	**/ɪə/**	374	29	351	24
	/eɪ/	308	21	256	17

The diphthong duration of /ɪə/ produced by the intervention group decreases from pre- to post-test (395 ms to 376 ms), and there is a slight decrease in the follow-up test (374 ms). Similar to the intervention group, the duration decreases in the control group between the pre- and post-test (332 ms to 320 ms). However, the follow-up-test reveals an increased diphthong duration (351 ms) for /ɪə/. Comparing both groups, the control group overall shows shorter diphthong durations.

The mean diphthong durations for /eɪ/ show a similar picture, with the mean durations being higher for the intervention (309.67ms) than the control group (253.34 ms). The diphthong durations of the production of /eɪ/ of the intervention group increase between pre- and post-test (301ms to 320 ms) and then decrease again in the follow-up test 308ms. This rise in duration might indicate that the learners changed their way of pronouncing the diphthong between the pre-and post-test and then produced a more stable version of /ɪə/ in the follow up test. However the durations of the diphthong /eɪ/ of the control group show a different picture. There is no difference in the durations produced in the pre- and post-test (252 ms and 252 ms) and also only a slight increase in the duration of the follow-up-test (256 ms). These results indicate that the control group does not considerably change in the pronunciation of /eɪ/.

The standard errors for both groups behave similarly. There is an increase between the pre- and post-test in the intervention (SE 11, SE 21) and control group (SE 10, SE 17) and then the values do not alter between the post- and follow-up test in both groups. This rise of the SE might also indicate an alteration in how the learner produces the diphthong between the pre- and post-test. However, the rather high SE of the follow-up test points to ongoing pronunciation difficulties.

Comparing the duration results of both diphthongs, it seems that there are systematic differences between /ɪə/ and /eɪ/ which are analysed in more detail with the help of an ANOVA. In addition, Section 8.3.2.2 has a closer look at the diphthong quality.

A mixed model analysis of variance with the within subject factors time, frequency, training and diphthong, and the between subject factor group (intervention vs. control) was carried out to determine the differences in diphthong duration. The ANOVA revealed (see Table 8-18) that the between subjects effect group and all three within subject effects diphthong, training and frequency were significant with strong effects sizes (see Table 8-18). However, there was no significant main effect of time.

Table 8-18: Significant between and within subjects effects for the diphthong data

Effects[53]	df	Error	Mean Square	F	Sig.	Partial Eta Squared
group (intervention/ control)	1	10	186904.169	5.741	.038	.989
training (trained/ not trained)	1	10	158267.533	59.158	.000	.855
frequency (higher/ lower frequency)	1	10	121469.914	249.852	.000	.962
diphthong (/ɪə/, /eɪ/)	1	10	406751.422	76.722	.000	.855
training x frequency	1	10	107850.278	149.158	.000	.937
training x frequency x group	1	10	6277.979	8.682	.015	.465
training x diphthong	1	10	69396.264	113.181	.000	.919
training x diphthong x group	1	10	2820.106	4.599	.058[54]	.315
frequency x diphthong	1	10	75525.801	36.007	.000	.783
training x frequency x diphthong	1	10	222438.525	126.529	.000	.927

The impact of the intervention is shown in the significant main effect of group ($F(1,10) = 5.741$, $p < 0.05$, partial $\eta^2 = .989$). Post-hoc tests with a Bonferroni adjustment revealed that the intervention group (M = 345.750 ms, SE = 16.471) produced significantly longer durations than the control group (M = 294.077 ms, SE = 13.921), ($M = 51.673$ ms, SE = 21.566, $p < 0.05$).

The significant main effect of diphthong ($F(1,10) = 76.722$, $p < 0.001$, partial $\eta^2 = .855$) showed the fundamental differences in the durations of the two diphthongs with significantly longer durations for /ɪə/ (M = 358.028 ms, SE = 13.991) than for /eɪ/ (M = 281.800 ms, SE = 8.641), ($M = 76.228$ ms, SE = 8.703, $p < 0.001$).

The strong impact of the factors training and frequency (see chapter 5.5) is shown in the significant main effects of training ($F(1,10) = 59.158$, $p < 0.001$, partial $\eta^2 = .855$) with longer duration for the trained (M = 343.688 ms, SE = 12.307) compared to the untrained items (M = 296.139 ms, SE = 10.010), ($M = 47.550$, SE = 6.182, $p < .001$), and frequency ($F(1,10) = 249.852$, $p < 0.001$, partial $\eta^2 = .962$) with shorter durations for the higher frequency items (M = 299.085 ms, SE = 10.863) in comparison to the lower frequency items (M = 340.742 ms, SE = 10.846), ($M = -41.657$ ms, SE 2.635, $p < 0.001$).

[53] Effects are adjusted with the Greenhouse-Geisser correction.
[54] approaches significance

Similar to the ANOVA carried out for the vowels (see 8.3.1), the presented results have to be interpreted with great caution due to the inherent bias in the nature of the test items and the different coarticulatory contexts of the diphthongs (see table 7.3.3). It is for this reason, that the significant interactions (see Table 8-18) are not discussed in further detail.

8.3.2.2 Formant movement

With regard to the formant analysis, Figure 8-13 and Figure 8-14 display the relative positions and formant movement of the scatter and means at four equidistant points in the central 60% of the diphthong /ɪə/ in the F1 x F2-F1 plane by time and group. Similarly, Figure 8-15 and Figure 8-16 show the relative positions and formant movement of the scatter and means of the diphthong /eɪ/ in the F1 x F2-F1 plane by time and group. The colours and arrows indicate the direction of the formant movements.

The visual inspection of the scatter plots of /ɪə/ (Figure 8-13) shows that the intervention and control group both show substantial variability in the production of the diphthong at the pre-test compared to the post-test. However, the cluster of the control group shows more variation in comparison to the intervention group with regard to the 60% and 80% portion of the vowel. The scatter for the intervention group becomes substantially less dispersed at the post- test and also decreases again at the follow-up-test showing an increasing stability in the productions of the diphthong /ɪə/. Although the cluster of the control group becomes slightly denser at the post- and follow-up test, the production of /ɪə/ shows more variation in comparison to the intervention group. This variation is especially evident in the production of the 60% and 80% point of the diphthong and might be due to fact that some learners might confuse the English /ɪə/ with the more open diphthong /ɪɐ/. The decreasing spread of the control and intervention group indicates a more stable production of the diphthongs.

Figure 8-14 shows substantial differences in the mean relative formant positions of /ɪə/ between the intervention and control group at the pre-test. Compared to the intervention group, the formant positions of the control group are less fronted at the opening phase of the diphthong and the trajectory ends in a more open position.

Additionally, the F1 values of the control group decrease between the pre- and post-test. This results in a less open production of /ɪə/ at this stage. Additionally, /ɪə/ is

produced slightly more fronted at the post-test. At the follow-up test the trajectory length decreases and while the F1 values of the 20% and 40% portions of the vowel increase, the F1 values of 60% and 80% decrease. This results in a smaller glide between /ɪ/ and /ə/. As a result a rather neutral and monophtongal version of /ɪə/ is produced. The mean formant positions and trajectory length of the production of /ɪə/ of the intervention group remain rather stable throughout the three data collection points.

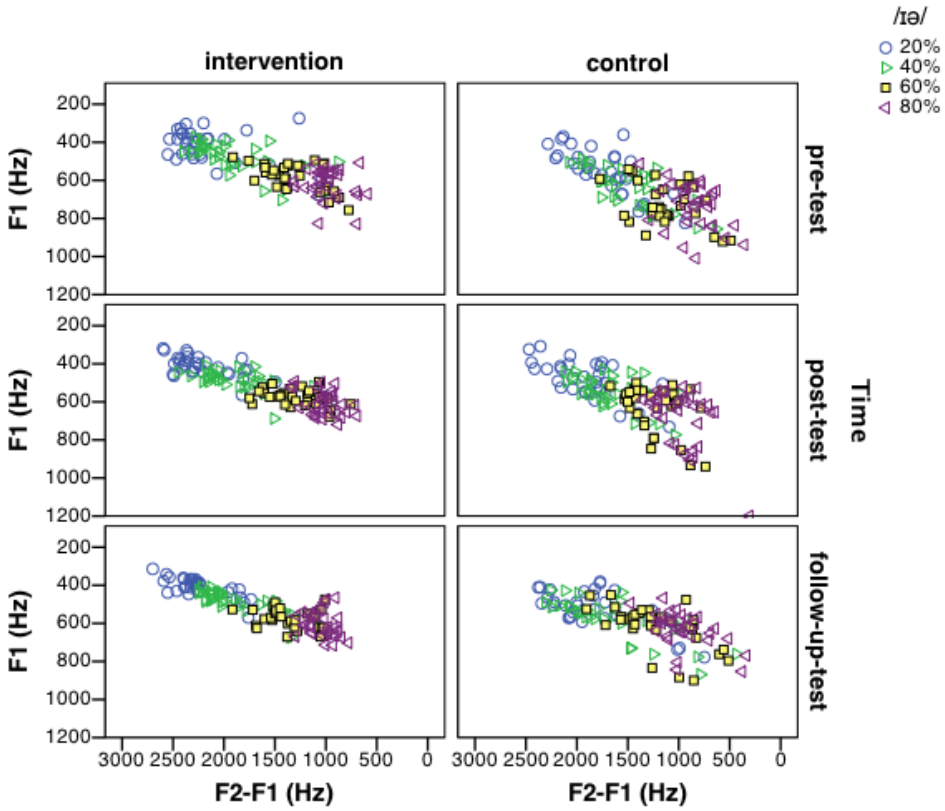


Figure 8-13: Formant movement (in Hz) of the 20-80% of /ɪə/ by group and time

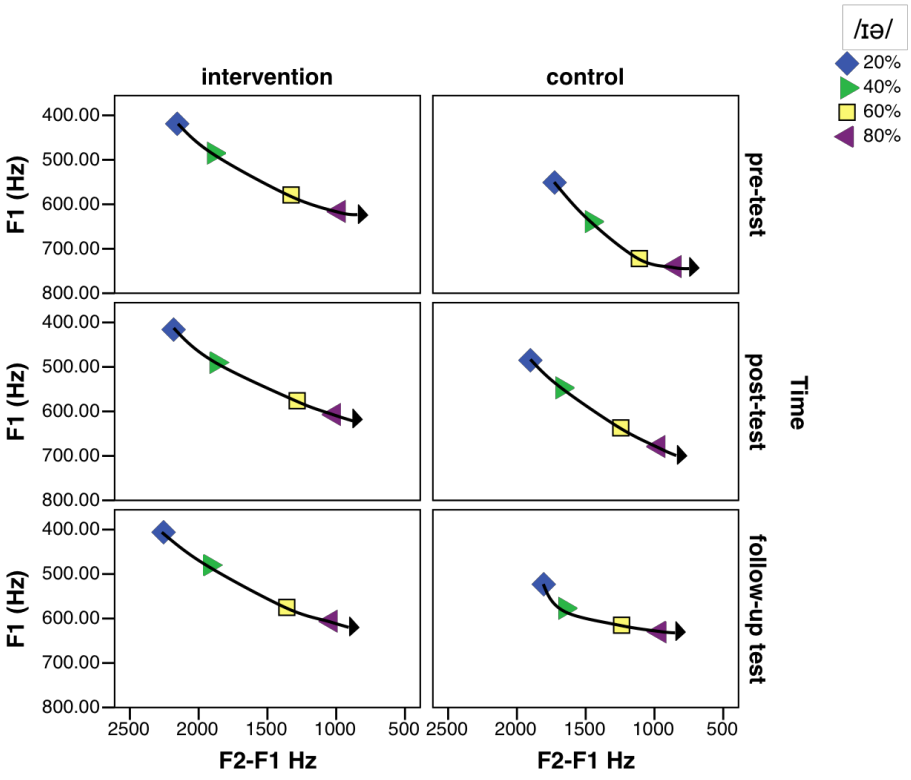


Figure 8-14: Mean formant movement (in Hz) of the central 20-80% of /ɪə/ by group and time

Figure 8-15 displays the relative positions and formant movement at four equidistant points (20%, 40%, 60%, 80%) of the central 60% of the diphthong /eɪ/ in the F1 x F2-F1 plane by time and group. Similar to the plots of /ɪə/ in Figure 8-13, the visual inspection of the scatter plots of /eɪ/ in Figure 8-15 shows that the intervention and control group both show substantial spread in the production of the diphthong at the pre-test phase of the study. Again, the control group shows more variation in comparison to the intervention group. This dispersion is especially evident in the more backed production of the 20% and 40% portions. Although the cluster of the control group shows overall less scatter at the post-test phase, the variation increases again for the follow-up-test. In contrast, the scatter for the intervention group becomes less dispersed at the post-test and the follow-up-test showing more stability in the production of the diphthong /eɪ/.

The visual inspection of the mean formant movement in Figure 8-16 shows that the diphthong /eɪ/ produced by the intervention group becomes slightly more open and fronted at the post-test phase compared to the pre-test phase. In addition, the glide between the two vowels changes and there is more formant movement. However, this trend declines at the follow-up test showing a slightly more open but otherwise very similar production of the diphthong compared to the pre-test. In comparison to the intervention group, the control group produces a less fronted and more open version of the diphthong /eɪ/ compared to the control group at the pre-test. The diphthong characteristics of the control group at the post-test also becomes more fronted and less open and the vowel space increases showing a production pattern similar to the pre-test of the intervention group. Yet, the vowel space of the control group shows increased spread at the follow-up-test, a smaller trajectory length and the 60% and 80% portions of the diphthong display decreased F2-F1 values leading to a more neutral production of /eɪ/ in comparison to the post-test and the follow-up test of the intervention group.

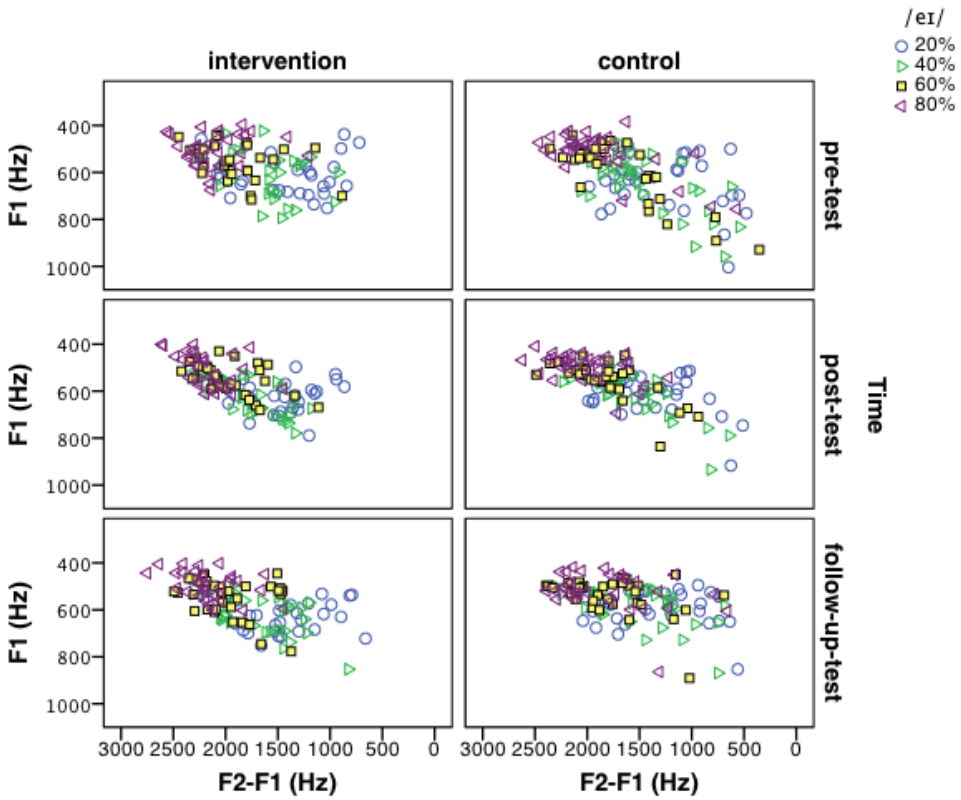


Figure 8-15: Formant movement (in Hz) of in the central 20%-80% of /eɪ/ by group and time

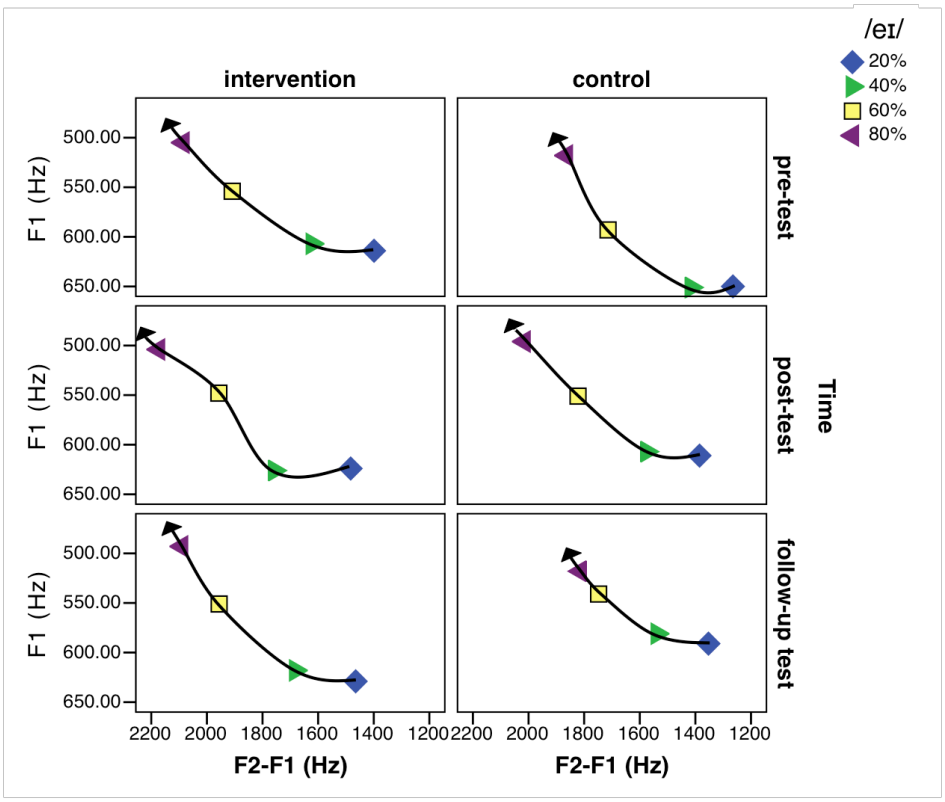


Figure 8-16: Mean formant movement (in Hz) of the mean 20%-80% of /eɪ/ by group and time

In Figure 8-13 to Figure 8-16 the formants are plotted at four equidistant points during the diphthong. Therefore, the frequency measurements are time normalised and do not take the duration into account. This issue will be addressed in Section 8.3.2.4

8.3.2.3 Vector length (VL) and trajectory length (TL)

The vector length (VL) in the F1 x F2 plane indicates the amount of formant change over the diphthong's duration. However, VL only calculates the Euclidean distance between the 20% and 80% portion of the diphthong and it fails to account for the non-linear nature of the formant tracks. In contrast, the trajectory length (TL) divides the diphthong in several sections and adds up the Euclidean distances for each portion of the diphthong (Fox & Jacewicz, 2009) and thus provides a more detailed account of the formant change across the diphthong (see Sections 7.5.2.2.2 and 7.5.2.2.3). On this basis, and also because the results of the VL and TL analyses showed comparable trends (see Figure 8-17), the following analysis concentrates on the results of the trajectory lengths (see Figure 8-18).

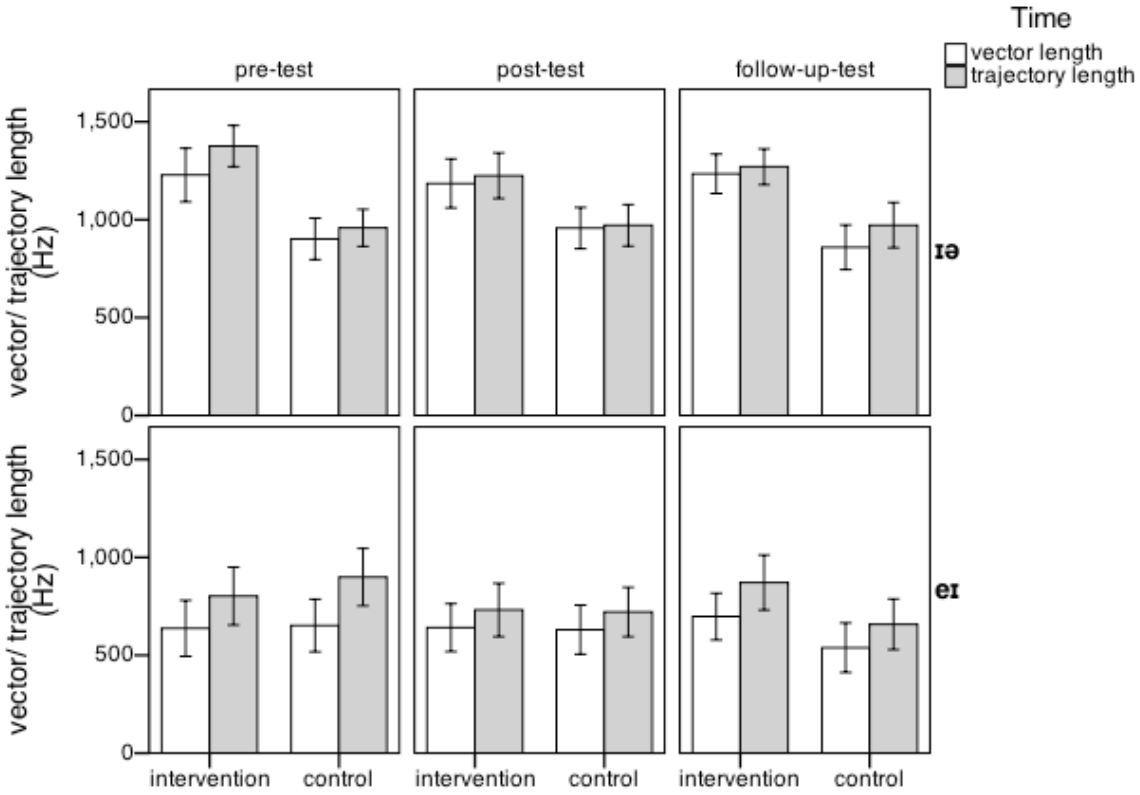


Figure 8-17: Mean VL and TL values and SEM[55] for the diphthongs /ɪə/ and /eɪ/ per group

[55] The error bars in the bar chart represent the standard error of the mean (SEM)

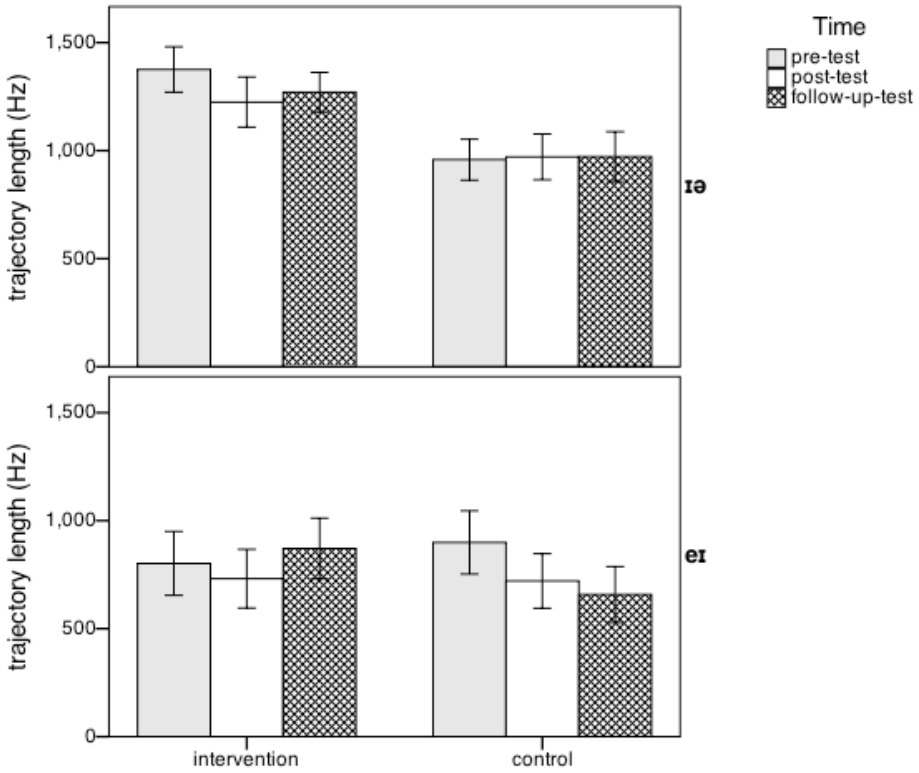


Figure 8-18: Mean TL values and SEM for the diphthongs /ɪə/ and /eɪ/ per group

As might be expected, Figure 8-18 reveals that there are systematic differences between the trajectory lengths of the two diphthongs /ɪə/ and /eɪ/. With regard to /ɪə/ produced by the intervention group the trajectories decrease between the pre- and the post-test and there is again a slight increase at the follow-up test. This decrease indicates less formant change. The control group only shows a small increase in TL over the three data collection points, which does not show much change in the production of the diphthong. In general, the intervention group already shows longer trajectory lengths at the pre-test indicating a bias between the two groups.

The trajectory lengths for /eɪ/ show a different picture. Although the intervention group also displays a small decrease in TL between the pre- and the post-test, the trajectory length surpasses the pre-test in the follow-up test. In the case of the control group, TL shows a steady decrease from pre-test to post-test and follow-up test. Overall, the intervention group displays smaller TL values in the pre-test in comparison to the control group.

Comparing the TL's for both diphthongs it becomes clear that on the whole TL's are greater for /ɪə/.

The confidence intervals in Figure 8-17 indicate significant effects for /ɪə/. Therefore, a separate repeated-measures ANOVA was carried out for trajectory length of the diphthong. Group, frequency and training were included as between subject factors.

Table 8-19: Significant between subjects effects for TL

Effects	df	Error	Mean Square	F	Sig.	Partial Eta Squared
group (intervention/ control)	1	56	1669694.69	37.464	.000	.401
training (trained/untrained)	1	56	278368.340	6.241	.015	.100
group x training	1	56	437199.188	3.270	.076[56]	.055

The results (see Table 8-19) show significant main effect of group ($F(1,56) = 37.464$, $p < 0.01$, partial $\eta^2 = .401$). Post-hoc tests with a Bonferroni adjustment revealed that the TL produced by the intervention group (M = 1290.427 Hz, SE = 37.319) were significantly longer than those of the control group (M = 967.385 Hz, SE = 37.319), (*M* = 323.042 Hz, SE = 52.778, $p < 0.00$).

The impact of the training is shown in the significant main effect of training $F(1,56) = 6.241$, $p < 0.05$, partial $\eta^2 = .100$. Post-hoc tests with a Bonferroni adjustment revealed significantly longer length for the trained items (M = 1194.833 Hz, SE = 37.319) in comparison to the untrained items (M = 1062.979 Hz, SE = 37.319), (*M* = 131.854 Hz, SE = 52.778, $p < 0.05$).

8.3.2.4 Spectral roc

The spectral **r**ate **o**f **c**hange (**roc**) gives evidence of how quickly the formant frequency changes occur in time (Fox & Jacewicz, 2009) as it divides the trajectory length of each of the three portions of the diphthong (section 1: 20-40%, section 2, 40-60%, section 3: 60-80%) by its duration. Figure 8-19 shows the mean spectral **roc** for each section by time, group and diphthong.

[56] approaches significance

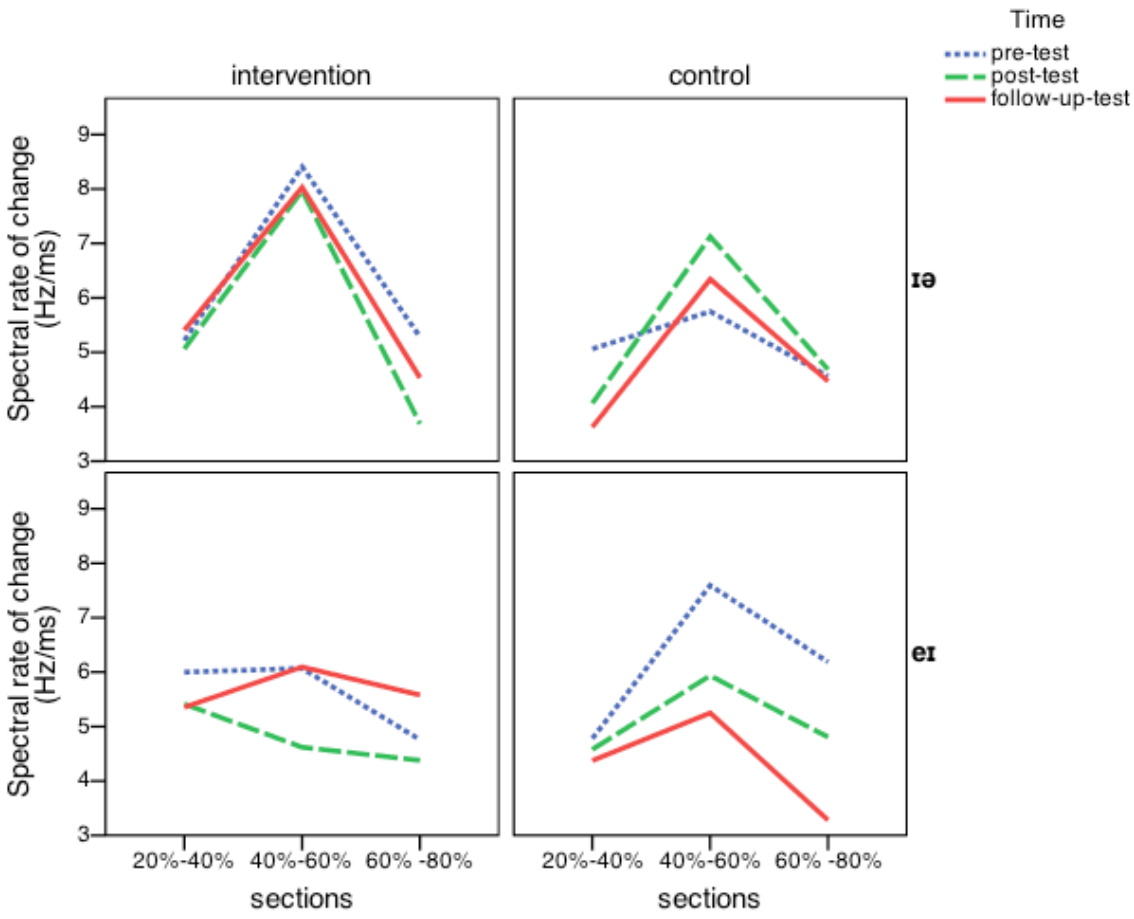


Figure 8-19: Spectral rate of change for each section by time, group and diphthong

The visual inspection of Figure 8-19 indicates considerable differences in the spectral change between the two diphthongs.

With regards to /ɪə/, the greatest spectral change is displayed for both groups at the 40% to 60% portion at each of the three data collection points indicating a typical diphthong pattern with fast articulatory movements between the two vowels at this stage of the diphthong. In comparison to the post- and follow-up test, the control group produces the 20% to 40% portion of the diphthong at the pre-test with a rather high articulation rate. Looking at both groups, the intervention group displays an overall higher mean spectral roc at the 40% to 60% portion of the diphthong indicating faster articulatory movements to reach the vowel target (Fox & Jacewicz, 2009).

Although the spectral rate of change of /eɪ/ produced by the control group shows a typical diphthong pattern, the intervention group clearly displays a different picture. The spectral rate of change of the intervention group does not change a lot between the first sections of the diphthong at the pre-test but decreases at the 60% to 80%

portion of /eɪ/. The post-test phase reveals decreasing spectral rates of change. In contrast, the follow-up-test again displays a rather typical graph for a diphthong with a higher articulation rate at the middle portion of the diphthong.

The spectral roc analysis of /eɪ/ (see Figure 8-18) reveals a pattern that was not depicted in the outcomes of the trajectory length (TL) (see Figure 8-19). On this account a separate repeated measures analysis was carried out with the within-subject factor spectral roc and the between subject factors group, training and frequency.

Figure 8-19 indicates some effects for /eɪ/ between the intervention and control group. Therefore, a separate repeated-measures ANOVA was carried out for the spectral roc of /eɪ/. Group, frequency and training were included as between subject factors. However, as Mauchly's test was significant, the results are not presented in this thesis.

8.3.3 Acoustic analysis of the approximant /w/

Similar to diphthongs, approximants are characterised by the time-varying shift of frequencies from a constriction to the following vowel. To assess the position and dynamic movements of the approximants in the F1 and F2-F1 plane, the acoustic measurements included the approximant-vowel duration which are presented in the first section of this chapter and the formant values of F1 and F2-F1 taken at the 20%, 40%, 60%, and 80% temporal points of the approximant and vowel to have a closer look at formant movements. The test items used to elicit the pronunciation data for the approximant /w/ were `one´, `swing´, `wait´ and `witch´. The approximant /w/ appears in word initial position as well as in a cluster, Figure 8-20 displays the mean approximant vowel durations for each test item by group and time.

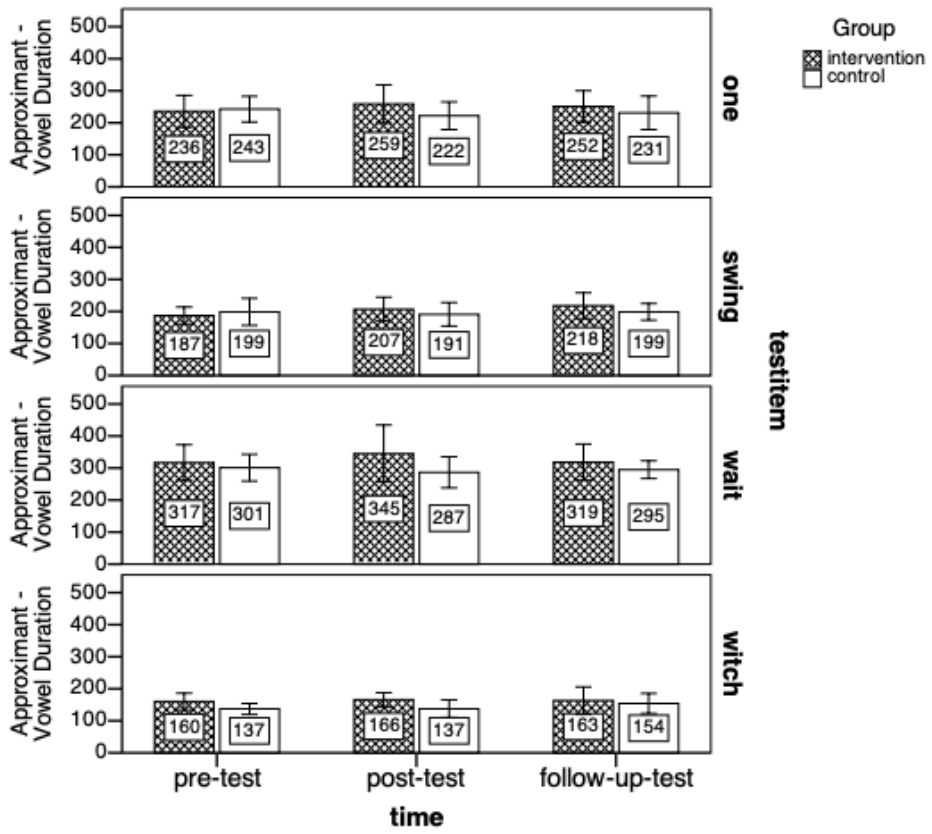


Figure 8-20: Mean approximant-vowel durations (in ms) and SEM by item, time and group

The data (see Figure 8-20) indicate systematic differences in the mean approximant-vowel durations between the four test items with `wait´ displaying the longest durations, followed by `one´, `swing´ and `witch´ (see Appendix Table 10-6 for mean approximant vowel duration and standard error).

The longer approximant vowel durations of `wait´ in contrast to the other test items is due to the fact that the /w/ is not followed by a vowel but by a diphthong. In the intervention group there is an increase in duration between the pre- (317.1ms) and post-test (345.4ms) and then the duration decreases again at the follow-up-test (318.8ms). The control group shows a reversed behaviour with the approximant vowel duration decreasing from the pre- (300.8ms) to the post-test (268.7ms) and then displaying a slight increase at the follow-up test (295.0ms). Overall, the control group displays shorter approximant vowel durations. The same duration pattern can be detected for the production of `one´. So similar to `wait´ the approximant vowel duration of `one´ produced by the intervention group increases between the pre- (235.8ms) and post-test (258.8 ms) and then there is again a slight decrease in the follow-up test (251.5 ms). With regards to the control group the approximant-vowel duration also decreases from pre- (242.7ms) to post-test (222.1ms) and then slightly increases again

in the follow-up test (231.3ms). Once more, the control group overall displays shorter durations in the post- and follow-up test. Yet, `wait´ and `one´ both groups display an increase in the standard error at the post-test demonstrating a non-stable production the test items.

The approximant-vowel duration of `swing´ produced by the intervention group increases over the three time points (187.1 ms, 207.1 ms, 217.9 ms) and so does the standard error (11.3, 15.7, 17.2). In the control group, the pre-test (198.8 ms) displays longer durations compared to the post-test (191.0 ms) and the durations of the follow-up test (198.5 ms) draw near the pre-test. However, the standard error of the control group decreases over the three data collection points (17.7, 15,5, 11.0). The trajectory analysis (see Figure 8-21) might reveal facts for a further investigation.

Similar to `swing´, the approximant in `witch´ is also followed by the vowel /ɪ/. The on the whole shorter approximant-vowel durations of `witch´ might be due to the fact that the vowel in `swing´ is followed by a voiced nasal in contrast to the voiceless post-vocalic affricate /tʃ/ in `witch´. With regards to the intervention group, there is an increase in the duration between the pre- (160.0 ms) and post-test (165.6 ms) and then a slight decrease in the follow-up test (162.9 ms). The standard error behaves conversely showing a more stable but yet longer production of `witch´ at the post-test. However, the data has to be treated with great caution as there is only a small amount of data and there are no remarkable differences. The control group displays similar durations at the pre- (137.1 ms) and post-test (137.5 ms) and there is an increase in duration at the follow-up-test (153.5 ms). The standard error increases over the three data collection points. Overall the control group displays shorter durations than the intervention group.

Due to the different coarticulatory contexts of the approximant in `one´, `swing´, `wait´ and `witch´, each test items needs to be examined separately. On this basis there is too little data to run an ANOVA. Therefore, similar to the diphthong analyses in Section 8.3.2.2, the trajectories of the mean formant movements of the central 60% of the approximant vowel section are used to analyse the data (see Figure 8-21).

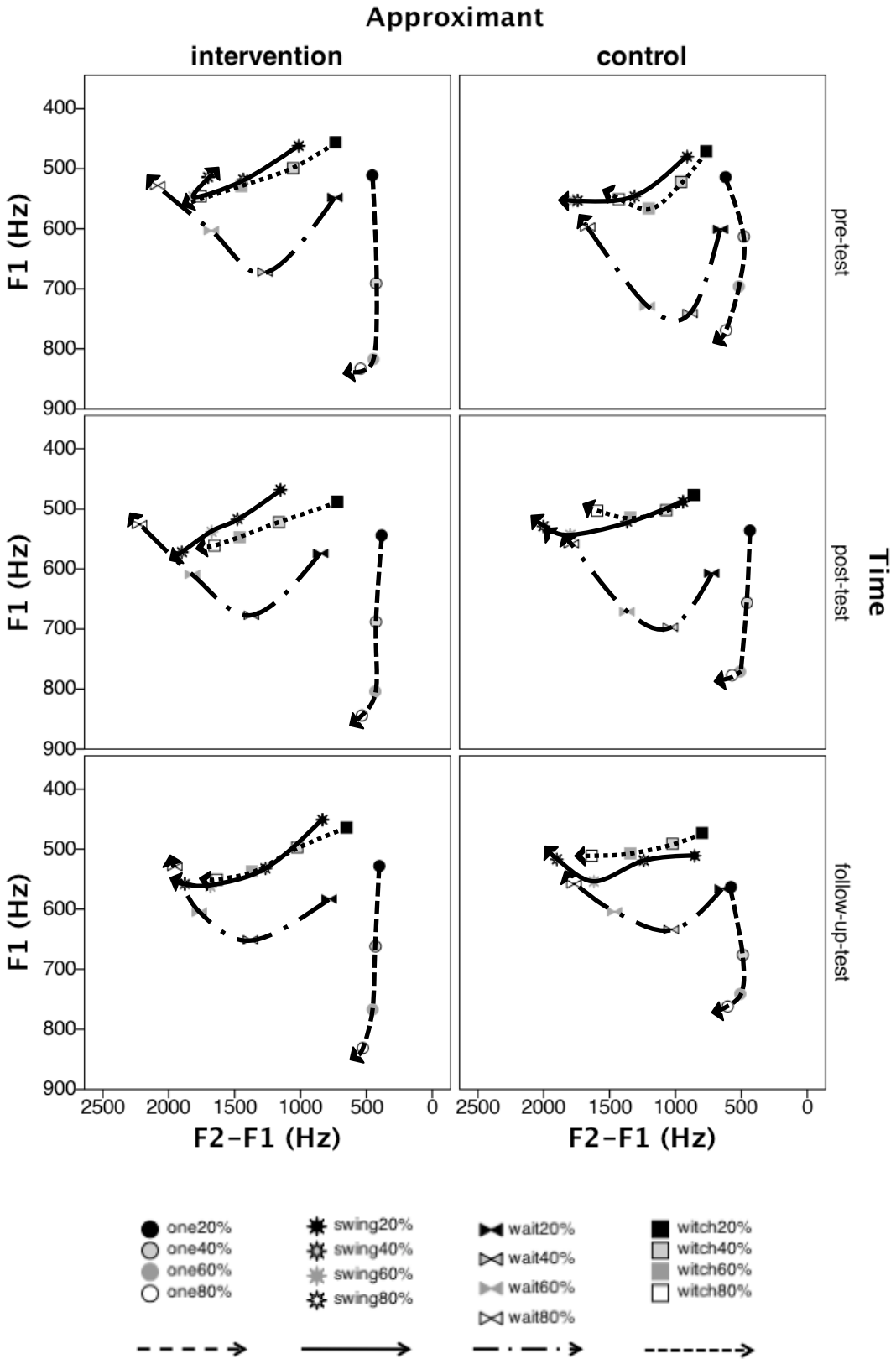


Figure 8-21: Mean formant movement (in Hz) of the central 60% of the approximant vowel portion for one, swing, wait and witch by group and time.

Figure 8-21 displays the relative positions and formant movement at four equidistant points of the central 60% of the approximant vowel portion of `one´, `swing´, `wait and `witch´ in the F1 x F2-F1 plane by time and group.

The visual inspection of the approximant vowel trajectory of `one´ produced by the intervention group shows a similar display of the trajectories at the pre-, post- and follow-up-test with the F2-F1 values for 20%, 40% and 60% section at around 500 Hz and a slightly more fronted production of the 80% of the vowel (about 600 Hz). However, the spacing between the four sections at 20%, 40%, 60% and 80% becomes more evenly distributed between the three time points. In comparison to the intervention group, the spacing of the approximant vowel portions of the control group is evenly distributed at the pre-test and becomes less regular from post- to follow-up-test. Moreover, the pre-test and the follow-up test of the control group reveals a more backed production of the 40% and 60% section of the approximant vowel. Although the F1 values of the first 20% section of the approximant and vowel are similar between both group (about 500 Hz), the F1 values for the 80% section of the control decrease between pre-, post- and follow-up test and are smaller compared to the intervention group.

The inspection of the trajectory of `swing´ produced by the intervention group displays similar F2-F1 values of the 20% section for the pre- and post-test but lower F2-F1 values at the follow-up test indicating a more backed pronunciation of the approximant. The pre-test reveals a backwards curve at the 60% to 80% section of the vowel. As the trajectory displays the mean group values, it seems that the subjects are unsure about the production of the vowel /ɪ/. The trajectory straightens at the post- and follow-up test. The 20% section of `swing´ pronounced by the control group displays similar F2-F1 values but slightly increasing F1 values throughout the three data collection points indicating a more open production of the approximant /w/ from pre- to post- to follow-up test. The F2-F1 values for the 80% section increase between pre- and post-test and then slightly decrease again at the follow-up test. This indicates a more fronted production of /ɪ/ at the post and follow-up test. Comparing the trajectories of both groups, the intervention group displays a more open production of the 20% section at the post- and follow-up test and there are especially changes in the production of the vowel /ɪ/ for both groups.

`Witch´ shows similar trajectories to `swing´. This can be expected as both test items include the same approximant vowel sequence /wɪ/. However, as already shown

above (see Figure 8-15) the mean duration for `witch´ are shorter compared to `swing´ due to the voiceless post-vocalic affricate. The intervention group displays similar F2-F1 values of the 20% portion of the approximant and vowel between pre- and post-test. However, they decrease at the follow-up test showing a more backed production of /w/ at the follow-up test. The F2-F1 values of the 80% section decrease slightly between pre- post- and follow-up- test showing a more backed production of /ɪ/. The F1 values for the 20% and 80% portion indicate a more open production at the post-test compared to the pre- and follow-up test for approximant /w/ and /ɪ/. The visual inspection of the approximant vowel trajectories of the control group reveals similar F2-F1 values for 20% portion of the approximant and vowel in `witch´. However, the F1 values increase slightly at the post-test revealing a more open production. Both, the F2-F1 and the F1 values for the 80% portion increase between pre- and post-test and present similar values for the follow-up-test showing a more fronted and open production of /ɪ/.

A visual inspection of `wait´ produced by the intervention group reveals similar F2-F1 values for the 20% section of /weɪ/. The F2-F1 values at 80% increase between pre- and post-test and then decrease at the follow-up-test. This shows a more fronted production of /ɪ/ after the intervention. The F1 values at 20% and 80% increase over the three data collection showing both the greatest increase after the intervention indicating a more open pronunciation of /w/ and /ɪ/. The control group displays similar F2-F1 values at the 20% portion of /weɪ/ over the three data collection points. Additionally, the F1 values at the 20% portion do not show much change between the pre-and post-test. However, there is a slight decrease at the follow-up test resulting in a more close production of /w/. The F2-F1 values of the 80% section increase slightly between the pre- and post-test and then show similar values at the follow-up test. However, the F1 values decrease over the three data collection points. This results in a more close production of /ɪ/. Comparing both groups, all six trajectories display a curved trajectory with the highest F1 values for the 40% section of the approximant vowel. Both, the intervention and the control show decreasing F1 values over the three data collection points. Overall, the intervention group shows lower F1 values. A reason for the more closed production of the 40% portion of /weɪ/ might be that in the German language most diphthongs shift to /ɐ/ (see chapter 3.5.2).

Summarising the presented information above, the approximant vowel trajectories do not reveal a lot of change in the approximant /w/ but in the production of the following vowel.

8.3.4 Acoustic analysis of the plosives in final position

In the German language final plosives are always devoiced. As voiced plosives in post-vocalic position are preceded by considerably longer vowels than their voiceless counterparts (Pickett 2001), the following section looks at the vowel duration data which are displayed in Figure 8-22.

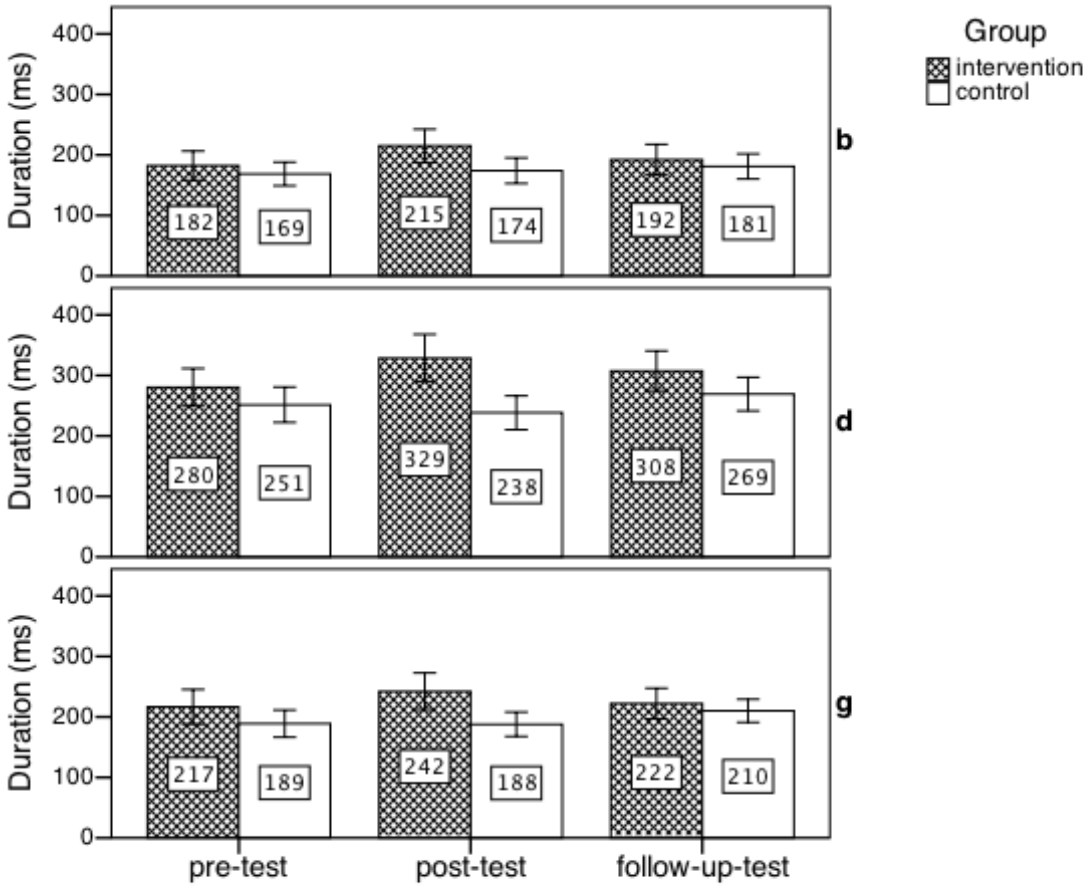


Figure 8-22: Mean pre-plosive vowel durations and SEM by time, group and plosive

Figure 8-22 show that the mean pre-plosive vowel durations by time group and plosive (see Appendix Table 10-7 for mean pre-plosive approximant vowel duration and standard error).

The mean values for the pre-vocalic vowels produced for /b/ for the intervention group increase from the pre-test (182.4 ms) to the post-test (215.0 ms) and then decrease again at the follow-up-test (192.2 ms). In contrast, the pre-plosive vowel duration of the control group increases over the three data collection points (168.6 ms to 174.0 ms to 181.1 ms).

The intervention group shows a similar pattern for /b/ and /d/ with a strong increase at the post-test (from 280.5 ms to 328.7 ms) and a decrease in duration at the follow-up-test (307.6 ms). However, for the control group, there is a small decrease between pre-and post-test (251.5 ms to 238.5 ms), and then the durations increase again at the follow-up test (269.3 ms).

Analogous to /b/ and /g/, the pre-vocalic durations produced for /g/ by the intervention group show the highest increase at the post-test (from 216.8 ms to 242.1 ms) and again a decrease at the follow-up test (222.3 ms). The pre-plosive vowel durations for /g/ produced by the control group behave in a manner similar to the values produced for /d/ with a slight decrease between pre- (188.9 ms) and post-test (187.8 ms) and an increase in duration at the follow-up-test (210.0 ms).

The presented results reveal that the shortest mean pre-plosive vowel durations are pronounced for /b/ followed by /g/ and /d/. In addition, the data show that there is an increase in vowel duration for all of the three plosives /b/, /d/ and /g/ over the three data collection points. However, the strongest increase can be observed for the pre-plosive vowel duration at the post-tests of the intervention group. A mixed model analysis of variance with the within subjects factors time, frequency, training and diphthong, and the between subjects factor group (intervention vs. control) was carried out to have a closer look at the differences in pre-plosive vowel duration. The ANOVA revealed (see Table 8-20) that within subject effects of training, frequency and plosive were significant with strong effect sizes. However, there was no significant main effect of time, $F(1.621, 21.078) = 1.756$, $p = .200$, $\mu = .119$, and group, $F(1,13) = 2.964$, $p = .109$, $\mu = .186$.

Table 8-20: Significant within subjects effects for the plosive data

Effects[57]	df	Error	Mean Square	F	Sig.	Partial Eta²
training (trained/ not trained)	1	13	106283.190	66.376	.000	.836
frequency (higher/ lower frequency)	1	13	208300.495	85.925	.000	.869
plosive (/b/, /d/, /g/)	1.990	25.868	429915.797	99.637	.000	.885
training x frequency	1	13	172739.560	101.106	.000	.886
training x plosive	1.616	21.009	72299.343	35.200	.000	.730
frequency x plosive	1.871	24.322	100707.046	49.732	.000	.793
training x frequency x plosive	1.712	22.251	267843.754	89.763	.000	.873

The significant main effect of plosive (F (1,13) = 66.376, $p < 0.001$, partial $\eta^2 = .836$) showed the key differences in the vowel durations preceding the three plosives increasing from /b/ (M = 184.609 ms, SE = 8.534) to /g/ (M = 210.416 ms, SE = 12.603), (*M* = 25.807 ms, SE = 6.698, $p < 0.001$) and to /d/ (M = 279.130, SE = 12.501), (*M* = 68.713 ms, SE = 6.913, $p < 0.001$). This result supports the results of the descriptive statistics (see above).

The strong impact of the factors training and frequency (see chapter 5.5) is shown in the significant main effects of training (F (1,13) =66.676, $p < 0.001$, partial $\eta^2 = .836$) with shorter durations for the trained (M = 210.658 ms, SE = 10.765) compared to the untrained items (M = 238.779 ms, SE = 10.806), (*M* = 28.121, SE = 3.452, $p < .001$) and frequency (F (1,13) = 85.925, $p < 0.001$, partial $\eta^2 = .869$) with shorter durations for the higher frequency items (M = 205.034 ms, SE = 10.519) in comparison to the lower frequency items (M = 244.403 ms, SE = 11.183), (*M* = -39.368 ms, SE 4.247, $p < 0.001$).

Similar to the ANOVA carried out for the vowels and diphthongs (see Table 8-13 and Table 8-18) the presented results have to be interpreted with great caution due to the biases in the nature of the test items, and the different coarticulatory contexts of the plosives (see Table 7.3.3). It is for this reason, that the significant interactions (see Table 8-18) are not discussed in further detail.

8.3.5 Acoustic analysis of the fricatives /z/, /θ/ and /ð/

This study looks at the interdental voiceless and voiced fricatives /θ/ and /ð/, and the voiced alveolar fricative /z/. Fricative noises are produced when a turbulent airflow hits a downstream obstacle (Johnson, 2005). The acoustic consequence is aperiodic noise whose spectral shape is determined by the length of the cavity in front of the

[57] Effects are adjusted with the Greenhouse-Geisser correction.

constriction. Therefore, the fricative data is analysed by its `centre of gravity´ (see Section 5.4.3) which reflects the formant structure of the dominant front cavity (Wrench, 1995). In addition to fricative noise, voiced fricatives also show evidence of periodicity (Harrington 2013). However, the total airflow of voiced fricatives is lower as the vocal folds are held closer to produce the voicing and thus their intensity is weaker (Pickett 2001, p. 141). Moreover, the fricative durations are shorter for voiced fricatives, but similar to the plosive features (see Section 8.3.4) the duration of the preceding vowel is longer for voiced fricatives compared to their voiceless counterparts (see Section 5.4.3).

This section first presents the mean vowel and fricative durations and then goes on to have a closer look at the fricative noise. The fricatives appear in the test items in pre- and in post-vocalic positions which affect the durational measures. Therefore, Table 8-21 presents the mean vowel and pre-vocalic fricative durations of the test items `zoo´, `that´, `them´, `the´, `thing´, and `thin´ and Table 8-22 the mean fricative and post-vocalic vowel durations of the test items `close´, `jeans´, `goes´, `smooth´, `month´ and `teeth´.

Table 8-21: Mean vowel and pre-vocalic fricative durations (in ms) and SE of `zoo´, `that´, `them´, `the´, `thing´ and `thin´ by time and group.

	Intervention				Control			
	Mean vowel duration (ms)	SE	Mean fricative duration (ms)	SE	Mean vowel duration (ms)	SE	Mean fricative duration (ms)	SE
pre-test	199.0	9.9	131.2	9.4	184.5	7.3	104.2	8.0
post-test	215.7	10.8	147.2	10.2	176.8	7.8	107.7	8.3
follow-up-test	199.3	9.9	138.2	9.4	188.3	7.4	104.8	8.6

Table 8-21 shows an increase in the mean vowel duration produced by the intervention group from pre- (199.0 ms) to post-test (215.7 ms) and a reversed behaviour for the follow-up-test (199.3 ms). The same pattern is displayed for the mean fricative durations (131.2 ms to 147.2 ms to 138.2 ms) and even for the standard errors. The mean fricative durations of the control group behave similarly with slightly increasing durations from pre- (104.2 ms) to post-test (107.7 ms) and a decrease (104.8 ms) at the follow-up test. However, the mean vowel durations display a reversed behaviour, as the mean vowel durations show a small decrease between pre- (184.5 ms) and post-test (176.8 ms) and then increase again at the follow-up-test (188.3 ms). Comparing

both groups, the intervention group displays overall longer duration and the changes in durations are smaller for the control group.

The second data subset includes fricatives in post-vocalic positions. It contains the following test items: `close´, `jeans´, `goes´, `smooth´, `month´ and `teeth´ (see Table 8-22).

Table 8-22: Mean vowel and post-vocalic fricative durations (in ms) and SE of `close´, `jeans´, `goes´, `smooth´, `month´, and `teeth´ by time and group.

	Intervention				Control			
	Mean vowel duration (ms)	SE	Mean fricative duration (ms)	SE	Mean vowel duration (ms)	SE	Mean fricative duration (ms)	SE
pre-test	286.2	9.9	226.3	6.7	270.1	10.2	199.3	7.0
post-test	326.4	13.6	231.0	6.8	268.0	9.2	208.3	8.2
follow-up-test	284.4	9.3	221.4	7.0	277.1	8.1	198.4	7.3

The mean vowel durations of the intervention group display an increase from the pre- (286.2 ms) to the post-test (326.4 ms) and then decrease again at the follow-up-test (284.4 ms). The mean fricative durations also increase slightly at the post-test (from 226.3 to 231.0) and reverse again at the follow-up-test (221.4 ms). Similarly, the mean fricative durations of the control group increase at the post-test (from 199.3ms to 208.3 ms) and decrease again at the follow-up-test (198.4 ms). Conversely, the mean vowel durations decrease slightly from pre- (270.1 ms) to post-test (268.0 ms) and show an increase at the follow-up-test (277.1 ms). Overall, the intervention group displays longer mean duration than the intervention group.

Comparing the two data subsets, both groups display a similar behaviour notwithstanding whether the fricative appears in pre- or post-vocalic position. Figure 8-23 gives an overview of the mean fricative and vowel durations for /θ/, /ð/ and /z/ by time and group. Similar to the mean values presented in Table 8-21 and Table 8-22, the mean fricative and vowel durations for each of the fricatives /z/, /θ/ and /ð/ of the intervention group show the highest values at the post-test and reverse to the pre-test at the follow-up-test. This pattern can also be found for the fricative durations of /θ/. The mean vowel durations of voiceless /θ/ do not change a lot between pre- and post-test and then increase at the follow-up test. In contrast, the vowel durations for the voiceless fricative /ð/ decrease between pre-and post-test and then increase at the follow-up-test whereas the respective fricative durations decrease to some extent from

pre- to post- to follow-up test. On the contrary, the vowel and fricative durations for /z/ increase slightly over the three data collection points.

Comparing all the data, the intervention group overall displays longer fricative and vowel durations than the control group. Moreover, the duration values for /z/ are higher than for /θ/ and /ð/. The mean vowel durations for the voiced fricative /ð/ are only slightly higher than for /θ/. However, the mean fricative durations are considerably lower for the voiced /ð/ than for /θ/. This difference is even higher for the fricative durations produced by the intervention group.

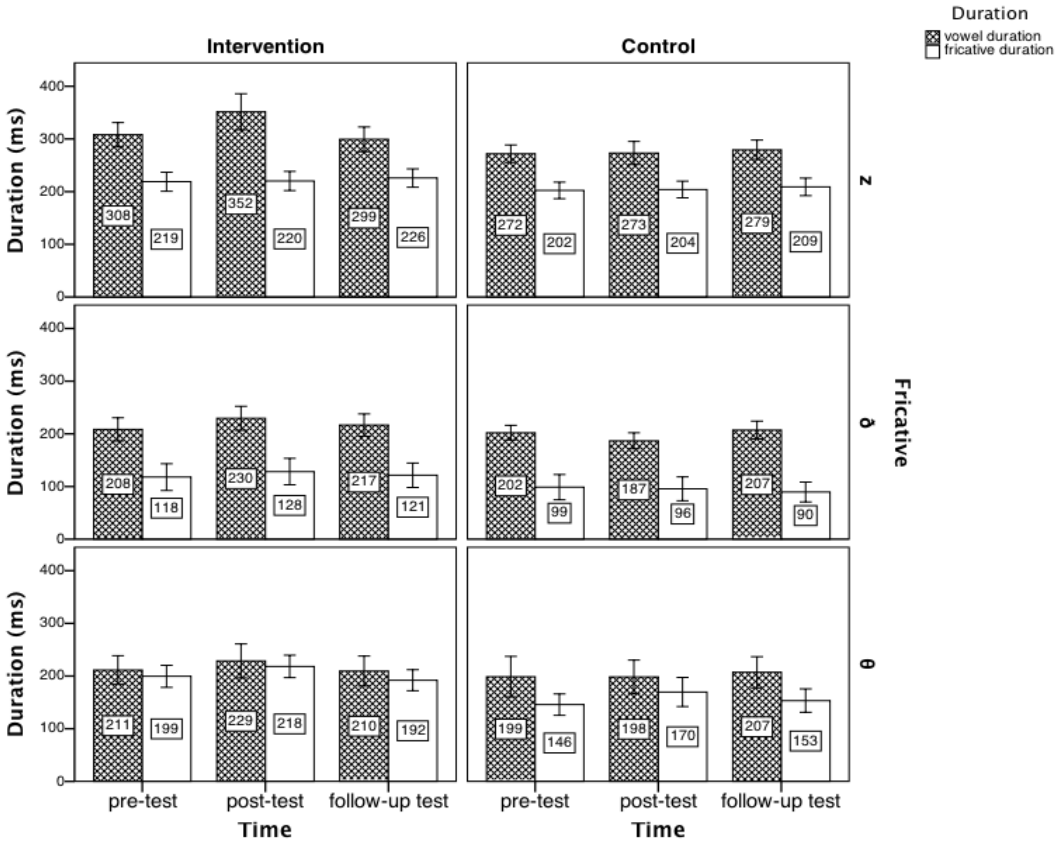


Figure 8-23: Mean fricative and vowel duration durations (in ms) and SEM by fricative, time and group

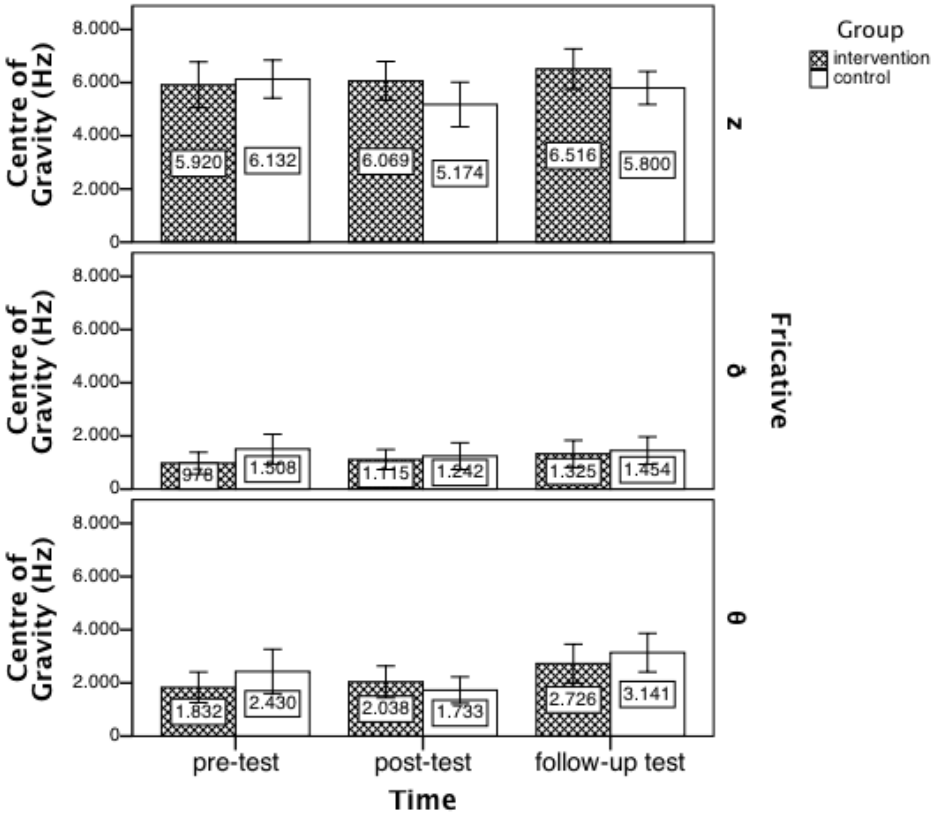


Figure 8-24: Mean centre of gravity (in Hz) for the fricatives /z/, /ð/ and /θ/ by time and group

Figure 8-24 displays the mean centre of gravity of the fricatives /z/, /θ/ and /ð/ by time and group. It is evident that the alveolar /z/ displays significantly higher centre of gravity values than the two dental fricatives /θ/ and /ð/. However, the voiceless /θ/ also shows a higher centre of gravity than the voiced /ð/. For all three fricatives, the values of the intervention group increase over the three data collection points. The centre of gravity values of the control group decrease at the post-test and then increase again approaching the pre-test values for /z/ and /ð/ and surpass the pre-test for the voiceless /θ/.

8.3.6 Acoustic analysis of the affricate /dʒ/

Affricates are seen as an additional manner of fricative articulation but in contrast to fricatives, affricates are preceded by an occlusion instead of more open articulation (Pickett, 1999). With reference to the acoustic analysis, the affricate /dʒ/ which was part of the intervention, is treated similarly to the fricatives (see Section 8.3.5 above) and therefore, the durations of the preceding or following vowels and the affricate durations are presented in the first part of this section (see Figure 8-25) and after that the centre of gravity measurements (see Figure 8-26) will be taken into account. Four test items

were used to elicit the affricate data: `job´, `jeans´, `John´ and `cage´. As the affricate appears in pre- and post-vocalic positions and in different coarticulatory contexts, the data will be displayed for each test item.

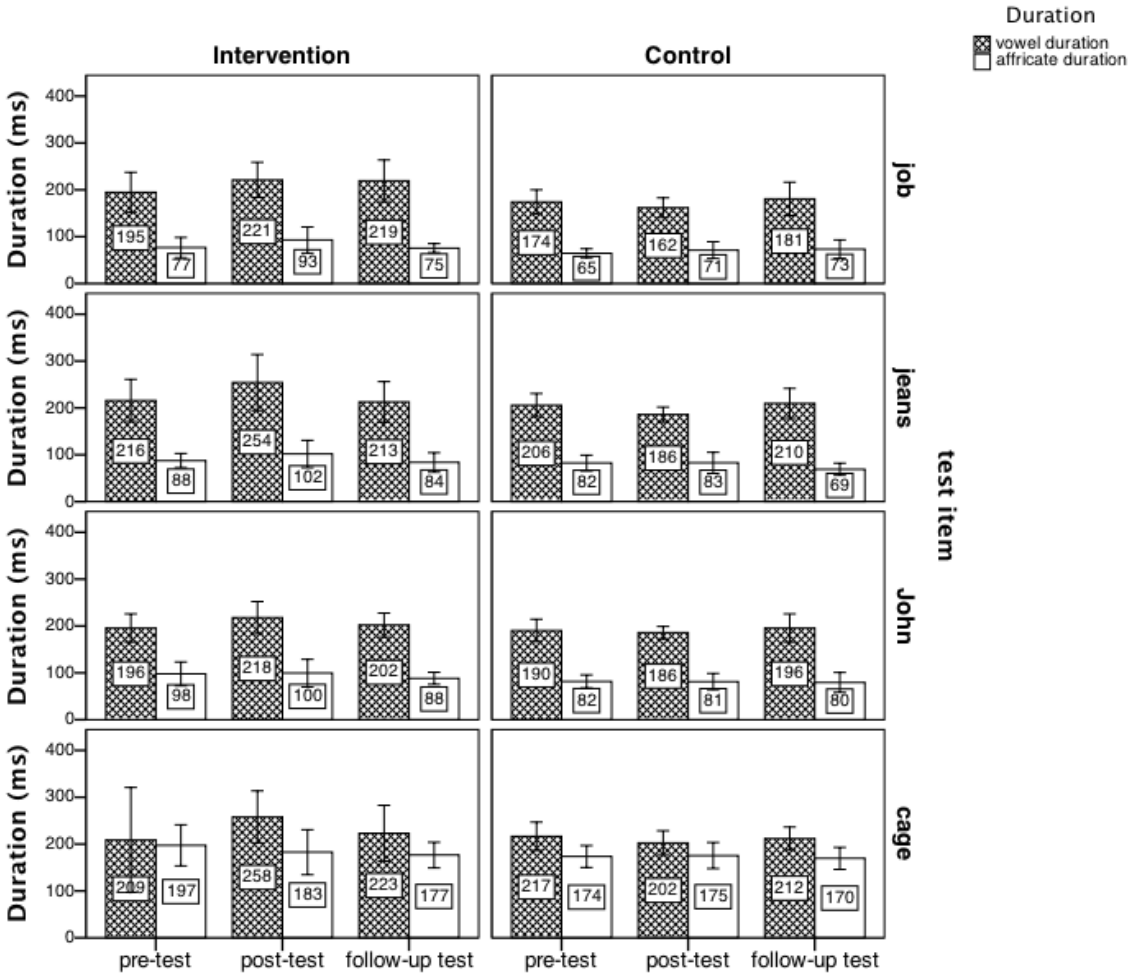


Figure 8-25: Mean vowel and affricate durations (in ms) and SEM by test item, group and time.

Figure 8-25 displays the mean vowel and affricate durations by test item, group and time. With regards to the mean vowel durations of all test items of the intervention group, the data shows a considerable increase between pre- and post-test and then a reversed decrease for the follow-up-test. The affricate data of the intervention group displays overall smaller durations but the same pattern with the largest duration values at the post-test with the exception of `cage´. The affricate durations for cage decreases slightly over the three data collection points. Then again, the data for `cage´ displays quite large confidence intervals in comparison to the other three test items.

The mean vowel durations produced for all test items by the control group show the reversed behaviour of the intervention group with a decrease between pre-and

post-test and an increase at the follow-up-test. However, the affricate durations display different patterns for each test item. With regards to `job´ there is a slight increase over the three data collection points. The affricate duration for `jeans´ increases slightly between pre- and post-test and then decreases at the follow-up test. A similar pattern can be seen for `cage´ and the affricate durations for `John´ decrease slightly from pre- to post- to follow-up-test. Despite the presented results, the fricative durations of the control group only show slight changes over the three data collection points and due to the small amount of data, the outcomes have to be treated with great caution.

Similar to the outcomes of the fricative data (see chapter 8.3.5), the vowel and affricate durations of the intervention group are on the whole longer than for the control group. Moreover, with the exception of `cage´ the vowel and fricative durations of the intervention group show similar patterns with an increase from pre- to post-test and a decrease at the follow-up-test. Again, similar to the fricative data, the reversed behaviour for the vowel durations of the control group takes place.

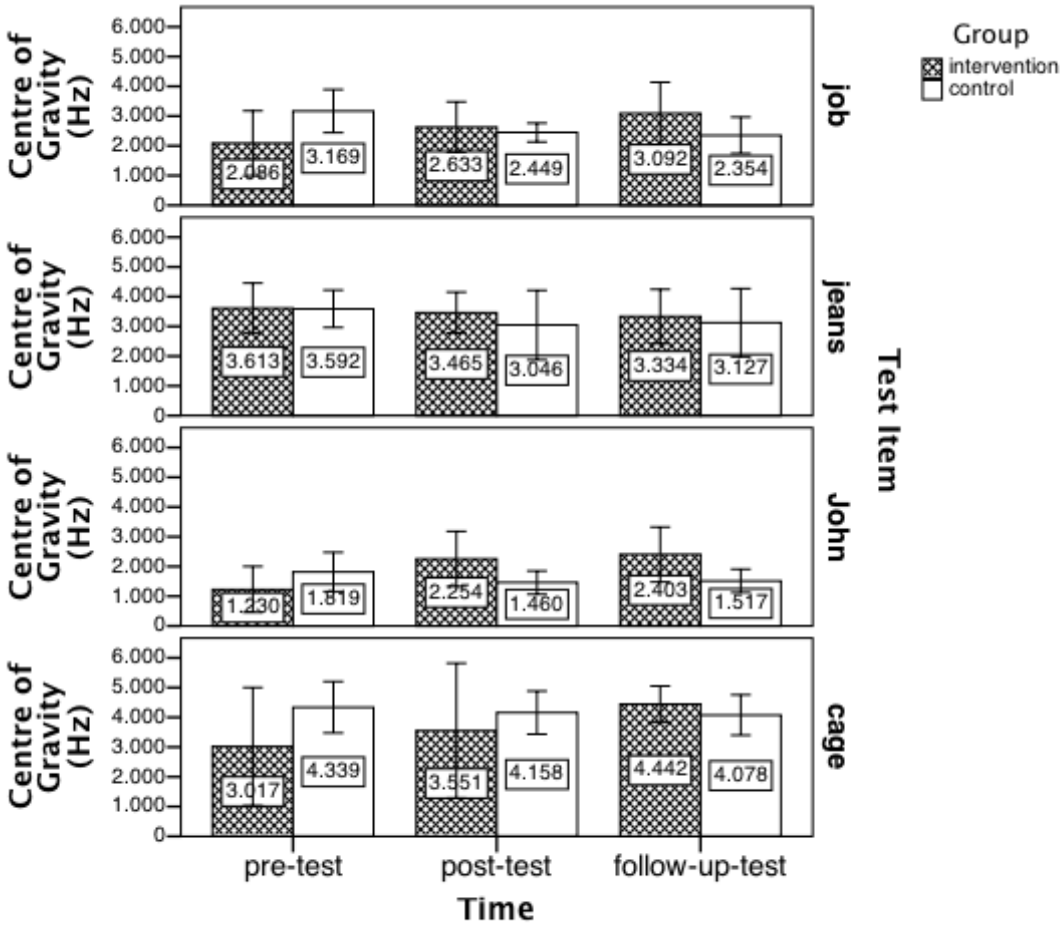


Figure 8-26: Mean centre of gravity (in Hz) and SEM for the affricate /dʒ/ by time, group and test item

Figure 8-26 displays the mean centre of gravity for the affricate /dʒ/ in the test items `job´, `jeans´, `John´ and `cage´.

With regard to `job´ produced by the intervention group there is an increase in the centre of gravity values over the three data collection points. However, the control group displays a reversed behaviour. Similar to `job´ the centre of gravity values for `jeans´ the intervention group increase from pre- to post- to follow-up-test. With regard to the control group, there is a decrease between pre- and post-test and the values increase again at the follow-up test. The same pattern is displayed for the centre of gravity of `John´ produced by the control group while the values of the intervention group increase over the three data collection points. In a similar fashion, the centre of gravity increases from pre- to post- to follow-up for the test item `cage´ produced by the intervention group. However, the values of the control group decrease over the three data collection points.

On the whole, the highest values for the centre of gravity are displayed for `cage´, followed, by `jeans´, `job´ and `John´. For all three test items the biggest increase in centre of gravity can be found at the post-test of the intervention group whereas there is a decrease in the centre of gravity values of the control group between pre- and post-test.

8.4 Discussion acoustic analysis

The previous section 8.3 presented the results of the acoustic analysis. This section follows on to discuss the results.

8.4.1 Vowels /ɑː/, /ɔː/, /ə/ and /æ/

The vowels /ɑː/, /ɔː/, /ə/ and /æ/ were selected for the pronunciation intervention programme (see Chapter 4). A look at the mean durations revealed a robust effect between the long vowels /ɑː/, /ɔː /and short vowels /ə/ and /æ/ (see Table 8-12). Moreover, an analysis of the vowel duration by group and time showed an increase in duration between the pre- and post-test phase. Despite the assumption that durations should decrease after training due to increased motor control (see Section 3.3.2) and dual route effects (see Section 5.5), the data show a different picture. In addition, the standard error increases between pre- and post-test for both groups. A reason for the increase in duration and error might be the tension of the subjects during data collection or a hyperarticulation of the test items due to the test scenario. Both are known to affect vowel durations. Regardless of this increase, the data also show some patterns with regards to vowel durations: Whereas the vowel duration of the control group only shows slight differences between pre- and post-test and a further increase in duration between post- and follow-up-test, the duration of all vowels in the intervention group increased between pre- and post-test and decreased between the post- and the follow-up-test indicating that the biggest change takes place after the intervention and that maturation effects take place at the follow-up-test. An ANOVA run on the duration data revealed statistically significant effects for training, frequency, and vowels. However, the group effect did not show significant results. Although the duration gives us some information, the first two formant frequencies mainly account for vowel perception. Therefore, the F1/ F2-F1 spaces are looked at with more detail.

In contrast to the ANOVA, the vowel spaces show apparent differences by group and between the three data collection points (see Figure 8-11). Comparing both groups, the vowel spread clearly displays a more variable and disrupted production of the vowels (with the exception of /ə/) for the control group in comparison to the intervention group at the pre-test. Similar to the results of the auditory analysis, this indicates a bias in the matching of both groups. Overall, for both groups all vowel ellipses are getting smaller from pre- to post-test indicating more stability in the pronunciation of the vowels. With the exception of /æ/ this trend continues in the follow-up-test with the biggest shift between pre- and post-test. This impression is supported by the decreasing standard deviations in the intervention group (see Table 8-14). This might point towards the fact that the intervention programme might have had an effect. Although there is a similar trend in control group and the decrease in SD values for F2mid, they increase for F1mid showing more spread in the open-closed dimension. As the four different vowels do show slightly different pictures, the following paragraphs look at them one after the other.

The dispersion of the vowel production of /æ/ increases in the follow-up-test for control group and reveals a rather neutral production of the vowel while the production of /æ/ by the intervention group is still rather neutral but breaks away from the production of /ə/. This might be due to the fact that the front open position of the German vowel space is not occupied (see Figure 3-3) the subjects might confuse /æ/ with the less open /e/. In addition, the overlap in the vowel spaces diminishes clearly in the intervention group. This shows a higher awareness of the English vowels and an increase in English vowel space. As /æ/ is not part of the German sound inventory (see Figure 3.3.), it seems that the control group replaces the vowel with the German less open vowel /e/. Although the intervention and control group both pronounce the /æ/ with similar fronting at the follow-up-test, the vowel remains more open in the intervention group which corresponds to the English target pronunciation of /æ/.

The control group pronounces the vowel /ɑː/ with a comparable backwards constriction but with a less open quality. Throughout the three test periods there is a slight shift towards a more anterior pronunciation of the /ɑː/ but it becomes more constricted and loses its open quality being pronounced rather neutrally. This might be due to the fact that there is no open back German vowel and the control group might confuse it with the less open German /ɔ/ (see Figure 3.3).

Moreover, the increase in the Euclidean distance between open-back vowel /ɑː/ and back vowel /ɔː/ (134Hz, 142Hz, 174 Hz) shows the spread of the vowel space for the intervention group (see Table 8-16). The increasing vowel spaces might be due the fact that the children grow and as a consequence their vowel spaces increase. Another reason might be the hyperarticulated pronunciation of the test items (Harrington, 2013, p. 91f) and more confidence in the production of the sounds (Banse & Scherer, 1996, p. 615)

An inspection of the vowel spaces of the control group indicates that vowels shift apart in the anterior and posterior direction, especially after the post-test. However, with regard to the open closed dimension, all vowels seem to be pronounced with a more closed quality after the follow-up-test. This is also supported by the upwards shifting centroid and a smaller Euclidean distance between the open-back vowel /ɑː/ and neutral –back vowel /ɔː/ between the post- and follow-up-test (644 Hz, 604 Hz). The production of more lax vowels points to a less hyperarticulated speech (Lindblom, 1996) and might also indicate and insecure pronunciation (Harrington, 2013, p. 92).

8.4.2 Diphthongs /ɪə/ and /eɪ/

The diphthongs /ɪə/ and /eɪ/ were trained in the intervention. Table 8-17 presents the mean diphthong durations for the two diphthongs by group status. On the whole, longer mean diphthong durations and standard errors were displayed for /ɪə/ compared to /eɪ/ supporting the fact that /ɪ/ and /ə/ are further apart in the vowel space than /e/ and /ɪ/.

With regards to /ɪə/, the diphthong duration produced by the intervention group decreases from pre- to post-test (395 ms to 376 ms), and there is a slight decrease in the follow-up test (374 ms). The strong decrease between the pre- and post-test might indicate more automated motor control processes and thus a higher articulation rate which can be a result of the intervention, whereas the slight decrease in the follow-up test might be due to maturation processes. The diphthong duration of /ɪə/ produced by the intervention group decreases from pre- to post-test (395 ms to 376 ms), and there is a slight decrease in the follow-up test (374 ms). Similar to the intervention group, the duration decreases in the control group between the pre- and post-test (332 ms to 320 ms). However, the follow-up-test reveals an increased diphthong duration (351 ms) for /ɪə/. Whereas the decrease might be explained by maturation, it seems that the students hyperarticulated the test items in the follow-up-test or attempted new ways to

articulate the diphthong. It is also possible that they have to adjust their articulatory movements to their growing body and as a consequence adapt the movements of their articulators. In both groups the standard error decreases slightly from pre- to post-test and then increases in the follow-up-test indicating that both groups present less stable versions of the diphthong. Comparing both groups, the control group overall shows shorter diphthong durations. This might be due to the fact that the intervention group is more aware of the pronunciation project and therefore might try to pronounce the test items particularly carefully.

The mean diphthong durations for /eɪ/ show a similar picture, with the mean durations being higher for the intervention (309.67 ms) than the control group (253.34 ms). The diphthong durations of the production of /eɪ/ of the intervention group increase between pre- and post-test (301ms to 320 ms) and then decrease again in the follow-up test 308ms. This rise in duration might indicate that the learners changed their way of pronouncing the diphthong between the pre-and post-test and then produced a more stable version of /eɪ/ in the follow up test.

To determine the differences in diphthong duration, a mixed model ANOVA with the within subjects factor time, frequency, training and diphthong, and the between subject factor group was carried out (see Table 8-18). The main effect of group revealed that the intervention group produced significantly longer durations than the control group. This might be due to the fact that the intervention group was more aware of the research project and hyperarticulated the test items. The significant main effect of diphthong showed significantly longer durations for /ɪə/ than for /eɪ/. This result supports the fact that /eɪ/ has a shorter trajectory length in comparison to /ɪə/ and therefore smaller articulatory movements are necessary to pronounce /eɪ/. The strong impact of the factors training and frequency (see chapter 5.5) is shown in the significant main effects of training with longer duration for the trained compared to the untrained items. Moreover, the frequency effect revealed shorter durations for the higher frequency items in comparison to the lower frequency items. The longer duration for the trained items points to the fact that the students might have pronounced the test items trained the intervention extra carefully. However, the shorter durations for the higher frequency items that also included trained and untrained items indicate more automated articulatory movements and thus learning effects (see Section 5.2.1.2)

The relative positions of the diphthongs in the F1 / F2-F1 space and the formant movements are displayed in Figure 8-13 to Figure 8-16. Figure 8-13 shows the scatter

plots of the 60% portion of /ɪə/ by group and time. The cluster of the control group becomes slightly denser at the post- and follow-up test. However, the production of /ɪə/ shows more variation in comparison to the intervention group. This variation is especially evident in the production of the 60% and 80% point of the diphthong and might be due to fact that some learners might confuse the English /ɪə/ with the more open diphthong /ɪɐ/. The decreasing spread of the control and intervention group indicates a more stable production of the diphthongs and indicates increasing motor-control abilities. This holds especially true for the intervention group. However, the intervention group already shows less scatter in the pre-test. This might be due to biases in the set-up of the groups. Figure 8-14 shows substantial differences in the mean relative formant positions of /ɪə/ between the intervention and control group at the pre-test. Compared to the intervention group, the formant positions of the control group are less fronted at the opening phase of the diphthong and the trajectory ends in a more open position. This might be due to the fact that the control group confuses the English /ɪə/ with the German /ɪɐ/. Moreover, the trajectory of the diphthong produced by the control group is smaller in comparison to the invention group which shows the more stable and longer production of /ɪə/ of the intervention group.

The spectral roc analysis (see Figure 8-19) reveals systematic differences between /ɪə/ and /eɪ/. On the whole, that /ɪə/ displays higher spectral rocs than /eɪ/. A reason for this difference is that /ɪə/ has a longer trajectory length than /eɪ/ and therefore faster articulatory movements are needed to reach the vowel target. Moreover, the test items for /ɪə/ have similar articulatory contexts: `year´, `dear´, `hear´, and `beer´. However, the test items for /ɪ/ are `page´, `cage´, `wait´ and ´grey´. In the first two items the diphthong is followed by a voiced affricate and in `wait´ there is a post-vocalic voiceless plosive. In `grey´ the diphthong is in final position. The difference in the production of /eɪ/ between the control and intervention group might be a result of hyperarticulation.

8.4.3 Approximant /w/

The test items containing the approximant /w/ are `witch´, `swing´, `one´ and `wait´. As expected, the measured approximant-vowel duration increased according to the coarticulatory context from `witch´, that entails a voiceless post-vocalic affricate, to `swing´, that includes a nasal, to `one´, that entails an open mid vowel, and `wait´ that

contains a diphthong. With regards to the intervention group, the durational patterns are similar to the ones displayed by the vowel and diphthongs with an increase in duration between pre- and post-test and a slight decrease at the follow-up test. On the contrary, with the exception of witch, that does not show much change in the intervention group either, the durations decrease between pre- and post-test. Overall the control group displays shorter durations than the intervention group. Similar to the diphthong analysis, the trajectories of central 60% of the approximant vowel section reveal more information. Therefore, the following paragraph discusses the trajectories for each test item at the three data collection points (see Figure 8-21).

The trajectory for `witch´ in the intervention group displays evidence of equal spacing at the pre-test. At the post-test the trajectory is slightly shortened with biggest space between the 20 - 40% portion of the vowel, indicating a focus on the glide between the approximant and the vowel. This might be due to the intervention that focusses on the glide. The follow-up-test shows the longest trajectory with equal spacing indicating the use of a bigger vowel space which might be due to maturation processes. The control group displays a slightly curved trajectory at the pre-test with the /ɪ/ being pronounced slightly less tense in comparison to the intervention group. The production of the /wɪ/ is more neutral and gets slightly tenser at the post-test. The follow-up-test of the control group shows a fairly similar picture, with a slightly more anterior production of the /w/ and more equal spacing. On the whole, the pronunciation of /w/ in `witch´ does not seem to pose a lot of problems for both groups, which is also indicated by the auditory analysis in which /w/ got the third highest mean pronunciation score (see Figure 8-2). This might be due to the fact that /w/ is part of the used curriculum and text books. The equal spacing at the follow-up-test indicates, that maturation takes place for both groups.

As swing also entails the /wɪ/ cluster, the trajectories look - as expected - rather similar. The trajectory displayed at the pre-test by the intervention group shows a backwards curve at the 60%-to 80% section. As the trajectory entails the mean formant movement, this curve might be due to the data of some subjects who might have problems with the pronunciation of the nasal or just with the item itself. At the post-test the production of /ɪ/ changes and a rather straight trajectory is produced which gets even longer and more equally spaced at the follow-up-test with a slightly more posterior and closed production of /w/. The 20% portion of the trajectory is produced like the /ʊ/. The trajectory of /wɪ/ pronounced by the control group displays similar F2-F1 as well as

F1) values for the 20% section of the /wɪ/ with a slightly more open production over the three data collections points. However, the 80% section of the trajectory the post- and follow-up-test is produced more tense resulting a longer trajectory and thus bigger vowel space. The curve displayed at the follow-up-test is probably due to individual differences in the mean data. Comparing the trajectories of both groups, the intervention group displays a slightly more open production of the 20% section, at the post- and follow-up test and there are especially changes in the production the vowel /ɪ/ for both groups.

The trajectories for the stimulus `one´ look fairly similar between pre-test, post-test and follow-up-test in the intervention group. However, the spacing between the four sections becomes more evenly distributed over the three data collection points indicating confidence in the pronunciation of the testitem. In contrast, the spacing of the control group shows a reversed picture. Additionally, F1 is decreasing between pre-test, to post-test, to follow-up-test which results in smaller trajectories. This leads to a less open production of the /wʌ/ in `one´. Either strong individual differences are responsible for this development or the subjects are mumbling as the very close spacing between the 60-80% portion of the approximant indicates. The effect might also be due to an antiformant as a result of the nasal in `one´ and lip rounding which results in a lower velum.

Contrary to the stimuli for /w/ presented above, in `wait´ the approximant is followed by a diphthong (/weɪ/) and not by a vowel. Thus, all six trajectories are curved with the highest F1 values for the 40% section of the approximant vowel. Both, the intervention and the control group show decreasing F1 values over the three data collection points. As most German diphthongs shift to /ɐ/ that has very low F1 values (see Figure 3-3), learning effects of the English pronunciation might take place. In contrast to the intervention group, the control group shows a more open production of the 40% section - especially at the pre-test. This might be due to the reading of the testitems and a more German accented pronunciation of `wait´.

Looking at the mean formant movement of the central 60% of the approximant vowel portion for `witch´, `swing´, `one´ and `wait´ at the three data collection points, it becomes apparent that overall the intervention group uses longer trajectories and a bigger vowel space and also shows more equal spacing at the post-test and the follow-up-test. Although `swing´ shows a slightly different picture, these results indicate that the intervention group pays more attention to the vowels in all dimension and shows

more consistency in their production (see Section 5.2.1.2). The approximant vowel trajectories do not reveal a lot of change in the in the approximant /w/ itself but in the production of the following vowel. This is a consequence of the sequences being analysed. The similar production of /w/ supports the outcome of the auditory analysis in which it was given high pronunciation ratings.

8.4.4 Plosives /b/, /d/ and /g/

In the German language final plosives are always devoiced (see Section 3.5.3). This thesis looked at the English voiced plosives /b/, /d/ and /g/ in final position. The pattern of the plosive durations mirrors the previous analyses showing an increase between pre- and post-test in the intervention group and a slight decrease between post-test and follow-up-test. The presented results reveal that the shortest mean pre-plosive vowel durations are shown for /b/ followed by /g/ and /d/. This might be due to the fact, that the test item `child´ entails a diphthong in contrast to a vowel. The data also reveals, that there is an overall increase in vowel duration for all of the three plosives between pre- and follow-up test. As the vowel duration in devoiced sequences is shorter than in similar voiced sequences, this lengthening indicates the emergence of increase vowel duration as a cue for voicing (see Section 3.5.3). So, although the intervention group shows a stronger effect, both groups show an increase in the performance of final devoicing. An ANOVA run on the duration data did not show a group effect. However, there is a significant main effect of plosive which supports the results of the descriptive statistics mentioned above. Moreover, the significant main effect of training reveals shorter duration for the trained items compared to the untrained items, which indicate higher articulatory speed and thus learning effects (see 5.2.1.2).

8.4.5 Fricatives /z/, /θ/ and /ð/

To analyse the fricatives, the measurements included mean vowel and fricative durations as well as the spectral features of the fricatives. The fricatives appear in the test items in pre- and post-vocalic position which affect the durational measures. Thus, the durations are analysed according to their location. For the duration of the fricatives in pre-vocalic position, the intervention group displays the same pattern as the analyses carried out before with an increase in fricative and vowel duration between the pre- and post-test and a decrease at the follow-up-test. In contrast, the durations of the control

group did not show a lot of change over the three data collection points (see Figure 8-23). As expected, the mean durations for vowels and fricatives in post-vocalic position are longer than in pre-vocalic position due to the voicing. Nevertheless, the post-vocalic durations behave similar to the pre-vocalic duration patterns with an increase between pre- and post-test and a decrease in the follow-up-test in the intervention group and with only slight changes over the three data collection points in the control group.

The mean fricative and vowel duration by fricative reveals that the duration values for /z/ are higher than for /θ/ and /ð/. This might be due to fact that the test items for /z/ include diphthongs in `goes´ (gəʊz) and `close` (kləʊs) and a long vowel in `jeans´ (dʒiːnz). The intervention group again displays a similar duration pattern for each of the fricatives with a rather strong increase between pre- and post-test. The biggest increase can be seen for /z/ indicating a strong impact of the pronunciation intervention, which focuses strongly on the voicing features, e.g. bumblebee /z/ (see Figure 8-23). This also corresponds to developmental patterns, in which the acquisition of the alveolar /z/ is starting very early whereas the dentals /θ/ and /ð/ are developing much later. Moreover, the auditory data has shown, that the subjects frequently devoiced the /ð/ which affects duration. Contrary to the results of of the intervention group, the durations for /z/ do not change a lot over the three data collection points (see Figure 8-23).

The length of the cavity in front of the constriction determines the spectral shape of the noise. As the front cavity becomes smaller from the production of the fricatives, the strongest resonances move upwards in frequency. Yet, the alveolar /z/ displays significantly higher centre of gravity (CoG) values than the two dental fricatives. This might be due to the fact that dental fricatives might be produced with little energy and show a diffuse spectrum. In addition, they are rather prone to the development of antiformants. The voiceless /θ/ shows a higher centre of gravity than the voiced /ð/. This was expected as voiced fricatives show weaker intensities as the vocal folds are held closer and limit the degree of airflow (see Section 5.4.3). For the three fricatives /z/, /θ/ and /ð/ the CoG values of the intervention group increase over the three data collection points which might indicate that the subjects are actually producing /θ/ and /ð/ instead of replacing it with an /s/ or /z/ which is typical for a German accent (see 3.5.3). In contrast, the centre of gravities decreases at the post-test and then increase again approaching the pre-test values for /z/ and /ð/ and surpass the pre-test for the voiceless /θ/ and /z/.

8.4.6 Affricate /dʒ/

As affricates can be seen as an additional manner of fricative articulation (Pickett 1999, p. 121), the analysis was treated similar to the fricatives. Equally to the durations presented in the previous sections of this chapter, the intervention group displays an increase in affricate and vowel durations between the pre-test and post-test with the exception of the affricate duration of the test-item ‘cage’. As the vowel duration for `cage´ had increases at the same time, it seems that the lengthening of the diphthong /eɪ/, to which special attention was paid in the intervention programme (see Figure 8-25), might have led to shorter affricate durations indicating devoicing. In contrast to the intervention data, the durations of the control group always show a slight decrease for the vowel durations between pre- and post-test and then an increase at the follow-up-test whereas the affricate duration do not display a lot of change over the three data collection points. The centre of gravity was used as an additional cue for the affricate analysis. However, the CoG results did not provide a lot of information as the phonetic context of the test itmes influenced the centre of gravity with the lowest values produced for `John´, then for `job´, that has a similar phonetic context, and finally for `jeans´ and `cage´ in that order.

So far, this chapter 8 has presented the results and outcomes of the auditory and acoustic analyses of the pronunciation intervention. The following chapter 9 of this thesis moves on to draw conclusions from the presented material

9. Conclusions

It was the purpose of this doctoral study to develop a new pronunciation intervention programme grounded in theories of second language acquisition, which can be easily integrated into English pronunciation teaching in German schools. Secondly, the study set out to determine whether and to what extent the pronunciation intervention can improve the pronunciation skills of young second language learners. The results of the data analyses showed a mixed picture of the effectiveness of the pronunciation training programme. This chapter's concluding analysis of the outcomes is structured by reference to the research focus and hypotheses (see Chapter 6).

Given that all the young learners who participated in this study were undergoing normal maturation processes and receiving ongoing English language input, the performance of both the intervention and control groups should have increased over time (Hypothesis 1). The study set out to determine whether and to what extent the pronunciation intervention could improve second language pronunciation over and above those maturational advances. Thus, the group receiving the pronunciation intervention should have performed better on the post-test and follow-up test of their L2 pronunciation abilities compared to the control group (Hypothesis 2). It is well established in intervention studies that the best performance scores are usually obtained at the time directly after the treatment (Shadish et al., 2001), so in this investigation it was predicted that the intervention group's post-test performance would be superior to the follow-up test five months later (Hypothesis 3). In this thesis, pronunciation performance was determined by mean accuracy scores and acoustic measures.

The auditory analysis revealed that the mean accuracy ratings for both groups increased over the three data collection points (see Table 8-1) and this trend is supported by the main effect of time showing that for both groups pronunciation ratings increased significantly between pre-test and post-test and increased further at the follow-up-test (see Section 8.1.2). These findings support Hypothesis 1 that the performance of both groups increased through maturation and ongoing English language input. However, the interaction between time x sound was not significant and scrutiny of the descriptive data revealed that the increase in the rating scores between pre- and post-test could not be clearly attributed to the intervention because the largest increase was seen in the control group (see Figure 8-3). However, despite the fact that

both groups had fairly similar ratings at the post-test, the intervention group surpassed the control group at the follow-up test and this difference was significant.

The acoustic analysis also showed mixed results. Learning effects should have led to increased articulatory speed and more consistent pronunciation. Thus, maturation would have resulted in decreasing durations and smaller standard deviations. Interestingly, the control and intervention groups display completely different patterns in the acoustic analysis with regard to duration scores. In the control group, for most data[58], there is a slight decrease in vowel duration at the post-test indicating that maturation effects are taking place and that the traditional teaching seems to have positive impacts on pronunciation learning. However, the control group data predominantly do not show carry-over effects for the follow-up test, and this finding does not support ongoing maturation effects. In the intervention group, the acoustic analysis showed increases in duration at post-test with the exception of the affricate duration for the test item 'cage'. This pattern was very consistent over all acoustic duration analyses. It seems that the duration patterns found in the acoustic analysis were overly influenced by hyperarticulation due to the test scenario or Hawthorne effects. In addition, time did not show up as a main effect in the ANOVAs run on the vowel, diphthong and plosive durations. Therefore, the duration data remain inconclusive despite the fact that the groups are obviously behaving very differently. Thus, the acoustic duration scores must be regarded with caution; other analyses and cues might reveal more illuminating information.

The F1 x F2-F1 planes for the vowels, diphthongs and approximants show clear patterns of results. With regards to the vowels, for both groups the vowel ellipses are getting smaller and show less scatter indicating more precise production of the vowels (see Figure 8-11). In addition, the vowel spaces are getting bigger (see Figure 8-11) and, apart from the neutral vowel /ə/, the Euclidean distances increased between pre- and post-test for both groups (see Figure 8-12). The overlap in the vowel spaces clearly diminished in the intervention group showing more awareness of the English vowel spaces and thus more precision in the production of the vowels. In contrast, the control group showed more overlap. In addition, the visual inspection of the diphthong scatter plots for /ɪə/ and /eɪ/ over the three data collection points revealed decreasing dispersion, indicating increasing stability in the pronunciation of the diphthong and

[58] Exceptions are: /æ/, /eɪ/, /w/ in `witch´, fricative and affricate durations (although the vowel duration for the affricates and fricatives show a decrease between pre-and post-test).

maturation effects for both groups (see Figure 8-13 and Figure 8-15). However, the intervention group clearly shows less dispersion at the pre-test, and this was carried through to the follow-up-test indicating better performance than the control group. Similarly, the intervention group's performance of the approximant /w/ displayed longer trajectories and a bigger vowel space and even exhibited more equal spacing at the post-test and the follow-up test compared to the control group.

A number of studies (Best, 1995; Flege et al., 2003; Iverson et al., 2003; Kuhl, 2000) have shown that different L2 sounds pose different degrees of difficulty for language learners. Therefore, this investigation included sounds that are different in the English and German sound inventories, as well as sounds that are similar in the two languages (see Section 3.5) The auditory analysis showed a strong significant effect of sound indicating that different sounds posed different problems for the German learners of English (Hypothesis 4). The mean sound ratings revealed that productions of the vowel /ə/, the diphthong /eɪ/ and the approximant /w/ yielded the three highest scores (see Figure 8-2). The sounds /ə/ and /w/ are similar to their German counterparts, whereas there is no German diphthong akin to /eɪ/. Conversely, the vowel /ɔ:/ and the fricatives /θ/, which have no equivalents in the German sound inventory, and the plosives /b/ and /z/ which are similar to their German counterparts, received the three lowest mean ratings. The interaction between sound x group (see Section 8.1.6) revealed that the fricatives /ð/ and /θ/ and the affricate /dʒ/, which are not part of the German sound inventory, improved significantly after the intervention indicating that the training programme might have had positive effects particularly on "new" sounds. Addressing Hypothesis 4, the results of the study did not provide enough evidence to answer the question of whether similar or different sounds are easier to acquire. However, this study did not employ minimal pairs, but instead used authentic vocabulary for the test items. Thus, the different coarticulatory contexts might have masked sound effects.

To explore the prediction that the pronunciation performance of the treatment group would improve not only as a result of the training effects of the intervention but also due to the intervention itself, test items that were explicitly trained in the treatment were contrasted with untrained test items (Hypothesis 5). The main effect of training revealed that the trained test items had significantly higher mean pronunciation ratings compared to the untrained items (see Section 8.1.3). However, there was no significant interaction between training x time. Inspection of the descriptive data (Table 8.2)

revealed that the means for both groups increased over the three data collection points indicating that maturation effects were taking place. As the control group did not receive a special treatment for the trained test items, ideally the ratings for the trained and untrained test items should have been fairly similar and indeed this assumption is true for all sounds with the exception of /b/, ð/, /θ/ and /z/ (see Figure 8.6). The interaction between training x sound x group (Figure 8.5) revealed that for nine of the 14 sounds higher ratings were given for the trained items compared to the untrained items, although significant training effects could be found for /b/ and /d/. However, training showed a significantly negative impact for /z/ and /w/. Review of the test items containing these sounds suggested that either the coarticulatory context or the within-task sorting (trained and untrained status) might have been confounding factors (see discussion below of the interaction of sound x frequency x training). The ANOVAs run on the durations for the vowels and diphthongs revealed a significant main effect of training, with longer durations for the trained test items compared to the untrained test items. This result is similar to the duration pattern over the time intervals, where the biggest increase was found from pre- to post-test suggesting that the students might have been making efforts to pronounce the trained items extra carefully. Overall, the results indicate that training effects are taking place.

Frequency effects are known to influence speakers' performances (Bishop & Keating, 2012; Cholin, 2008); more frequent use tends to lead to increased skill. Thus, higher frequency items should have outscored the lower frequency items (Hypothesis 6). The simple effects analysis (see Section 8.1.9) of sound x frequency revealed that the test items containing the vowels /ɔ:/ and /æ/, the plosive /b/ in final position, the fricative /ð/ and the approximant /w/ received significantly higher ratings in line with higher frequency status. Moreover, the interaction graph (see Figure 8-8) shows that the higher frequency items (with the exception of the fricatives /θ/ and /z/) generally got slightly higher ratings compared to the lower frequency items, showing a strong effect of frequency. In addition, the acoustic analysis run on the duration data revealed a statistically significant effect of frequency, with longer durations for the higher frequency items containing the vowels and the approximant, and shorter durations for the diphthongs (see Section 8.3).

As the interactions between sound x frequency (see Figure 8-7) and sound x frequency x training (see Figures 8-8 and 8-9) revealed, frequency effects sometimes surpassed training effects and it is conjectured that the set-up of the stimuli according

to frequency and training might have had additional effects (see discussion at Section 8.2).

Comparing both data sets, a reason for the partly different results of the two analyses of the same data might be that the auditory analyses were more strongly affected by the inherent features of the test items compared to the acoustic analysis that despite coarticulation effects focused more on the specific sound.

9.1 Key findings, limitations and outlook

In the literature on pronunciation, most studies have focused only on single aspects of pronunciation and were set in laboratories with mostly adult participants (Archibald & Young-Scholten, 2003; Bohn & Flege, 1992; Gut, 2009 Iverson et al., 2003; Neri et al., 2006; Strange et al., 1998) and there are only few pronunciation studies that present data for classroom- based interventions with younger learners in primary and secondary school education (Thomson and Derwing (2015)). This thesis is one of them. According to Thomson and Derwing's (2014) benchmark with regards to a "good" pronunciation training study this thesis fulfils almost all of the criteria (see Section 3.1). It has a very high ecological validity as the pronunciation intervention can be easily integrated in the general communicative L2 classroom of German English L2 learners at the ages of 10 to 12. Not only were the curricular guidelines considered, but also great care was taken to implement the intervention in the everyday communicative foreign language classroom. For this reason, the results of a teacher questionnaire (see Section 2.4) and a textbook/material analysis (see Section see 2.3.) fed into the set-up of the pronunciation training programme (see Chapter 4). The analysis revealed, that these materials do not provide a theoretical background on pronunciation acquisition and there are barely any guidelines on methods to implement and systemise pronunciation teaching in the foreign language classroom. Thus, this thesis provides a rationale for pronunciation teaching grounded in theories of language perception and production (e.g. Best (1995), Flege (1995)), the language magnet (Kuhl and Iverson (1995)) and takes the outcomes of motor theory and mirror neurons into account. In addition, the thesis includes a detailed teachers' manual that can be used by foreign language teachers (see Appendix Section 10.4). In addition, methods and participants of the data analyses are described in detail (see Section 7.5) to allow a replication of the study. To verify that any improvement in pronunciation abilities is a result of

instruction, a control group was employed and a follow-up-test was included to determine lasting effects of the intervention.

The main point of concern regarding this study relates to the selection of the test items. Due to the limited vocabulary of English learners at beginner level, the test items could not be matched as minimal pairs. The use of authentic vocabulary as test items meant that the sounds appeared in different coarticulatory contexts which affected the outcomes of the analyses (see Section 8.4 and Chapter 9 above). In addition, the results also showed clear frequency effects (see Chapter 8 and 9) which seemed to surpass training effects and might have affected the overall results. Thus, to avoid additional effects, in a further study, test items should be chosen according to similar frequency. Moreover, the choice of the data elicitation method (see Section 7.4.3) might have affected the results, as reading problems could have contributed to poor pronunciation ratings. Also, some of the results – although not statistically significant – indicated biases in the matching of the intervention and control group, which was done semi-randomly. Due to the constraints of the PhD project, only 16 monolingual German learners of English were chosen for the final analysis in this study (see Section 7.2.4). Despite the investigation being comparable in scale to other acoustic studies of pronunciation that have an average sample size of 11.61 (Sakai & Moorman, 2018), the small sample size (n = 16) partly limited the validity of the results. A further study could assess the robustness of the presented findings, especially as originally datasets were completed for 181 students. In addition, for all of these students quantitate and qualitative questionnaire data on the participants' sex, age, language background, language experience and motivation were gathered which was not used in this study (see Section 10.3.3.in the Appendix for full questionnaire). So complementary analyses could be conducted to provide more insight in social factors affecting learning and the nature of interactions in the L2 and to increase the power of the statistical analyses.

10. Appendix

10.1 Literature

Table 10-1: Pronunciation training programmes carried out with younger learners of English

Suggested Literature	Cardoso 2010	Chen & Goswami, 2011	Lima, 2010	Kennedy, 2003	Trofimovich, Light-bown, Halter, & Song, 2009	Tsiartsioni, 2010	von Rekowski
subject language	Portuguese monolingual (BP)	Spanish (Mexico)	Portuguese	Francophone	Francophone	Greek	German, monolingual
L2	Slavir (fake language)	English	English	English	English	English	English
n	30	44	28	47	74	72	16
age	16.3	15-19	11-13	17-20	8.4	10-16	10.95
place of instruction	secondary school classroom	high school, intact classes	non-native teacher, two intact classes	junior college intact classes	20 intact classes	state school intact classes,	secondary school intact classes
design	quasi-experimental, within groups pretest/ posttest design.	quasi-experimental, within groups pre-post test two treatment	interventionist action research intervention study, pre- post-test, delayed post-test	Intervention study control group, pretest, posttest	Intervention study control group design,	control group design	intervention study quasi-experimental control group design pretest, posttest, delayed posttest
training scope	foreign homorganic onset clusters (/sl/, /sn/, /st/ - sC)	/t, d, v, z. ð,θ š/	Sounds particularly difficult for Brazilan speakers of English: /θ/,/ð/, /æ/, /ɬ/,/uː/, final /t, d, ɪd/,/iː/, /h/, /t/, /iː/, /h/, /r/, /ɪ/,/ə/, /ə/, /ɽ/	/θ/,/ð/	6 simple English sentences, particular interest /h/, possessive -s	Vocalic and intervocalic duration production of aspirated /p, t, k/, the devoicing of /b, d, g/ and vowel, lengthening before word final voiced stops.	Vowels: /ɑː/, /ɔː//æ/, /ə/ Diphthongs: /eɪ/, /ɪə/ Word-final plosives /b/, /d/, /g/ Fricatives: /θ/, /ð/, /z/ Affricate: /dʒ/ Approximant: /w/
theoretical paradigm	Teachability ´, Markedness	Cooperative Learning	Explicit teaching	peer corrective feedback	sustained, long-term comprehension practice in both listening and reading	/	Motor Theory, Language Magnet, Speech Learning Model
Analysis	auditory and acoustic ratings	audio and video tape, soundswere as correct (+) or incorrect.	recordings were phonetically transcribed and analyzed.	rating scale according to tongue position, accuracy scores	sentences were transcribed, judged by phonetically trained judge, listener ratings	Duration (VOT, Vowel length before final stops) Acoustic ratings	auditory and acoustic ratings

Table 10-2: Pronunciation training programmes focussing on segments

Suggested Literature	Elliott, 1995	Garcia, 2005	Huthaily, 2008	Liu & Fu, 2011	Warsi, 2002	von Rekowski
subject language	English	English	English	Mandarin	Japanesse	German, monolingual
L2	Spanish	Spanish	Arabic	English	English	English
n	66	53	46	60	16-18	16
age	/	18-32		juniors at a University	adult learners	10.95
place of instruction	University students	University students	University students	University students	State College	secondary school intact classes
design	pre- and post-test	pre- and post-test 3 treatment groups + control group	control group, pre- and post-test	intact classes control group	control group, pre- and post-test	intervention study quasi-experimental control group design pre-test, post-test, delayed post-test
training scope	mimicking discrete words, sentences, pronunciation of written words, spontaneous pronunciation focus on allophones	/l/, /ɬ/ /d/, /t/, /ɾ/	sound recognition, minimal pairs Arabic sounds, 40 items	10 sounds that were often mispronounced including /v/ and /w/.	/l/, /r/	Vowels: /ɑ:/, /ɔ://æ/, /ə/ Diphthongs: /eɪ/, /ɪə/ Consonants: Word-final plosives /b/, /d/, /g/ Fricatives: /θ/, /ð/, /z/ Affricate: /dʒ/ Approximant: /w/
theoretical paradigm	pronunciation attitude, field independence		Contrastive Analysis	Monitor and effects of instruction	Effects of visual feedback (diagram with articulatory movements)	Motor Theory, Language Magnet, Speech Learning Model (SLM)
Analysis	3 raters auditory accuracy ratings	pronunciation accuracy 3-point scale rated by a native speaker of English	minimal pair ratings	auditory ratings	list of 5 words. Auditory analysis (rating scale)	auditory and acoustic ratings

Table 10-3: Pronunciation training programmes involving awareness raising

Literature	Couper, 2011	Ingels, 2011	Nagamine, 2011	Sardegna, 2011	von Rekowski
subject language	Korean, Mandarin	Mandarin	Japanese	Chinese, Vietnamese, Korean, Thai, Turkish, French, Portuguese, Spanish	German, monolingual
L2	English	English	English	English	English
n	24	7	30	38	16
age	adults	23-28	adults	22-47	10.95
place of instruction	University students	University students	University students	University students	secondary school intact classes
design	pre- and post-test 2 ×2 factorial design	pre- and post-test	pre- and post-test	Pre-test, post-test, delayed post-test	intervention study quasi-experimental control group design pre-test, post-test, delayed post-test
training scope	epenthesis, or the inappropriate addition of an extra vowel, often a schwa.	suprasegmental features	voiceless bilabial, alveolar, and velar stops/plosives	linking sounds within and across words	Vowels: /ɑː/, /ɔː/, /æ/, /ə/ Diphthongs: /eɪ/, /ɪə/ Consonants: Word-final plosives /b/, /d/, /g/ Fricatives: /θ/, /ð/, /z/ Affricate: /dʒ/ Approximant: /w/
theoretical paradigm	CP, socially constructed metalanguage (SCM) and critical listening (CL)	3 levels of self-monitoring	Hyper-Pronunciation Training Method	Covert Rehearsal Model (CRM)	Motor Theory, Language Magnet, Speech Learning Model (SLM)
analysis	intelligibility ratings	accuracy scores on message unit boundaries, primary phrase stress, and intonation	VOT, pitch	pronunciation accuracy scores	auditory and acoustic ratings

Table 10-4: Pronunciation training programmes including auditory rating scales

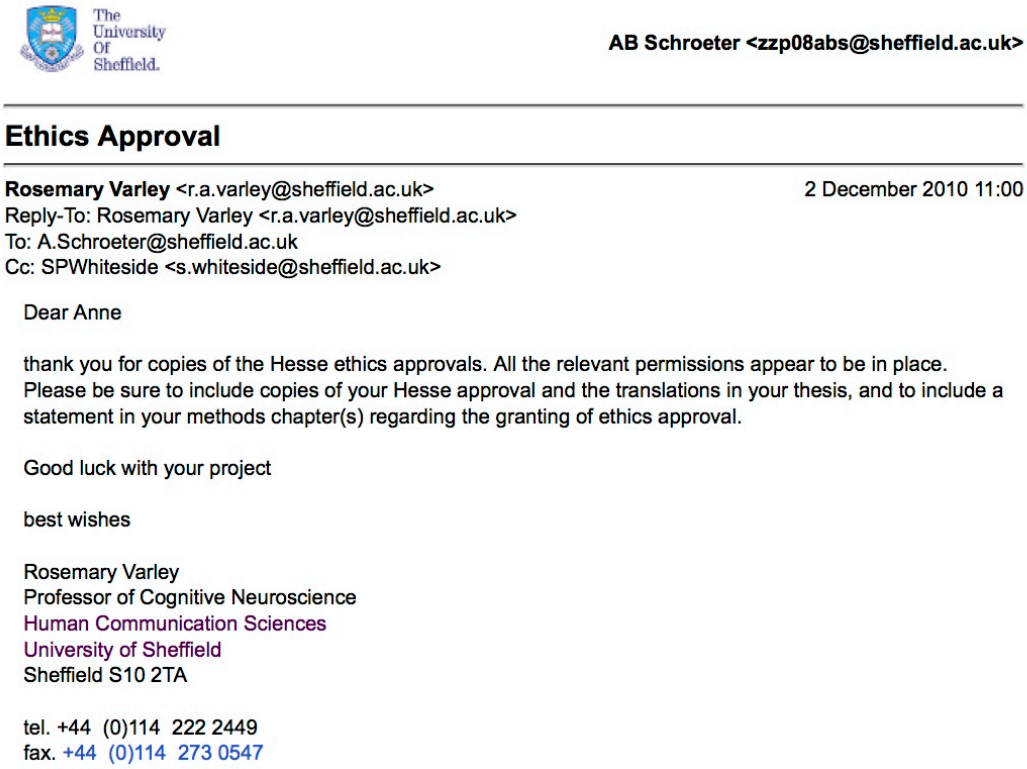
Literature	Henderson, 2008	Munro & Derwing, 2008	Perlmutter, 1989	von Rekowski
subject language	Bulgarian, French, Japanese, Greek	Mandarin, Russian, Ukrainian, and Croatian	China, Korea, Sri Lanka, Malaysia, Austria, Germany, Poland, Ghana	German, monolingual
L2	English	English	English	English
n	5	44	24	16
age	23-58	19-49	19-43	10.95
place of instruction	University Course	University Course	University students/ teaching assistants	secondary school intact classes
design	intervention study pre- and post-test	Repeated measures 2 experimental groups (Asian vs. Slavic language background)	pre- and post-test	intervention study quasi-experimental control group design pretest, posttest, delayed posttest
training scope	spontaneous speech	10 English vowels in CVC context word frequency	intelligibility ratings, subject matter identification	Vowels: /ɑː/, /ɔː//æ/, /ə/ Diphthongs: /eɪ/, /ɪə/ Consonants: Word-final plosives /b/, /d/, /g/ Fricatives: /θ/, /ð/, /z/ Affricate: /dʒ/ Approximant: /w/
theoretical paradigm	Contrastive Analysis, Functional Load			Speech Motor Control, Language Magnet, Speech Learning Model (SLM)
analysis	speech rate, pace, and word stress	Intelligibility Judgment Task word frequency	intelligibility ratings, conversational topic identification	auditory and acoustic ratings

Table 10-5: Pronunciation training programmes including acoustic analyses

Literature	Counselman, 2010	Hincks & Edlund, 2009	Suarez, 2013	von Rekowski
subject language	English	Mandarin (and other Chinese dialects)	Spanish, French	German, monolingual
L2	Spanish	English	English	English
n	28	14	8	16
age		22.3 (test group), 24.5 (intervention group)		10.95
place of instruction	University conversation course	English language class at University		secondary school intact classes
design	control group design	control group design	pre- and post-test	intervention study quasi-experimental control group design pretest, posttest, delayed posttest
training scope	/e, o, a, p, t, k, b, d, g, ɾ, l/	intonation, on-line visual feedback on the presence and quantity of pitch variation	/p/, /t/, /k/	Vowels: /ɑː/, /ɔː//æ/, /ə/ Diphthongs: /eɪ/, /ɪə/ Consonants: Word-final plosives /b/, /d/, /g/ Fricatives: /θ/, /ð/, /z/ Affricate: /dʒ/ Approximant: /w/
theoretical paradigm	production vs perception		Markedness	Motor Theory, Native language Magnet, Speech Learning Model (SLM)
analysis	F2/F1 analysis for /o/ and /e/ auditory ratings for the consonants	pitch, duration	Duration (VOT)	auditory and acoustic ratings

10.2 Ethical clearance

As the study was conducted in schools in Hesse in Germany, ethics approval was sought from the Hessian Ministry of Education in Germany (see chapters 10.2.1 and 10.2.3). The Chair of the ethics review procedures at the time of the approval stated that the Department of Human Communication Sciences in Sheffield recognises the German ethics approval (see Figure 10-1):



The University Of Sheffield.

AB Schroeter <zzp08abs@sheffield.ac.uk>

Ethics Approval

Rosemary Varley <r.a.varley@sheffield.ac.uk> 2 December 2010 11:00
Reply-To: Rosemary Varley <r.a.varley@sheffield.ac.uk>
To: A.Schroeter@sheffield.ac.uk
Cc: SPWhiteside <s.whiteside@sheffield.ac.uk>

Dear Anne

thank you for copies of the Hesse ethics approvals. All the relevant permissions appear to be in place. Please be sure to include copies of your Hesse approval and the translations in your thesis, and to include a statement in your methods chapter(s) regarding the granting of ethics approval.

Good luck with your project

best wishes

Rosemary Varley
Professor of Cognitive Neuroscience
Human Communication Sciences
University of Sheffield
Sheffield S10 2TA

tel. +44 (0)114 222 2449
fax. +44 (0)114 273 0547

Figure 10-1: Ethics approval from the Department of Human Communication Sciences, Sheffield, UK

10.2.1 Ethics approval: Hessian Ministry of Education: Teacher Questionnaire

On September 10th, 2009, the Hessian Ministry of Education granted the approval of the questionnaire study. The English translation can be found in the following section 10.2.2.

Bearbeitungsnummer 660.003.000 - 304 -

Sehr geehrte Frau Schröter,

wir hatten in der vergangenen Woche miteinander telefoniert. Sie hatten mir mitgeteilt, dass Sie die Untersuchung an 500 Schulen durchführen möchten und mit Herrn Ilnitzky im Vorfeld bereits abgesprochen hatten, dass die Einverständniserklärungen der Schulen in Ihrem Falle nicht notwendig sind.
Die Stellungnahmen der Referatsleiterin für Grundschulen sowie die der Datenschutzbeauftragten liegen inzwischen vor, sie sind positiv.
Ich genehmige damit Ihren Antrag vom 26. August 2009, Sie können mit der Untersuchung beginnen.

Mit freundlichen Grüßen

Monika Hestermann-Retterath

HESSEN

Monika Hestermann-Retterath
Hessisches Kultusministerium
Referat I.4 (M.2)
Luisenplatz 10
65185 Wiesbaden
Tel.: +49 611 368 - 2223
Fax: +49 611 97100600
E-Mail: Monika.Hestermann-Retterath@hkm.hessen.de

Figure 10-2: Ethics approval of the questionnaire study by the Hessian ministry of education

10.2.2 Translation of the ethics approval of the questionnaire study

Reference number 660.003.000 -304-.

Dear Ms. Schröter,

We've talked to each other on the phone last week. As you've told me, you're planning to do your research in about 500 schools and that you've already agreed with Mr. Ilnitzky prior to our contact that the schools do not need to provide any consent forms.

The statements of the head of the primary school division and the data protection commissioner are positive with regards to your request. Therefore, I grant permission to your research application from August 26th, 2009. You may start with your research.

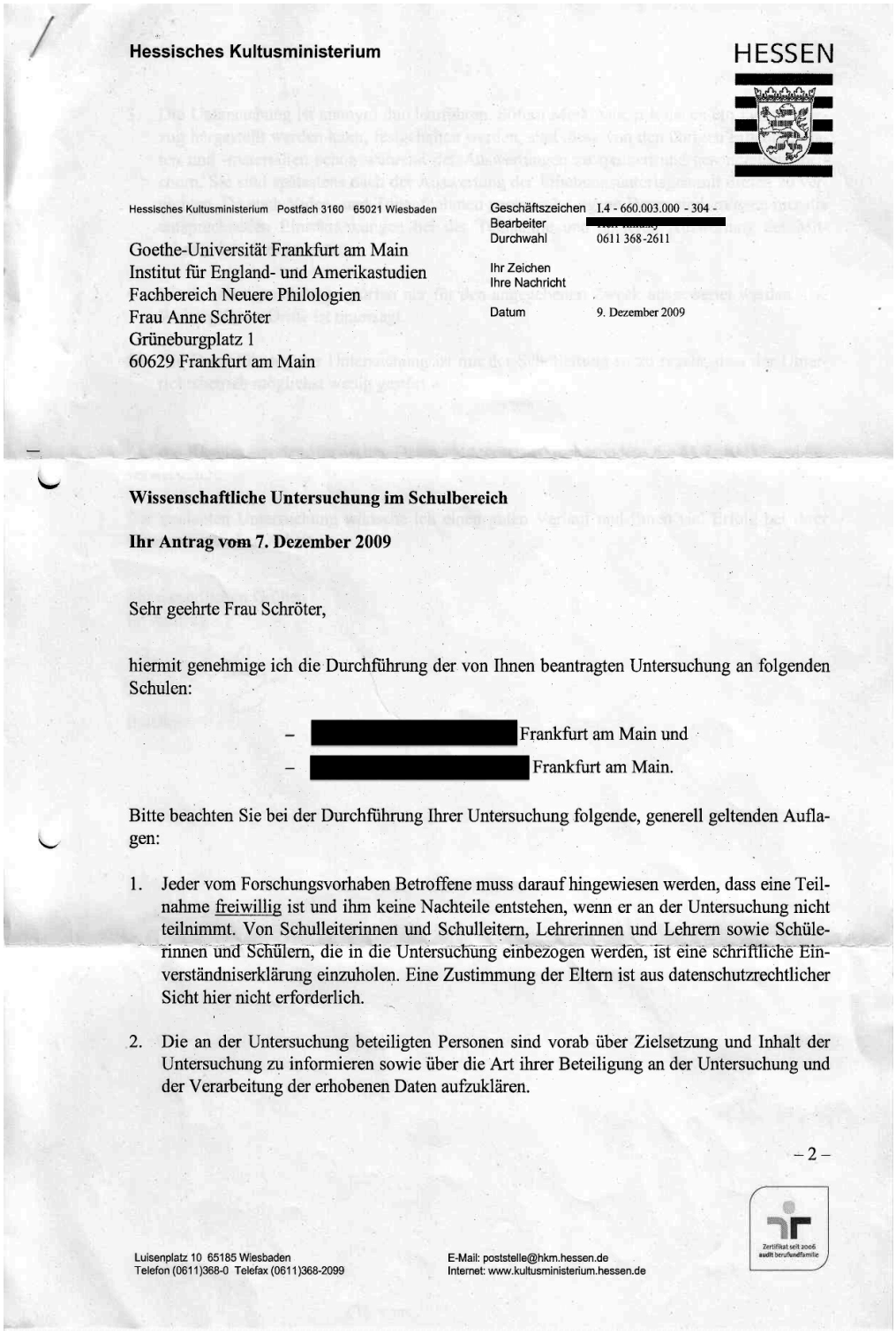
Yours sincerely

XY

Hessian Ministry of Education

10.2.3 Ethics approval: Hessian Ministry of Education: Intervention Study

On December 9th, 2009, the Hessian Ministry of Education granted the approval of the pronunciation intervention project. The English translation can be found in the following section 10.2.4.



Hessisches Kultusministerium

HESSEN

Hessisches Kultusministerium Postfach 3160 65021 Wiesbaden

Goethe-Universität Frankfurt am Main
Institut für England- und Amerikastudien
Fachbereich Neuere Philologien
Frau Anne Schröter
Grüneburgplatz 1
60629 Frankfurt am Main

Geschäftszeichen I.4 - 660.003.000 - 304 -
Bearbeiter [redacted]
Durchwahl 0611 368-2611

Ihr Zeichen
Ihre Nachricht

Datum 9. Dezember 2009

Wissenschaftliche Untersuchung im Schulbereich

Ihr Antrag vom 7. Dezember 2009

Sehr geehrte Frau Schröter,

hiermit genehmige ich die Durchführung der von Ihnen beantragten Untersuchung an folgenden Schulen:

- [redacted] Frankfurt am Main und
- [redacted] Frankfurt am Main.

Bitte beachten Sie bei der Durchführung Ihrer Untersuchung folgende, generell geltenden Auflagen:

1. Jeder vom Forschungsvorhaben Betroffene muss darauf hingewiesen werden, dass eine Teilnahme freiwillig ist und ihm keine Nachteile entstehen, wenn er an der Untersuchung nicht teilnimmt. Von Schulleiterinnen und Schulleitern, Lehrerinnen und Lehrern sowie Schülerinnen und Schülern, die in die Untersuchung einbezogen werden, ist eine schriftliche Einverständniserklärung einzuholen. Eine Zustimmung der Eltern ist aus datenschutzrechtlicher Sicht hier nicht erforderlich.

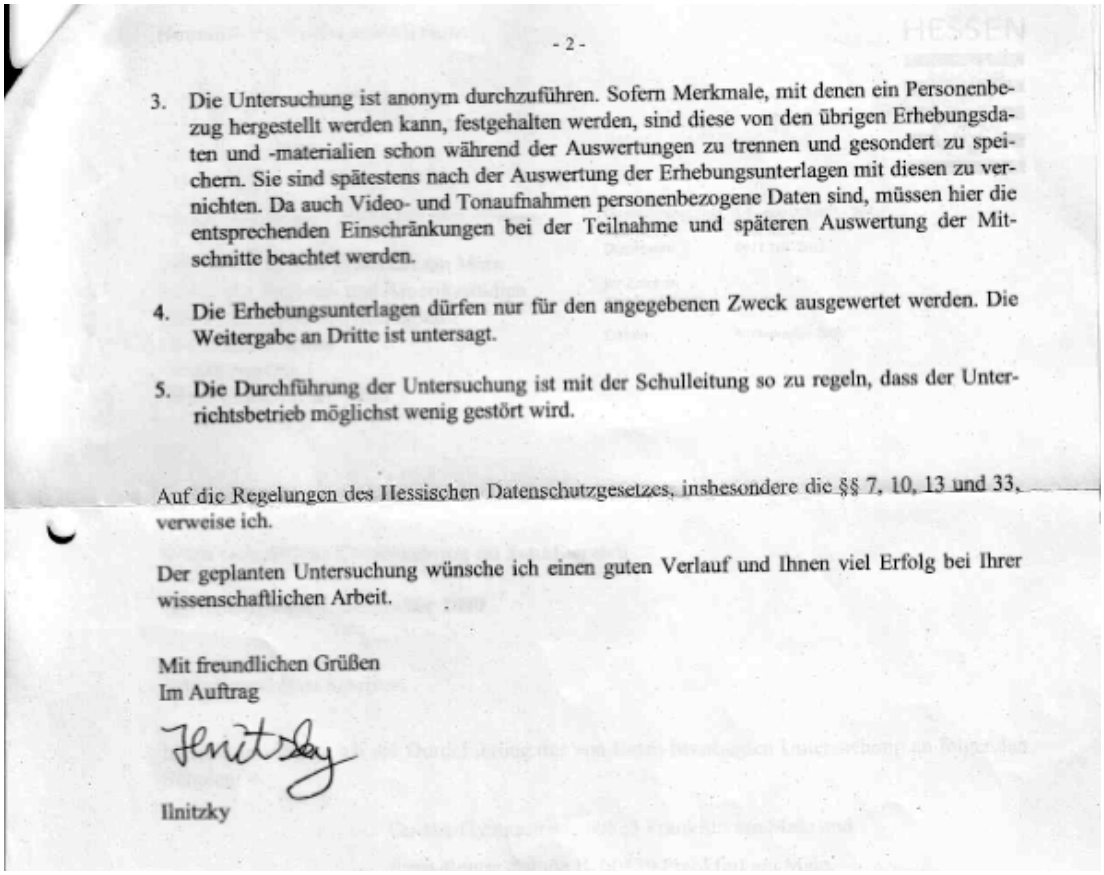
2. Die an der Untersuchung beteiligten Personen sind vorab über Zielsetzung und Inhalt der Untersuchung zu informieren sowie über die Art ihrer Beteiligung an der Untersuchung und der Verarbeitung der erhobenen Daten aufzuklären.

– 2 –

Zertifikat seit 2006 audit berufundfamilie

Luisenplatz 10 65185 Wiesbaden
Telefon (0611)368-0 Telefax (0611)368-2099

E-Mail: poststelle@hkm.hessen.de
Internet: www.kultusministerium.hessen.de



- 2 -

3. Die Untersuchung ist anonym durchzuführen. Sofern Merkmale, mit denen ein Personenbezug hergestellt werden kann, festgehalten werden, sind diese von den übrigen Erhebungsdaten und -materialien schon während der Auswertungen zu trennen und gesondert zu speichern. Sie sind spätestens nach der Auswertung der Erhebungsunterlagen mit diesen zu vernichten. Da auch Video- und Tonaufnahmen personenbezogene Daten sind, müssen hier die entsprechenden Einschränkungen bei der Teilnahme und späteren Auswertung der Mitschnitte beachtet werden.

4. Die Erhebungsunterlagen dürfen nur für den angegebenen Zweck ausgewertet werden. Die Weitergabe an Dritte ist untersagt.

5. Die Durchführung der Untersuchung ist mit der Schulleitung so zu regeln, dass der Unterrichtsbetrieb möglichst wenig gestört wird.

Auf die Regelungen des Hessischen Datenschutzgesetzes, insbesondere die §§ 7, 10, 13 und 33, verweise ich.

Der geplanten Untersuchung wünsche ich einen guten Verlauf und Ihnen viel Erfolg bei Ihrer wissenschaftlichen Arbeit.

Mit freundlichen Grüßen
Im Auftrag

Ilnitzky

Figure 10-3: Ethics approval of the intervention study by the Hessian Ministry of Eduation

10.2.4 Translation of the ethics approval by the Hessian Ministry of Education:

Scientific School Research
Your application from December 7th, 2009

Dear Ms. Schroeter,

I grant approval for scientific research in the following requested schools:

1. School A, Frankfurt am Main
2. School B, Frankfurt am Main

Please adhere to the following standard practice conditions in your research:

1. Every subject concerned by the study has to be informed that the participation is voluntary and that there are no consequences from not taking part. A written consent form is needed from the head of schools, teachers and students who are taking part in the research. According with data protection regulations a parent consent form is not needed in this study.
2. Every subject concerned by the study has to be informed about the aims and contents of the research. Moreover, the nature of the participation and data analysis has to be made clear prior to the study.

3. The scientific research has to be done anonymously. In case biographical characteristics are collected, they have to be separated from other collected data and material already during the data analysis and they have to be saved in different places. At the latest they are to be deleted after the data analysis. As video and audio are counted as biographical data, specific restrictions of the participation in and analysis of the recordings have to be taken into account.
4. The data collected is only to be used for the requested aims. The transmission of data to third parties is not allowed.
5. The research done in school has to be conducted in such a way that there is a minimal interference with the regular school education.

I refer to the Hessian Data Protection Act, in particular to §7, 10, 13, and 33.

I wish you good luck for your planned study and success for your research.

Yours sincerely

10.2.5 Parent Consent Forms

10.2.5.1 Parent consent form: control group

Anne Schröter
Goethe-Universität Frankfurt
Institut für England- und Amerikastudien
Abteilung Sprachlehrforschung und Didaktik
Grüneburgplatz 1
60629 Frankfurt am Main

Telefon +49 (0) 69 26024492
E-Mail schroeter@em.uni-frankfurt.de

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Frankfurt, 11.1.2010

Informationsschreiben zu einer wissenschaftlichen Untersuchung zur Aussprache-förderung im Rahmen des Englischunterrichts am XXX.

Liebe Eltern!

Derzeit wird am Institut für Anglistik, Abteilung Sprachlehrforschung und Didaktik der Goethe-Universität Frankfurt unter der Leitung von Herrn Professor Quetz ein wissenschaftliches Forschungsprojekt zur Aussprache im Englischunterricht der Unterstufe durchgeführt.

Um einen Einblick zu erhalten, welchen Standard die Aussprache von Schülern in der fünften Klasse hat, werden im Februar 2010, Juli 2010 und November 2010 Datenerhebungen an verschiedenen Schulen im Raum Frankfurt durchgeführt. Die Dauer der Untersuchung beträgt pro Kind jeweils ca. fünf bis zehn Minuten. Bei dieser Untersuchung werden ca. 50 aus dem Englischunterricht bekannte Wörter auf einem Computerbildschirm gezeigt. Die Schüler/innen werden gebeten, diese Wörter laut vorzulesen und in ein Mikrophon zu sprechen. Zudem werden die Kinder gebeten, verschiedene Ausspracheübungen, die sie im Unterricht kennen gelernt haben, zu bewerten.

Alle Aufnahmen, Angaben und Daten werden anonymisiert und vertraulich behandelt, nicht an Dritte weitergegeben, und die Auswertung der Erhebungsunterlagen erfolgt nur im Rahmen der angegebenen Studie.

Die wissenschaftliche Untersuchung wurde gemäß des Erlasses vom 9.12.2009 durch das Hessische Kultusministerium genehmigt. Das XXX hat dem Forschungs-projekt in der Schulkonferenz vom 5.5.2009 zugestimmt.

Wir bitten Sie, in der beigefügten Erklärung Ihr Einverständnis zu geben, dass Ihre Tochter/Ihr Sohn an der wissenschaftlichen Studie teilnehmen darf.

Gerne stehe ich Ihnen für Fragen auch gerne per Email oder Telefon zur Verfügung.

Anne Schröter (Untersuchungsleiterin)

Einverständniserklärung zur Teilnahme an der wissenschaftlichen Untersuchung zur Ausspracheförderung im Englischunterricht.

Ich bin damit einverstanden, dass

meine Tochter _____/ mein Sohn_____

an dem Forschungsprojekt zur englischen Aussprache teilnimmt.

- Die Beteiligung an der Studie ist freiwillig. Eine Nichtteilnahme an der Studie zieht keinerlei Konsequenzen nach sich.

Wenn zutreffend, bitte ankreuzen:

- ▫ Mein Kind hat eine Hörschwäche.

Name Schüler/in:

Name Erziehungsberechtigte/r:

Kontakt:

Figure 10-4: Parent consent form: control group

10.2.5.2 Parent consent form: intervention group

Anne Schröter
Goethe-Universität Frankfurt
Institut für England- und Amerikastudien
Abteilung Sprachlehrforschung und Didaktik
Grüneburgplatz 1
60629 Frankfurt am Main

Telefon +49 (0) 69 26024492
E-Mail schroeter@em.uni-frankfurt.de

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Frankfurt, 11.1.2010

Informationsschreiben zu einer wissenschaftlichen Untersuchung zur Aussprache-förderung im Rahmen des Englischunterrichts am XXX.

Liebe Eltern!

Derzeit wird am Institut für Anglistik, Abteilung Sprachlehrforschung und Didaktik der Goethe-Universität Frankfurt unter der Leitung von Herrn Professor Quetz ein wissenschaftliches Forschungsprojekt zur Aussprache im Englischunterricht der Unterstufe durchgeführt.

Das Erlernen der englischen Aussprache hat nicht nur Einfluss auf eine verbesserte Sprachkompetenz, sondern es kann auch das Hörverstehen erleichtern und somit den gesamten Lernprozess vereinfachen.

Daher soll in den zwei Parallelklassen von Frau X und Frau Y in der Jahrgangsstufe fünf des Goethe-Gymnasiums der Fokus des Englischunterrichts im Frühjahr 2010 vermehrt auf die englische Aussprache gelegt werden. Die Ziele und Inhalte des Englischunterrichts gemäß der hessischen Bildungsstandards ändern sich dadurch nicht und es fallen auch keine zusätzlichen Unterrichtsstunden an Auch geht kein anderer Unterrichtsstoff verloren.

Um auszuwerten, wie groß die Auswirkungen dieses leicht veränderten Sprachunterrichts sind, werden im Februar 2010, Juli 2010 und November 2010 in allen fünften Klassen Datenerhebungen durchgeführt. Die Dauer der Untersuchung beträgt pro Kind jeweils ca. fünf bis zehn Minuten. Bei dieser Untersuchung werden ca. 50 aus dem Englischunterricht bekannte Wörter auf einem Computerbildschirm gezeigt. Die Schüler/innen werden gebeten, diese Wörter laut vorzulesen und in ein Mikrophon zu sprechen. Zudem werden die Kinder gebeten, Ihre Meinung zu den verschiedene Ausspracheübungen, die sie im Unterricht kennen gelernt haben, abzugeben.

Alle Aufnahmen, Angaben und Daten werden anonymisiert und vertraulich behandelt, nicht an Dritte weitergegeben, und die Auswertung der Erhebungsunterlagen erfolgt nur im Rahmen der angegebenen Studie.

Die wissenschaftliche Untersuchung wurde gemäß des Erlasses vom 9.12.2009 durch das Hessische Kultusministerium genehmigt. Das XXX hat dem Forschungsprojekt in der Schulkonferenz vom 5.5.2009 zugestimmt.

Wir bitten Sie, in der beigefügten Erklärung Ihr Einverständnis zu geben, dass Ihre Tochter/Ihr Sohn an der wissenschaftlichen Studie teilnehmen darf.

Gerne stelle ich Ihnen das Forschungsprojekt auch persönlich auf dem nächsten Elternabend vor und stehe ich Ihnen für Fragen auch gerne per Email oder Telefon zur Verfügung.

Anne Schröter (Untersuchungsleiterin)

Einverständniserklärung zur Teilnahme an der wissenschaftlichen Untersuchung zur Ausspracheförderung im Englischunterricht.

Ich bin damit einverstanden, dass

meine Tochter _____/ mein Sohn_____

an dem Forschungsprojekt zur englischen Aussprache teilnimmt.

- Die Beteiligung an der Studie ist freiwillig. Eine Nichtteilnahme an der Studie zieht keinerlei Konsequenzen nach sich.

Wenn zutreffend, bitte ankreuzen:

- ▫ Mein Kind hat eine Hörschwäche.

Name Schüler/in:

Name Erziehungsberechtigte/r:

Kontakt:

Figure 10-5: Parent consent form: intervention group

10.3 Questionnaires

In this study a teacher and a student questionnaire were distributed. Both questionnaires can be found in the following sections 10.3.2.

10.3.1 Request for teachers to participate in the online survey

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Institut für England- und Amerikastudien ▪ Fachbereich Neuere Philologien ▪ Grüneburgplatz 1 ▪ 60323 Frankfurt am Main

An alle
Englischlehrerinnen und -lehrer
an weiterführenden Schulen in Hessen
(Bereich Sek. 1)

(mit der Bitte um Weiterleitung dieses Schreibens an die Englischlehrkräfte durch die Sekretariate)

Anne Schröter
(wissenschaftliche Mitarbeiterin)
Institut für England- und Amerikastudien
Fachbereich Neuere Philologien

Grüneburgplatz 1
60629 Frankfurt am Main

Telefon +49 (0) 69 5960 8939
E-Mail schroeter@em.uni-frankfurt.de

Betr.: Durchführung einer wissenschaftlichen Befragung als Teil eines größeren Forschungsprojekts zum Ausspracheunterricht im Englischunterricht der Unterstufe

Bezug: Genehmigung der Untersuchung durch das Hessische Kultusministerium vom 20.10.2009 (660.003.000 – 304)

2 iPod Shuffle 4GB[1] als Weihnachtsgeschenk zu gewinnen!

Sehr geehrte Damen und Herren!

Derzeit wird am Institut für Anglistik, Abteilung Sprachlehrforschung und Didaktik der Goethe-Universität Frankfurt (Professur Quetz) eine wissenschaftliche Untersuchung zum Thema „Identifizierung signifikanter Ausspracheprobleme und Möglichkeiten ihrer Verbesserung im Englischunterricht der Unterstufe" durchgeführt.

Der zu dieser Untersuchung eingesetzte Online-Fragebogen richtet sich an alle Englischlehrerinnen und -lehrer, die an **weiterführenden hessischen Schulen Englisch unterrichten** (kein LIVs).

Der inhaltliche Schwerpunkt der Fragebogenerhebung liegt auf der Erfassung des Stellenwerts der Förderung fremdsprachlichen Ausspracheunterrichts, der in der Unterstufe durchgeführt wird.

Als kleines Dankeschön für das Ausfüllen des Fragebogens werden unter allen teilnehmenden Lehrerinnen und Lehrern zwei Ipod Shuffle 4GB verlost.

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

Sie finden den Online-Fragebogen unter der Internet-Adresse:

http://limesurvey.uni-frankfurt.de/limesurvey/index.php?sid=69253&newtest=Y&lang=de

Die Bearbeitungszeit beträgt etwa 15 bis 20 Minuten. Der Fragebogen kann jederzeit zwischengespeichert und die Bearbeitung zu einem späteren Zeitpunkt fortgesetzt werden.

Der Befragungszeitrum endet am 30.11.2009

Die Teilnahme an dieser Lehrerbefragung ist freiwillig. Das Nichtausfüllen wird weder für die Lehrkräfte noch für die Schule irgendwelche Nachteile mit sich bringen.

Die Aussagekraft der mit dieser Studie gewonnenen Erkenntnisse und ihre Nutzbarmachung für die Unterrichtspraxis hängt ganz entscheidend von einer möglichst umfassenden Teilnahmequote ab. Daher bin ich, um eine valide Auswertung des Fragebogens zu gewährleisten, auf Ihre aktive Mithilfe angewiesen.

Bitte beachten Sie, dass Sie mit dem Ausfüllen des Fragebogens die nach den Datenschutzbestimmungen notwendige schriftliche Einwilligung zur Verarbeitung Ihrer Daten geben. Die erhobenen Daten werden nur für den mit der Untersuchung verbundenen Zweck verarbeitet, vollständig anonymisiert behandelt und ausschließlich für statistische Analysen verwendet.

Ich würde mich sehr freuen, wenn Sie diese Studie unterstützen würden und damit aktiv die Forschung in der englischen Fachdidaktik vorantreiben.

Ich bedanke mich schon im Voraus für Ihre Hilfe!

Mit herzlichen Grüßen

Anne Schröter

[1] Modell 3. Generation, Farbe und Gravur auf Wunsch

2

Figure 10-6: Request to take part in the online questionnaire

10.3.2 Teacher questionnaire with answers (English and German)

Online-Questionnaire: English Pronunciation Teaching at German Secondary schools in Hesse

A note on privacy: This survey is anonymous.

The record kept of your survey responses does not contain any identifying information about you unless a specific question in the survey has asked for this. If you have responded to a survey that used an identifying token to allow you to access the survey, you can rest assured that the identifying token is not kept with your responses. It is managed in a separate database, and will only be updated to indicate that you have (or have not) completed this survey. There is no way of matching identification tokens with survey responses in this survey.

Pronunciation & Speech Standard/ Aussprache & Sprachstandard

1. **How important is integrating pronunciation work into your English lessons?**
   Wie stark ist Ihr Interesse daran, Ausspracheunterricht in Ihren Englischunterricht zu integrieren?
   → Choose one of the following answers

	Very important	Important	Medium	Unimportant	Very unimportant
	Sehr stark	*Stark*	*Mittel*	*Wenig*	*Überhaupt nicht*
Valid Percentages					

2. **How would you evaluate your own pronunciation competence?**
   Wie schätzen Sie Ihre eingene Aussprache ein?
   → Choose one of the following answers

	Native-speaker competence	Near-native-speaker competence	Highly competent but with a strong German accent	Little competence
	Muttersprachliche Kompetenz	*Fast-Muttersprachliche Kompetenz*	*Hohe Kompetenz, jedoch mit starkem deutschen Einschlag*	*Eher niedrig*
Valid Percentages				

3. **How would you rate your own speech standards**[59]**?**
   Welcher Sprachstandard kommt Ihrem Englisch am nächsten?
   → Choose one of the following answers

	British English				American English	Australian English	Newzealand English	Other
	RP	General	Scots	Irish				
	`RP´	*Britisches Englisch Im weitesten Sinne*	*Schottland*	*Irland*	*Amerikanisches Englisch*	*Australisches Englisch*	*Neuseeländisches Englisch*	*Andere*
Valid Percentages								

[59] For sake of usability, many available speech standards are missing. The participants were asked to name any other speech standard.

4. **How would you describe the speech standard you use for teaching (either by yourself or with the help of media)?**
Welcher Sprachstandard kommt demjenigen am nächsten, den Sie selbst im Englischunterricht der fünften Klasse durch Ihre Sprache und durch Medien vermitteln?
→ Check any that apply

	British English				American English	Australian English	Newzealand English	Other
	RP	General	Scots	Irish				
	Britisches Englisch				*Amerikanisches Englisch*	*Australisches Englisch*	*Neuseeländisches Englisch*	*Andere*
	RP	*Im weitesten Sinne*	*Schottland*	*Irland*				
Valid Percent-ages								

Other/*Andere*:

5. **With reference to teaching English in the fith grade, what is your goal of pronunciation teaching?**
Welche Aussprachekompetenz streben Sie bei Ihren Schüler/innen im Englischunterricht der fünften Klase an?
→ Choose one of the following answers

	Native-speaker competence	Intelligibility	As long as the students talk at all, I'm happy.
	Muttersprachliche Kompetenz	*Verständlichkeit der Aussprache*	*Ich bin froh wenn die Schüler/innen reden, egal wie*
Valid Percentages			

Pronunciation & Error Correction/ Aussprache & Korrektur

6. **How important are pronunciation exercises for your English students of the fifth grade?**
Wie stark ist das Interesse Ihrer Schüler/innen im Englischunterricht der fünften Klasse an Aussprucheübungen?
→ Choose one of the following answers

	Very important	Important	Neither	Unimportant	Very unimportant
	Sehr stark	*Stark*	*Mittel*	*Gering*	*Kein Interesse*
Valid Percentages					

7. **How often do your students of the fifth grade want to be corrected with regards to their English pronunciation?**
Wie oft wollen Ihre Schüler/innen im Englischunterricht der fünften Klasse hinsichtlich ihrer Aussprache korrigiert werden?
→ Choose one of the following answers

	At every mistake	Only at profound mistakes	From time to time	Rarely	Rather not

Valid Percentages	*Bei jedem Fehler*	*Nur bei schwerwiegenden Fehlern*	*Hin und wieder*	*Selten*	*Am liebsten gar nicht*

8. **How do you correct the pronunciation of your pupils in the fifth grade of English?**
   Wie korrigieren Sie die Aussprache Ihrer Schüler/innen im Englischunterricht der fünften Klasse?
   → Check any that apply

	Present the correct version myself	Present similar sounding words	Present-ing listening examples	Presenting phonetic transcription	Drilling	Ask other students to help	Other
Valid Percen-tages	*Selbst korrekt Vor-sprechen*	*Verweis auf ähnlich klingende Wörter*	*Durch Hörbei-spiele*	*Durch phon-etische Lautschrift*	*Durch Drills*	*Andere Schüler vor-sprechen lassen*	*Andere*

9. **How often do you correct the pronunciation of your pupils in the fifth grade of English?**
   Wie oft korrigieren Sie Ihre Schüler/innen im Englischunterricht der fünften Klasse hinsichtlich ihrer Aussprache?
   → Choose one of the following answers

	At every mistake	Only at profound mistakes	From time to time	Rarely	Rather not
Valid Percentages	*Bei jedem Fehler*	*Nur bei schwerwiegenden Fehlern*	*Hin und wieder*	*Selten*	*Am liebsten gar nicht*

Pronunciation & Teaching resources/ Aussprache und Unterrichts-materialien

10. **Which of the following textbooks are you using in your English lessons in the fifth grade?**

Welches Lehrwerk benutzen Sie im Englischunterricht Ihrer fünften Klasse?
→ Choose one of the following answers

Publisher	Klett			Cornelsen		Diesterweg			
	Green Line	*Red Line*	*Orange Line*	*English G*	*Portobello Road*	*Let's go*	*Notting Hill Gate*	*Camden Town*	*Other Andere*
Valid Percentages									

11. **Do you think that there are sufficient pronunciation exercises in your textbooks and the related materials?**
Sind Sie der Meinung, dass in Ihrem Lehrwerk und den dazugehörigen Begleitmaterialien ausreichend Übungen zur Aussprache angeboten werden?
→ Choose one of the following answers

	Yes	*No*
Valid Percentages	*Ja*	*Nein*

If, yes, please answer question 12.

12. **Do you think that there should be more pronunciation exercises in the textbooks and the relating material?**
Sind Sie der Meinung, dass der Aussprachuntericht in Ihren Lehrwerken starker verankert werden sollte?
→ Choose one of the following answers

	Yes	*No*
Valid Percentages	*Ja*	*Nein*

13. **Do you use pronunciation exercises apart from your textbooks and the relating material?**
Verwenden Sie Material zur Aussprache, das nicht in dem von Ihnen benutzen Lehrwerk und den dazugehörigen Begleitmaterialien angeboten wird?
→ Choose one of the following answers

	Yes	*No*
Valid Percentages	*Ja*	*Nein*

If yes, please answer question 14:

14. **Please indicate what kind of pronunciation exercises (besides the ones in your school books) you use and where you can find these resources**[60]**.**
Geben Sie bitte an, welches Material (außer Ihrem Lehrwerk und den dazugehörigen Begleitmaterialien) Sie für Ihren Aussprachunterricht benutzen und wo Sie dieses Material finden.

[60] The participants answered the questions in German.

Pronunciation activities/ Aussprache & Unterrichtsaktivitäten

15. **How often do you use pronunciation exercises in your English lessons in the fifth grade?**
Wie oft führen Sie Übungen zur Aussprache in Ihrem Englischunterricht der fünften Klasse durch?
→ Choose one of the following answers

	Very often (in every English lesson)	Frequently (at least once a week)	From time to time (more than once a month)	Rarely (about once a month)	Hardly ever
Valid Percentages	*Sehr oft (in jeder Englischstunde)*	*Oft (mindestens einmal pro Woche)*	*Mittel (mehr als einmal im Monat)*	*Eher selten (ca. einmal pro Monat)*	*Fast nie*

16. **How often do you use each of the following pronunciation activities in your English lessons in the fifth grade?**
Wie oft benutzen Sie welche der folgenden Aktivitäten im Englischunterricht der fünften Klasse?
→ Please rate your answer in the following domains:

	Very often (in every English lesson)	Frequently (at least once a week)	From time to time (more than once a month)	Rarely (about once a month)	Hardly ever	I don't know this activity
	Sehr oft (in jeder Englischstunde)	*Oft (mindestens einmal pro Woche)*	*Mittel (mehr als einmal im Monat)*	*Eher selten (ca. einmal pro Monat)*	*Fast nie*	*Kenne ich nicht*
Drills						
Phonetic transcriptions						
Minimal pairs						
Teaching pronunciation with the help of words						
Teaching pronunciation with the help of phrases						
Teaching pronunciation with the help of sentences						
Reading passages						
Dialogues						
Tongue twisters						

Rhymes Games (e.g. Bingo) Songs Role-playes Back-chaining Vokis Speech spetrograms	

a) The answers are given with reference to total numbers (=n).

17. **When you teach pronunciation, what duration do these exercises have within an English lesson in the fifth grade**?
Wenn Sie Aussprache im Englischunterricht der fünften Klasse unterrichten, welche Dauer haben Ihre Ausspracheübungen in der Regel innerhalb einer Englischstunde?
→ Choose one of the following answers

	More than 10min	About 10min	About 5min	About 1-2min
	Mehr als 10min	*Ca. 10min*	*Ca. 5min*	*Ca. 1-2min*
Valid Percentages				

Pronunciation & Phonetics/ Aussprache & Phonetik

18. Do you think your degree in English has sufficiently prepared you to judge, correct, and teach pronunciation?
Fühlen Sie sich in Ihrer Lehramtsausbildung ausreichend dafür ausgebildet Aussprache zu korrigieren, zu beurteilen und zu unterrichten?
→ Choose one of the following answers

	Yes	No
	Ja	*Nein*
Valid Percentages		

19. **Do you think you are competent enough to use phonetic transcription in your English lessons?**
Fühlen Sie sich kompetent genug, um phonetische Lautschrift in Ihrem Englischunterricht zu verwenden?
→ Choose one of the following answers

	Yes	*No*
	Ja	*Nein*
Valid Percentages		

20. **Do you use the phonetic transcription in your English lessons in the fifth grade?**
Benutzen Sie phonetische Lautschrift in Ihrem Englischunterricht in der fünften Klasse?
→ Choose one of the following answers

	Yes	No
	Ja	*Nein*
Valid Percentages		

If yes, please answer the following question:

21. **My pupils in the fifth grade can _____ (read/ write) phonetic transcription.**
Meine Schüler/innen im Englischunterricht der fünften Klasse können die phonetische Schrift ___ (lesen/schreiben).
→ Check any that apply

	read	write
	lesen	*schreiben*
Valid Percentages		

22. **Do you think that the use of phonetic transcription helps to improve the pronunciation competence especially in a fifth grade?**
Haben Sie den Eindruck, dass der Einsatz von phonetischer Schrift gerade bei Schüler/innen einer fünften Englischklasse beim Erwerb einer guten Aussprache hilfreich ist?
→ Choose one of the following answers

	Yes, definitely	It helps for some students	No	I don't know
	Ja, eindeutig	*Bei einigen Schüler/innen*	*Nein*	*Weiß nicht*
Valid Percentages				

23. **Which of the following English sounds do you rate to be the most difficult for English learners of a fifth grade?**
Welche der folgenden Laute halten Sie für Englischlerner einer fünften Klasse am schwierigsten?

Vokale *Vowels*	ɪ sh**i**p	e b**e**d	ʊ t**oo**k	uː f**oo**d	æ fl**a**t	aː l**a**st		
	iː f**ee**t	ə p**o**tato	ɜː th**i**rd	ɔː p**ou**r	ʌ r**ou**gh	ɒ wh**a**t		
Diphthonge *Diphthongs*	ɪə f**ea**r	eə th**e**re	eɪ pl**ay**	ɔɪ t**oy**	aɪ l**i**ght	aʊ t**ow**er	əʊ fl**oa**t	
Konsonanten *Consonants*	p **p**ast	t **t**in	k pa**ck**	tʃ **ch**icken	f **f**eel	θ **th**ought	s **s**illy	ʃ **sh**op
	b **b**est	d **d**oes	g **g**reat	dʒ lar**ge**	v **v**olume	ð **th**ese	z **z**oo	ʒ lei**s**ure
	l							
	m **m**eeting	n **kn**ows	ŋ so**ng**	h **h**eavy	l **l**ovely	r **wr**iter	j **y**ellow	w **w**ild

b) The answers are given with reference to total numbers (=n).

Pronunciation & Language Background/ Aussprache & Sprachhintergrund

24. **What is the language background of your pupils in the fifth grade (in %)?**
Ihre Englischschüler/innen in der fünften Klasse haben folgende sprachliche Hintergründe (in %)?

25. **Do you think that the specific language background of an English learner accounts for specific pronunciation problems?**
Sind Sie der Meinung, dass Schüler/innen eines unterschiedlichen Sprachhintergrundes unterschiedliche Probleme in der Aussprache haben?
→ Choose one of the following answers

	Yes	No
	Ja	*Nein*
Valid Percentages		

If yes, please answer the following question:

26. **Which pronunciation problems do you consider to be due to a specific language background?**
Welche Aussprachepobleme führen Sie auf den unterschiedlichen Sprachhintergrund einzelner Schüler/innen zurück?

I) in total 142 participants gave valid answers (n=142).

27. **Which of the following sounds do you consider to be particular difficult for learners of the following languages?**
Welche der folgenden Laute halten Sie für Sprecher/innen der folgenden Sprachen für besonders schwierig?

		ɪ sh**i**p	e b**e**d	ʊ t**oo**k	uː f**oo**d	æ fl**a**t	aː l**a**st		
Vokale *Vowels*	German								
	Turkish								
	Russian								
	Polish								
	Other								
		iː f**ee**t	ə p**o**tato	ɜː th**i**rd	ɔː p**our**	ʌ r**ou**gh	ɒ wh**a**t		
	German						6		
	Turkish								
	Russian								
	Polish								
	Other								
		ɪə f**ear**	eə th**ere**	eɪ pl**ay**	ɔɪ t**oy**	aɪ l**i**ght	aʊ t**ow**er	əʊ fl**oa**t	
Diphthonge *Diphthongs*	German								
	Turkish								
	Russian								
	Polish								
	Other								
		p **p**ast	t **t**in	k pa**ck**	tʃ **ch**icken	f **f**eel	θ **th**ought	s **s**illy	ʃ **sh**op
Konsonanten *Consonants*	German								
	Turkish								
	Russian								
	Polish								
	Other								
		b **b**est	d **d**oes	g **g**reat	dʒ lar**ge**	v **v**olume	ð **th**ese	z **z**oo	ʒ lei**s**ure
	German								
	Turkish								
	Russian								
	Polish								

	Other								
		m **m**eeting	n **kn**ows	ŋ so**ng**	h **h**eavy	l **l**ovely	r **wr**iter	j **y**ellow	w **w**ild
	German								
	Turkish								
	Russian								
	Polish								
	Other								

c) The answers are given with reference to total numbers (=n).

Other/*Andere*

German Deutsch (n=467)	
Turkish/Türkisch (n=196) Russian/Russisch (n=80) Polish/Polisch (n=53) Other/Andere (n=54)	

Pronunciation & Teaching perspective/ Aussprache & Lehrerperspektive

28. **Please rank the following activities according to the importance they have for you in the English lessons of the fifth grade.**
Bitte ordnen Sie folgende Aktivitäten gemäß dem Stellenwert, den Sie ihnen in ihrem Englischunterricht in der fünften Klasse beimessen.

	Grammar	Vocabulary Training	Listening Comprehension	Pronunciation	Communicative Competence	Reading Comprehension
	Grammatik	*Wortschatzarbeit*	*Hörverstehen*	*Aussprache*	*Kommunikative Kompetenz*	*Leseverstehen*
Valid Percentages						

29. **Do you think that pronunciation exercises should be more strongly integrated into English lessons?**
Sind Sie der Meinung, dass der Ausspracheunterricht im Englischunterricht stärker verankert werden sollte?
→ Choose one of the following answers

	Yes	No
	Ja	*Nein*
Valid Percentages		

30. **Which aspects of pronunciation teaching do you find particularly interesting?**
Welche Aspekte des Ausspracheunterrichts interessieren Sie besonders?

d) in total 141 participants answered this question (n=141).

31. **Would you like to take part in teacher training concerning pronunciation?**
Würden Sie gerne an Lehrerfortbildungen zum Thema Aussprache teilnehmen?
→ Choose one of the following answers

	Yes	No
Valid Percentages	*Ja*	*Nein*

32. **At what type of school are you teaching?**
An welcher Schulform unterrichten Sie?
→ Check any that apply

	~ comprehensive school	~ integrated comprehensive school	school for children with learning difficulties	~ secondary modern school	~ secondary modern school	~ grammar school
Valid Percentages	*Gesamtschule*	*Integrierte Gesamtschule*	*Förderschule*	*Haupt-schule*	*Real-schule*	*Gymnasium*

33. **Which English students do you mainly teach?**
In welcher Schulstufe unterrichten Sie hauptsächlich?
→ Check any that apply

	Students at the ages of about 10-12	Students at the ages of about 13-15	Students at the ages of about 16-19
Valid Percentages	*Unterstufe*	*Mittelstufe*	*Oberstufe*

34. **How many years have you been teaching English?**
Seit wie vielen Jahren unterrichten Sie Englisch?
→ Choose one of the following answers

	Less than 5	6-10	11-14	15-19	20-24	25-29	More then 25
Valid Percentages	*Weniger als 5*	*6-10*	*11-14*	*15-19*	*20-24*	*25-29*	*Mehr als 25*

35. **What is your sex?**
Welchem Geschlecht gehören Sie an?
→ Choose one of the following answers

	Male	Female
Valid Percentages	*Mann*	*Frau*

10.3.3 Student Questionnaire

Online-Questionnaire: Student Questionnaire

A note on privacy: This survey is anonymous.

Dear Student

Thank you for taking part in this questionnaire about English pronunciation teaching. The following questions refer to your English lessons. It is very important that you answer the questions honestly. The questionnaire has NO impact on your English mark. It is only used to get your and your classmates' opinion on English pronunciation teaching. If you have any questions, please don't hesitate to ask and you'll get further explanations.

The questionnaire is anonymous, so, nobody – not even your teacher - knows about your answers. They are only important for the evaluation of the questionnaire.

The questionnaire consists of 17 questions.

Demographic Questions /Demographische Fragen

1. Dear student! To match your questions without giving up anonymity, you need to use a code word. It consists of the first two letters of your name, the first two letters of your surname and the first two numerals of your birth date e.g.: Franz Böhme 01.01.1999 → FR BÖ 01
→ Please, insert your code word here: __ __ __

2. Which is your class?
In welche Klasse gehst Du?
→ Choose one of the following answers.

	Grammar School B					*Comprehensive School A*				
Class Total	5a	5b	5c	5d	Overall	5c	5d	5e	5f	Overall

3. When is your birthday? (Example: 15.12.1997)
Wann bist Du geboren?
→ Please choose from the following chart.

4. What is your school? → see next question
In welche Schule gehst Du?
→ Choose one of the following answers.

- o Grammar school 1
- o Comprehensive school 2

5. Are you male/female?
Bist Du ein Junge/Mädchen?
→ Choose one of the following answers.

- o male
- o female

6. What is your nationality?
Welcher Staatsangehörigkeit gehörst Du an?
→ Check any that apply

- o German
- o English
- o French
- o Other : ____
- o Greek
- o Italian
- o Polish
- o Russian
- o Spanish
- o Turkish
- o American
- o Chinese
- o Korean

7. What date is it today?
Welches Datum haben wir heute?
→ Please, choose from the following chart.

8. Which languages can you speak? What's your level?
Welche Deiner Sprachen kannst Du wie gut sprechen?
→ Check the answer that is most suitable to you.

	Excellent; It doesn't pose any problems	Good; Although I sometimes make mistakes	Average skills; I can exchange information on familiar topics	Not very good; I can understand some words and contents	I cannot speak the language
German					
English					
French					
Latin					
Greek					
Italian					
Polish					
Russian					
Turkish					
Mandarin					
Persian					
Hindi					

9. Can you speak any other languages that are not listed above (see question 8)? If yes, which language is it and what is your level?
Gibt es eine Sprache, die Du sprichst, die aber auf der voherigen Listen nicht aufgetaucht ist? Wenn ja, welche und wie gut sprichst Du sie?
→ Please, write your answer in the following box.

	Excellent;	Good;	Average skills;	Not very good;	

	It doesn't pose any problems	Although I sometimes make mistakes	I can exchange information on familiar topics	I can understand some words and contents	I cannot speak the language
Other:___					
Other:___					

10. When did you start studying English at school?
Seit wann lernst Du Englisch in der Schule?
→ Choose one of the following answers.

o Year 1 o Year 2 o Year 3 o Year 4 o Year 5

11. I've already studied English in kindergarten.
Ich hatte schon Englisch im Kindergarten.
→ Choose one of the following answers.

o Yes o No

12. I've stayed longer than four weeks in an English-speaking country.
Ich habe schon einmal länger als vier Wochen im englischsprachigen Ausland verbracht.
→ Choose one of the following answers.

o Yes o No

If yes;

13. If yes, please tell me where you went. How long did you stay?
Wenn ja, wo und wie lange warst Du im englischsprachigen Ausland?
→ Please, write your answer in the following box.

Questions concerning pronunciation teaching /Fragen zur Aussprache

14. Below there are some statements with regards to pronunciation teaching. Please, check any fitting statement. If you are unsure what box to tick, please choose the answer that is most likely.
Bitte kreuze das für Dich zutreffende an. Wenn es Dir schwer fällt Dich zu entscheiden, wähle die Antwort aus, die am ehesten Deiner Meinung entspricht.
→ Check any that apply.

	Strongly agree	Agree	Rather disagree	Disagree
In my class everyone struggles with the pronunciation of English words	O	O	O	O

My English pronunciation is very good	O	O	O	O
I think that the English pronunciation sounds strange.	O	O	O	O
I really like English pronunciation exercises	O	O	O	O
I would love to have a better English pronunciation	O	O	O	O
One day I want to sound like a real English native speaker	O	O	O	O
I don't think that pronunciation is important because it is not part of any class tests.	O	O	O	O
Doing pronunciation exercises in class is a waste of time	O	O	O	O
I cannot understand the speakers on the audio material as they are always mumbling.	O	O	O	O
It's easy to understand the audio material in class	O	O	O	O
It's hard to understand each word in the English classroom I need to know how the words are pronounced in order to understand them.	O	O	O	O
It's possible to acquire a good pronunciation	O	O	O	O
It's difficult to pronounce English words as the writing doesn't always match the pronunciation.	O	O	O	O

15. What pronunciation aspects do you perceive as particularly difficult?
Was an der englischen Aussprache findest Du besonders schwierig?
→ Please, write your answer in the following box.

16. Do you like pronunciation exercises in your English lessons?
Machen Dir Übungen zur Aussprache im Englischunterricht Spaß?
→ Choose one of the following answers.

o Yes o No

17. Which pronunciation exercises do you know?
Welche Übungen zur Englischen Aussprache kennst Du?
→ Choose one of the following answers.

- o The teacher models the words/ sentences and we are repeating it/ them.
- o We use audio cassettes and CD's
- o Other: _____

10.4 Pronunciation Intervention: Teachers' Manual

10.4.1 Teacher's manual: Introduction and vowels

Teachers' Manual: Vowels /ɑ:/, /ɔ:/, /æ/, and /ə/

1. The teacher's manual: Some information

Children learn language through gestures (watching and learning how their caretakers produce words using their articulators), by listening and by repeatedly trying to produce sounds that sound like language until they are finally understood. Similar procedures take place when children learn a second language. Therefore, all tasks in this manual are divided into a presentation, discrimination and identification, and a production phase.

Before the training begins there is a short phase that introduces ideas to explain the parts of the mouth and the mouth movements needed for the pronunciation training. The sound chart with the phonetic code is also presented at this point.

After that the vowels, diphthongs, and consonants which seem to be especially difficult for learners of English are introduced. As the new-sound presentation phase is the most essential part of the training it is this phase which takes the longest to complete (about 25min). However, this can be split - if needed. The other activities take around five to ten minutes each.

A suggested timeline of when to teach which activity is included. The explanations and tasks are given in English but can be taught in German as well. The tasks and descriptions are suggestions only and can be varied.

The use of video is always indicated in red and needs only to be used in each presentation phase. Its use is optional in the discrimination and production phase. ITunes and Quick Time can be installed from the following web pages if necessary:

http://www.apple.com/de/itunes/download/

http://www.apple.com/de/quicktime/download/

At some point during the beginning of the pronunciation training the students will be asked to fill in an online questionnaire describing their language background and motivation in the computer room. This will probably take them around 15 minutes. I gladly help with this task.

Thank you very much for taking part in the pronunciation training!

1

2. Introduction

a) Task I1: Introduction to the pronunciation teaching

Aim: Introduction of the knowledge needed for the pronunciation work - raising awareness that all kinds of articulators are used for pronunciation.

Time		Media
Ca.20min	Andy the Pronunciation Android[1] Task 1, 2, 3, 4 (5 optional) on the handout	Worksheet I1 Andy the Android AudioCD Track I1 Andy the Android Track I2 Andy the Android

b) Task I2: Teaching the chart

Do not teach the whole chart in one lesson. This will be too much for students to take in at one time. Focus instead on the sounds as they appear in the activities. Put the sound chart and the sounds you are dealing with on the wall as you proceed with the activities.

Hand a sound chart to each of your students and explain that you are going to use the worksheet over the following weeks.

Aim: Students should understand that a "phonetic code" exists and that it is useful because English sounds are pronounced differently to their German equivalents.

Time		Media
Ca.5min	T.: shows S. the poster with the phonetic code. T.: What can you see on the poster? Have you seen it before? Where? Why do you need it? ➔ Phonetic Code ➔ These sounds tell you how to pronounce words T.: Why don't you just use "normal" letters? ➔ Teacher writes example words on the blackboard ➔ Because sometimes one English letter can be pronounced differently, e.g. bad, arm, small T. hands out the phonetic code charts. Explains that students are going to use it in the following weeks.	Phonetic Classroom Poster Phonetic Chart Phonetic Student Chart Blackboard T writes bad, arm, small on the blackboard

3. Vowels (-ca. 18.3.2010)

Presentation

c) Task V1: Long and short vowels (together about 25min)

Aim: Students are introduced to the vowels /ɑ:/, /ɔ:/, /æ/, and /ə/. They should understand that there are long and short vowels. Students watch and produce the movements needed in order to produce each one of the four sounds.

[1] (Bowler 2006: 45)

Long Vowel Sounds and /ɑ:/ and /ɔ:/ (about 15min)

If there is enough time, the following four vowels (the long and the short ones) should be presented in one lesson (this will mean that the video is only needed during one session). Otherwise, the task can be split and used in two or four lessons.

Time		Media
Ca. 8min	T. shows the poster/ɑ:/ T.: What is it? Can you find it on the poster with the phonetic code? Students should find out that /ɑ:/ is a phonetic code for the letter [a]. Students watch the video and repeat after the speaker. / ɑ:/ is a long sound. Relax your lips and open your jaw to make this sound. Make words very long Students repeat (also possible as a drill) the words from the poster: arm, far, grass, aunt Students should write one reference word for/ɑ:/ in their own phonetic chart T.: Why is there a colon after the sound? → Long vowels T.: How many other sounds with a colon can you find on your phonetic chart? (→ Answer: 5)	Poster V1: Vowel /ɑ:/ CD V: Video V1 Time: 00:49min Phonetic Code Chart
Ca. 7min	T.: There is one sound on your chart that looks like this: /ɔ:/ T. shows the poster /ɔ:/ T.: Do you already know this sound? What does it stand for? Students watch the video and repeat after the speaker. T.: /ɔ:/ is a long sound. Open your mouth a little and make your lips round. Students repeat (also possible as a drill) the words from the poster: walk, all, draw, Laura Students should write one reference word for /ɔ :/ in their own phonetic chart T.: Can you also find this sound in the German language? → No, → Thus, there is phonetic code. **Exercise book entry:** German and English sounds are pronounced differently. We use phonetic code to know how to pronounce the English words. A colon indicates a long sound.	Poster V2: Vowel/ɔ:/ CD Video V2 Time: 00:59min Phonetic Code Chart Exercise book entry Poster V1,/ɔ:/ and V2 /ɑ:/ are put on the wall

Short vowels sounds /æ/, /ə/

Time		Media
Ca. 10min	T.: Remember - we talked about long vowels. If there are long vowels there are also short vowels. Can you show me the short vowels on the phonetic chart?	Classroom Poster Phonetic chart

3

	T.: What sound is: /æ/ (T. draws the sound /æ/ on the blackboard) Teacher discusses ideas with the students. Do you already know some words that have this sound? Students watch the video and repeat after the speaker. /æ/ is a short sound. Spread your lips wide and open your jaw to make this sound (like eating an apple). Make the words very short. Teacher drills the words from the poster: bad, thank, clap, lamb T. : Do you know Tarzan? Tarzan and Jane? What sound does he make? Do you want to listen to him? Can you repeat after Tarzan? Relax your face and open your mouth a little. Imagine someone pushes you gently in the stomach. Make the sound as small and as short as you can. (It's also the sound made when s.o. vomits) Students repeat (also possible as a drill) the words from the poster: an, from, Laura, clever	Blackboard /æ/ CD Video V3 Time: 00:53min Poster Vowel /æ/ Audio CD Track V1, /ə/ Time: 00:16min Poster /ə//æ/ are put on the wall

d) Task V2: Mouth practice of long and short vowels

Time		Media
Ca. 5min	Students should use the words from the vowel posters and produce them silently while watching their mouths in the mirror. They should describe (and discuss) their mouth, tongue, and teeth movements to (with) their partner. T.: "Work in pairs and one student silently produces a word. The other student has to guess it." (and vice versa")	Mirrors (can be found in the tea boxes) Words from the posters: bad, thank, clap, lamb, an, from, Laura, clever, walk, all, draw, arm, far, grass, aunt (other words can also be used)

Identification & Discrimination

In case time is short, only task can be taught.

a) Task V3: Long or Short Vowel

Zeit		Media
Ca. 5min	T.: Can you hear the long (short) vowels? Teacher reads a list of words containing long and short vowels. The students should listen for the long sounds. Every time they hear a long sound they should raise the positive smiley and vice versa.	Smileys (☺ ☹). (a positive and a negative smiley for each student) For a start the known words should be used: thank, clap, lamb, an, from, Laura, clever, walk, all, draw, arm, far, grass, aunt. After that words can be taken from the text book or anywhere else

4

Production

a) Task V4: Practising /ɑ:/, /ɔ:/, /æ/ and /ə/

Time		Media
5min	T. puts all the posters /ɑ:/, /ɔ:/, /æ/ and /ə/ on the blackboard. (They'll glue on the blackboard if the back of them is moistened with a bit of water) Work in pairs. Write down **one** sentence. The sentence should have a meaning. Each of the vowels /ɑ:/, /ɔ:/, /æ/ and /ə/ has to appear at least **twice (2x)** in it. Underline the vowels in the sentence. Put on your English shoes and try to sound like the English Queen. T. has to model the Queen's accent Practise the sentence and your accent with your partner. Walk around in the classroom. Read your sentence to a student from another group. He/She should try to repeat it. Then change roles and find other partners. Which sentence is the best/funniest, etc.?	Poster V1, V2, V3, and V4 Vowel /ɑ:/, /ɔ:/, /æ/ and /ə/ Transparency V1 A transparency with the task is added (there is also a German translation)

Worksheet: Andy the Android

The worksheet: "Andy the Android" by Bill Bowler (Bowler, 2005, p. 45) cannot be displayed due to copyright. It introduces the vocabulary needed for the pronunciation intervention in a visual format (e.g. touch your voice box, put your tongue between your teeth, breath out, etc.). The task is to match the new words to the articulatory gestures displayed by Andy the Android. In addition, the students have to practice the gestures themselves with the help of a "Simon says" activity.

Vowel sounds /ɑ:/, /ɔ:/, /æ/ and /ə/

Work in pairs. Write down **one** sentence. The sentence should have a meaning(Bedeutung).

Each of the vowels /ɑ:/, /ɔ:/, /æ/ and /ə/ has to be used at least **twice (2x)** in the sentence.

Underline the vowels /ɑ:/, /ɔ:/, /æ/ and /ə/ in the sentence.

Put on your English shoes and try to sound like the English Queen.

Practise the sentence with your partner.

Walk around in the classroom. Read your sentence to a student from another group. He/She should try to repeat it.

Then change roles and find other partners.

Which sentence is the best/funniest, etc.?

Transparency V1

Vokale /ɑ:/, /ɔ:/, /æ/ und /ə/

Schreibt in Partnerarbeit einen englischen Satz. Der Satz soll einen Sinn ergeben.

Der Satz muss mindestens jeden der geübten Vokale /ɑ:/, /ɔ:/, /æ/ and /ə/ **zweimal** enthalten.

Unterstreicht die Vokale in dem Satz.

Lest euch gegenseitig den Satz vor.

Versucht dabei so wie die englische Königin zu reden.

Steht auf und präsentiert euren Satz einem Mitschüler/in eurer Klasse. Diese/r soll ihn möglichst auswendig wiederholen.

Präsentiere deinen Satz so vielen Mitschülern wie möglich.

Welcher Satz ist der beste, lustigste, etc?

Transparency V1

/ɔ:/	/ɑ:/	/æ/	/ə/
w**a**lk /wɔ:k/	**a**rm /ɑ:m/	b**a**d /bæd/	**a**n /ən/
all /ɔ:l/	f**a**r /fɑ:/	th**a**nk /θæŋk/	fr**o**m /frəm/
dr**a**w /drɔ:/	gr**a**ss /grɑ:s/	cl**a**p /klæp/	Laur**a** /lɔ:rə/
L**au**ra /lɔ:rə/	**au**nt /ɑ:nt/	l**a**mb /læm/	clev**er** /klevə/
V1	V2	V3	V4

Figure 10-7: Posters Vowels V1-V4

10.4.2 Teacher's manual: Diphthongs

Teachers' Manual: Diphthongs

1. Introduction

a) Task D1: /ɪ/, /i:/, /ɔ/, /e/

Introducing the sounds needed for the pronunciation of diphthongs:

Time		Media
20min	Does anyone know the following sound? → /ɪ/ (T shows Poster D1 /ɪ/ and puts it on the board) Do you know a word that contains this sound? (Examples are collected on the black board) T.: Is the sound /ɪ/ long or short? → short T.: How do you indicate a long vowel? →/:/ T.: Try to find the long /iː/ on the sound chart. (T shows Poster D2 and puts it on the board) T.: Do you know a word that contains this sound? (Examples are collected on the black board) Drill: Teacher shows two pictures of the neutral and spread mouth positions and asks which equals the long /i:/ and which the short /ɪ/. T shows the pictures to the students with varying speed. Every time the mouth is spread the students should say one example word with a long /i:/, eg. sheep. Every time the T shows the neutral mouth position the students should say an example word with the short /ɪ/, e.g. ship. Drill repetition with the two example words about 10 times e.g. Ship, sheep, ship, sheep, sheep, ship, ship, sheep, sheep, sheep, ship, ship, sheep etc.	Poster D1: Vowel /ɪ/ (without examples.) suggested blackboard sketch see below Poster D2: Vowel /iː/ (without examples) Picture D1 Spread mouth position Picture D2 Neutral mouth position
	T.: Do you remember the schwa /ɔ/? It was the sound that Tarzan made. Can anybody repeat Tarzan's sound? (T. shows schwa poster and puts it on the board) T.: Can you give me any examples for words with an /ə/? (Examples are collected on the black board, see below)	Poster D3: Vowel /ɔ/ (without examples)
	T.: Look at this /e/ sound (T puts the /e/ poster on the board). What does it look like? Can you think of any words with this sound? (Examples are collected on the black board, see below)	Poster D4: Vowel /e/ (without examples)
	T.: Fill in one example word for the vowels /ɪ/,/i:/,/ɔ/, and /e/ on your own sound chart.	Worksheet V1: Phonetic Chart

1

Possible use of the black board:

(Instead of drawing the vowels on the black board, the laminated posters can be put on the black board (they will stick if the back of the poster is moistened with a little water)

Long vowel	**Examples** (different examples are also possible)	**Short vowels**	**Examples** (different examples are also possible)
i:	sheep, people, etc	**ɪ**	ship, big
		ɔ	from, Laura
		e	egg head

2. Presentation

a) Task D2: Diphthong features

Students should understand what diphthongs are and how they are pronounced

Time		Media
5min	T. shows video at least twice: "When two vowels go walking" T.: What was the video about? (T. collects ideas) (ideas could be, two vowels, first one is louder, there is a glide (N.B: glide is a new piece of vocabulary) etc.) T.: When there are two vowel sounds together you call them: "diphthongs"	Video D1: "When two vowels go walking" Vocabulary: To glide (v.)→rutschen Diphthong (n.) → Doppellaut
5min	T. hands out Poster D1 containing the vowel /ɪ/ to one student and poster D3 containing the Vowel /ɔ/ to another student. Both students should stand facing the class, about 2m apart from one another, and displaying the posters to the class. Every time one of the students moves, the class shouts the sound. First the student with the /ɪ/ moves sideward towards the student with the /ɔ/. The class says /ɪ/. Then the student with the /ɔ/ moves sideward towards the student with the /ɪ/ and the class says /ɪ/ and so on. The closer the students move towards each other, the louder and longer the first vowel should be pronounced in relation to the second until a glide between the two vowels is established.	Poster D1: Vowel /ɪ/ (without examples) Poster D3: Vowel /ɔ/ (without examples) Poster D4: Vowel /e/ (without examples)

2

	/ɪ/ - - - /ə/ → /ɪ/ - - /ə/ → /ɪ/ - /ə/ → /ɪə/ As soon as the students are standing right beside each other the diphthong should be pronounced. T. repeats the same procedure as mentioned above with /eɪ/	
5min	T. shows the posters with the diphthongs /ɪə/ and /eɪ/ and repeats the example words: year/dear and page/grey with the class. T.: Fill in one example word for the diphthongs /ɪə/ and /eɪ/ on your own sound chart. T writes the following rule on the blackboard. The students should copy it in their exercise books: **Diphthong (Doppellaut)** **In a diphthong there is a glide from one vowel sound to another. The first sound is longer and louder than the second.**	Poster D5: Vowel /ɪə/ (with examples) Poster D 6: Vowel /eɪ/ (with examples) Phonetic Code Chart

b) Task D2: Getting to know the diphthongs

Time		Media
5min	Students should produce some words containing the diphthongs /ɪə/ and /eɪ/ silently while watching their mouths in the mirror. They should describe (and discuss) their mouth, tongue, and teeth movements to (with) their partner. T.: "Work in pairs. One student should silently produce a word and the other student has to try and guess it." (and vice versa)	Mirrors

3. Identification & Discrimination

c) Task D2: Diphthongs

Time		Media
15-20min	T.: Do you remember diphthongs. What are they? T puts the poster /eɪ/ on the blackboard T.: (/eɪ/) : Pretend not to hear someone and say, eh? (Repeat several times) T.: Try it aloud, whispered, and silently.	Poster D6: Vowel /eɪ/

3

	Then the T. repeats the example words "page" and "grey" T puts the poster /æ/ next to the poster /eɪ/ on the blackboard. T shows the mouth picture of /æ/ and repeats the sound with the students. Then the T repeats the example words: bad, clap, thank, lamb with the students. T. reads wordlist 1. Students should fill in the words in the correct column of the worksheet D1. Wordlist 1: say, clap, hat, cake, way, page, man, grey, fat, they, cat, lamb	Mouth picture D3 Worksheet D1
	T puts the poster /ɪe/ on the blackboard T.: Make the sound /ɪɔ/ while tugging your ear Then the T. repeats the example words "year" and "hear" with the students T puts the poster /i:/ next to the other poster /ɪe/ on the blackboard and asks the students to name some example words: e.g "sheep, key, people, cheese" T. reads a list of words on the worksheet. Students should fill in the words in the correct column of the worksheet D1. Wordlist 2: me, police, teen, key, fear, hear, keep, ear, idea, clear, sheep, year	Poster D5: Vowel /ɪɔ/ Worksheet D1
5min	Correction of the worksheet using transparency D1	Transparency D1

Solutions: Worksheet D1

/eɪ/	/æ/	/ɪɔ/	/i:/
page	hat	hear	sheep
grey	clap	year	keep
cake	lamp	fear	key
way	cat	ear	police
say	fat	idea	me
they	man	clear	teen

4. Production

d) Task D2: Diphthongs

This task can also be given as homework

Time		Media
Homework + some short presenta-tions	T.: Write a poem. The poem should have four lines. You have to use at least one of the new diphthongs (/eɪ/, /æ/, /ɪɔ/, /i/) in each line. Present your poem to the class and pay special attention to the pronunciation of the diphthongs.	Worksheet D1: (Transparency for the solution is added)

4

Worksheet D1

Task1: Diphthong: /eɪ/ vs. vowel /æ/

Listen to the words in the box on the right. Put the words into the correct column.

/eɪ/	/æ/	1. say 2. clap 3. hat 4. cake 5. way 6. page 7. man 8. grey 9. fat 10. they 11. cat 12. lamb

Task 2: Diphthong /ɪə/ vs. vowel/i:/

Listen to the words in the box on the right. Put the words into the correct column

/ɪə/	/i:/	1. me 2. police 3. teen 4. key 5. fear 6. hear 7. keep 8. ear 9. idea 10. clear 11. sheep 12. year

Task 3: Write a poem. The poem should have four lines. You have to use at least one of the diphthongs **/eɪ/** or **/ɪə/** and one of the vowels **/i:/ /æ/** in each line.
Present your poem to the class and pay especially attention to the pronunciation of the diphthongs.

/ɪ/ D1	/i:/ D2	/ə/ D3	/e/ D4	/ɪə/ D5	/əɪ/ D6
				year /jɪə/ d**ear** /dɪə/	p**a**ge /peɪdʒ/ gr**ey** /greɪ/

Figure 10-8: Diphthong Posters D1-D6

10.4.3 Teacher's manual: Plosives

Teachers' Manual: Plosives: /b/,/d/, and /g/ in final position

1. Vowels and Consonants

Because consonants are introduced at this point of the training, a brief definition of vowels and consonants will be given to the students at the beginning of the unit .

2. Voiced vs. voiceless

When working with consonants, it is helpful to understand the difference between voiced and voiceless consonants, because that is often the only distinguishing feature between two sounds.

a) Task P1: Introduction of the terms: consonant and vowel

This task can be done quite quickly or even omitted if the students already know about the difference between consonants and vowels.

Time		Media
5-10min	T. asks the students to name some vowels and then some consonants. T. writes /ɑ:/ and /d/ (other sounds are also possible as well) on the blackboard and asks the students to pronounce the /ɑ:/ and then the /d/. T. asks the students to concentrate on the movement of their tongue and mouth. They should recognize that, unlike vowels, there is an obstruction when pronouncing consonants. The students will then copy the definitions from the transparency. (If time is short, the definition on the transparency can be copied and given to the students as well.)	Vocabulary: Interruption (v.) – die Störung
	Vowels A speech sound by which the air from the lungs goes through the mouth WITHOUT interruption. **Consonants** A speech sound by which air from the lungs goes through the mouth WITH an interruption.	Transparency P1 → see below

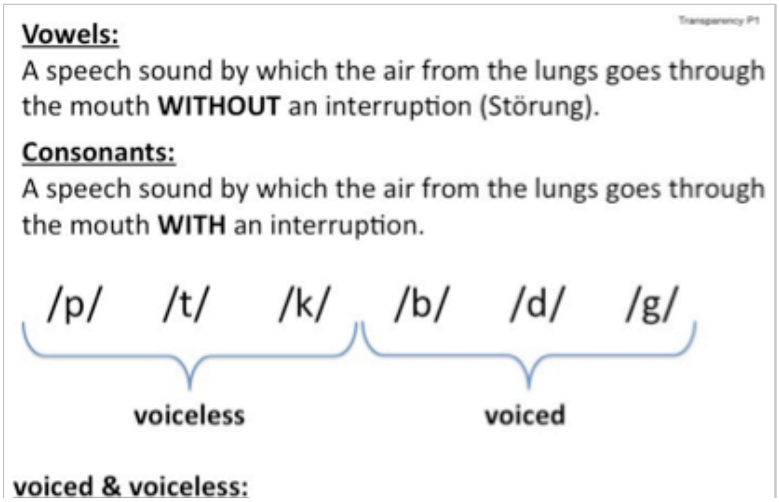
b) Task P2: Introduction to voiced and voiceless consonants

Time		Media
10min	T. writes the following sounds on the blackboard: /p/ /t/ /k/ /b/ /d/ /g/	blackboard → see below

1

	T. hands out a piece of paper to each student. The students are asked to hold it in front of their lips and pronounce the sounds on the blackboard. They should exaggerate the pronunciation. T. asks the students whether something happens when pronouncing the plosives → The paper should move for /p/, /t/, and /k/, but not for /b/, /d/, and /g/.	Envelope with pieces of paper
	T.: Touch your voice box and make all the sounds again. Do you feel a difference? → There should be a vibration of the voice box/vocal chords for for /b/, /d/, and /g/ but not for/p/, /t/, and /k/.	
	T. marks the voiced and the voiceless consonants on the blackboard and explains the words voiced and voiceless.	blackboard → see below
	Students shall copy the definitions from the transparency. (If time is short, the transparency can be copied and given to the students as well.) Voiced & voiceless The voice box vibrates for voiced consonants but there is **no** vibration for voiceless consonants.	Transparency P1

Possible use of the blackboard (also included as a transparency and as a handout)



The vocal cords (Stimmbänder) vibrate for **voiced** consonants but there is **NO** vibration for **voiceless** consonants

3. Introduction to /b/, /d/, and /g/ in final position

Most Learners of English do not have any problems in pronouncing the sounds /b/, /d/, and /g/ correctly but especially German learners of English typically encounter problems with the voiced consonants such as /b/, /d/, and /g/ at the end of words e.g. club, bad, and frog. This is mainly due to the fact that the German language always asks for a terminal devoicing (Auslautverhärtung) as in "Tag"→ /ta:k/, and "wird" → /wirt/.[1] As a consequence, when /b/ /d/ and /g/ occur in final position learners will tend to use /p/, /t/, and /k/ respectively, such that "bag" may sound like "back". It is therefore important to emphasize in the classroom that the final /b/ /d/ and /g/ are pronounced very "heavily" in English[2].

4. Presentation

Because the use of /b/, /d/, and /g/ is problematic, rather than the production of these sounds, there will be no video for this unit.

a) Task P3: b, d, g in final position

Time		Media
5min	T. shows an index card with "cap" and one with "cab" to the students. T. then pronounces one of the words and the students have to find out whether the T. has spoken "cab"or "cap". Then the students have to close their eyes and the teacher will ask them again to distinguish between "cap" and "cab". T. asks students why they might think it is so difficult to hear the difference between the /p/ and /b/?	Flash Cards P1, P2 cap & cab
	T. says: Guten Tag!" Which sound do you hear at the end of Tag? →/k/ and not /g/ T. : In the German language the voiced consonants /b/, /d/, /g/ are always pronounced as the voiceless consonants /p/,/t/, /k/ when they occur at the end of a word. Therefore, in English the students have to over-pronounce and exaggerate /b/, /d/, and /g/ when they occur at the end of a word.	Worksheet P1 see "Rule 1"

b) Task P4: Features of the plosives

Time		Media
10min	T. hands out the worksheet and asks the students to listen to the pairs of words on the CD. They are to pay	Worksheet P1: CD P1: (38s) Cap or

[1] (Quetz 1998: 52)
[2] (Kenworthy 1987: 137)

	particular attention to the sounds of the underlined letters.	Cab?
	T. draws attention to the length of the words on the left and right side of the worksheet asks students to listen again.	
	T. asks one student to read the following rule on the worksheet: The vowels before the voiced consonants /b/, d/ and /g/ (e.g. bag) are longer than before the voiceless sounds /p/ /t/ /k/ (e.g. back).	Worksheet P1 See "Rule 2"
	T. reads the 6 words (cap, cab, bed, bet, bag, and back) together with the student and exaggerates the pronunciation.	

5. Identification & Discrimination

a) Task P5: Discrimination of the plosives /b/, /d/, and /g/ in final position

If its too much trouble to use the CD, the teacher can also read the sentences.

Time		Media
5min	T. works again with the worksheet P1 and asks the students to listen to the sentences and to circle the words they hear on the CD. cab, cap, bed, bet, bag, back	Worksheet P1 CD P2 (16s) What's wrong?
	The transparency can be used for the correction	Transparency of the Worksheet P1

6. Production

a) Task P6

Time		Media
5min	T. puts the phonetic posters P1, P2, and P3 on the wall and drills the pronunciation of the words on the posters several times: job, web, bad, bread, big, egg,	Phonetic Poster P1, P2, P3
	Students fill in an example word for /b/,/d/,/g/, /p/, /t/, and /k/ on their own sound chart.	Worksheet V1 Phonetic Chart

a) Task P7: Speaking as fast as possible

Time		Media
2min	T. puts the transparency P2 on the OHP. The students should practise the sentence several times with their partner and say it as fast as possible but with the correct pronunciation: The sad bad dog did his job and went after the fat cat in the bed.	Transparency P2

Worksheet P1

Cap or Cap?[3]

Task P1: Listen to the pairs or words. Pay attention to the sounds of the underlined letters.

In the original manual, pictures of a cap and a cab, a bed and a bet, as well as a bag and a back taken from Bowler (2005: 34) are displayed.

/b/ bag	/p/ cap
/d/ bed	/t/ bet
/g/ bag	/k/ back

Task P2. Listen to these sentences. Circle the words you hear.

1. What's wrong with your **bag / back ?**
2. Let's get a **cab / cap !**
3. Do you want a **bed / bet ?**

Rule 1

In the **German** language the voiced consonants /b/, /d/, /g/ at the **end of words** are always pronounced as the voiceless consonants /p/,/t/, /k/.

Example → Guten Ta**g**! is pronounced as "Guten Ta**k**" !

In English you have to pronounce /b/, /d/, and /g/ at the end of the words.

Rule 2

The vowels before the voiced consonants /b/, d/ and /g/ (e.g. ba**g**) are long.
The vowels before the voiceless sounds /p/ /t/ /k/ (e.g. bac**k**) are short.

[3] Adapted from Bowler 2005: 34

/b/	/d/	/g/	
job /dʒɒb/ web /web/ P1	bed /bæd/ bread /brɛd/ P2	big /bɪg/ egg /ɛg/ P3	

Figure 10-9: Plosive posters P1-P3

10.4.4 Teacher's manual: Fricatives

Teachers' Manual: Fricatives /s/ /z/ /θ/ /ð/

1. Introduction

German speakers usually don't have any problems with pronouncing the /s/ or /z/ correctly but German speakers of English often pronounce the /s/ especially at the beginnings of words as voiced, because in the German language an /s/ at word beginnings is always voiced. So, for example, the English name /su:zn/ or /su:zi/ is often pronounced as /zu:zn/ or /zu:zi/ by German speakers of English. (Auch oft "sechs Filme" vs. "Sexfilme".)[1]

/θ/ and /ð/ (thin and mother) don't occur in the German language. Respectively, the most common substitutions are /s/ and /z/. [2]

a) Task F1: Introduction to /s/ /z/.

Time		Media
5min	T. puts the transparency F1 on the OHP and shows the first picture to the students. T. asks the students to put their fingers on their voice box and make the sound of a snake and then of a bee. After establishing the /s/ and the /z/ in class, ask the learners to alternate the sounds without stopping: /s…s….z….z…z…s..s../	Transparency F1
	Students should touch their voice box and determine which of the sounds is voiced and which voiceless.	
	T. practises and drills the example on the transparency F1 over-exaggerating the voiced and unvoiced "s2", e.g. zzzzzzZoo, sssssSue, etc.	Transparency F1

a) Task F2: Introduction to /θ/ /ð/

Time		Media
5min	T. writes /θ/ and /ð/ on the board and asks the students whether they already know these phonetic symbols. T. makes sure everybody knows that they represent the letters "th".	Blackboard th→/θ/ th →/ð/
	T. asks students what the do with their lips, mouth, etc when they pronounce the "th" and collects ideas.	
	T. explains how to do the "th". (Even if the students already know how to speak this sound, the over-exaggerated position is a useful and helpful reminder!) Place one finger against your lips. Try to touch your finger with your tongue. Breathe out. Now add your voice. [3]	

[1] (Quetz 1998: 51)
[2] (Kenworthy 1987: 137)

(Kenworthy, 1987, p. 137; Quetz, 1998, p. 51)

2. Presentation

a) Task F3: Watching the video

The video is rather silly but truly helpful. Please encourage the students to speak along and to exaggerate their mouth movements as well.

Time		Media
11min.	T. watches the video with the students. Students should speak along whenever possible.	Video F1: /ɵ/ /ð/

b) Task F4

Because this task can be regarded as a repetition of the video it makes sense to do it right after watching the video if possible.

Time		Media
5-10min	T. puts the phonetic posters F2, F3 on the blackboard and asks the students to decide which one is the voiced/voiceless "th".	Phonetic Posters F2, F3
	T. drills the example words on the posters: /ɵ/ thank, think, maths, thin /ð/ that, than, them, paths Students should touch their voice box, so that they not only hear but feel the difference as well.	
	T. puts the Phonetic Poster F1 /z/on the blackboard and practises the example words: Close, does, zip, paths	Phonetic Poster F1
	Students should write one example for /θ/, /ð/, and /z/ in their own sound chart.	Phonetic Code Chart

3. Identification & Discrimination

a) Task F5:

Time		Media
10-15min	T. presents Task 1 on Worksheet F1 and reads the example sentences. Students should work in pairs and decide which of the "ths" are voiced or voiceless. 1. What do you think? 2. Where is the bathroom? 3. He is my brother. 4. Jonathan is bigger than 7. The paths (Wege) are very long. 8. That is great! 9. The baby is very thin.	Worksheet F1

[3](Kelly, 2000, p. 56)

	Marc. 5. We have maths on Thursdays. 6. Thank you very much! 10. He likes them. 11. My mother likes swimming.	
	T. corrects the worksheet (and practises some words again in case there are difficulties).	Worksheet F1 on a transparency
	words with /θ/ / **words with /ð/** think / brother bathroom / than Jonathan / the maths / paths thursday / that thank / them thin / mother	Solutions

4. Production

a) Task F7: Tongue Twisters.

Time		Media
5min to open end.	T. reads the tongue twister several times with the students. Students in groups should come up with a tongue twister of their own using all four sounds: /s/, /z/, /θ/, and /ð/	Transparency F2

3

Worksheet F1

/θ/ or /ð/?

Task 1: Which words have /θ/, and which words have /ð/?

1. What do you think?
2. Where is the bathroom?
3. He is my brother.
4. Jonathan is bigger than Marc.
5. We have maths on Thursdays.
6. Thank you very much!
7. The paths (Wege) are very long.
8. That is great!
9. The baby is very thin.
10. He likes them.
11. My mother likes swimming.

words with /θ/	words with /ð/
think	

Remember:

/θ/ is voiceless → there is no vibration of the voice box
/ð/ is voiced → there is vibration of the voice box.

The voiced /ð/ occurs commonly in relationship terms, such as father, brothers, whereas the voiceless /θ/ occurs frequently in personal names e.g. Matthew, Theo, etc.

Transparency F2

"th"- Tongue Twisters

The original transparency in the teacher's manual included tongue twisters adapted from Language Avenue (2010).

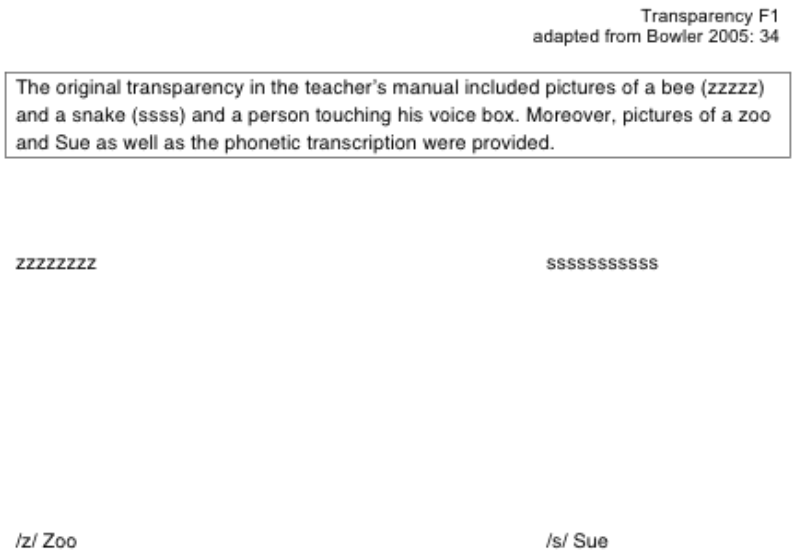
Susan and Theo thanked …

I thought a thought. But the thought I…

Is this …

(Language Avenue 2010)

(adapted from Language Avenue, 2009)



/z/		/θ/		/ð/	
close	/kləʊz/	thank	/θæŋk/	that	/ðæt/
does	/dʌz/	thing	/θɪŋ/	than	/ðən/
zip	/zɪp/	maths	/mæθs/	them	/ðem/
jeans	/dʒiːnz/	thin	/θɪn/	paths	/pɑːðz/
	P1		P2		P3

Figure 10-10: Fricative posters F1-F3

10.4.5 Teacher's manual: Approximant

Teachers' Manual: w/v

1. Introduction

The sound /w/ doesn't equal its German counterpart. Although a mispronunciation of this consonant usually doesn't lead to misunderstandings the German use of the /w/ is often ridiculed in comics, such as "Very vell".[1]

The English /w/ is pronounced like /u:/, e.g. "will" → "u:-ill".

A lot of German speakers have problems with the hearing of the /w/. They might confuse it with the /v/ or /f/. The letter "v" can also cause problems, because in English it is pronounced as /v/ but not as /f/ as in the German language[2].

a) Task W1

Time		Media
10min	T. writes "wet" and "vet" on the blackboard. T asks the students how to pronounce the two words and what the differences are between the words. T. explains (see below) and **models** how to pronounce the /v/ and /w/ and practises the two words. / v/ /v/ is a voiced consonant. Bite your bottom lip with your top teeth and push air out to make this sound. /w/ /w/ is the short form of the long vowel /u:/. Put a pencil in your mouth and put your lips round it. Take the pencil out and keep your lips round. Push air out to make this sound[3]	<u>Vocabulary:</u> Vet (n.) → Tierarzt
	Students should produce the two words "wet" and "vet" silently while watching their mouths in the mirror. They should describe (and discuss) their mouth, tongue, and teeth movements to (with) their partner. T.: "Work in pairs. One student will silently produce a word. The other student has to guess which word." (and vice versa")	<u>Mirrors</u>

[1] (Quetz, p. 50f).
[2] (Hooke, p. 244.)
[3] (Bowler 2005: 55)

1

(adapted from Bowler, 2005, p. 55; Hooke & Rowell, 1982, p. 244; Quetz, 1998, p. 50f.)

2. Presentation

a) Task W2:

Time		Media
3min	Student watch the video and speak along as often as possible.	Video W1 2min 10s

b) Task W3

Time		Media
5min	T. writes the word "one" on the blackboard. T. makes clear that although there is no "w" in one it is pronounced with the /w/. T. asks whether the students know similar words and collects them on the blackboard.	Blackboard
	T. puts the poster W1 on the board and drills the example words: walk, one, swing, sweet	Vocabulary: sweet (adj.) → süß Poster W1

3. Identification & Discrimination

a) Task W4:

The students are likely not to know the vocabulary in this task. It shouldn't matter too much as the pictures make the meaning quite clear and may even help them to focus on the minimal pair /v/ and /w/.

Time		Media
6min	Task1 T. : Listen to the pairs of words. In the words on the left the first sound is /v/. In the words on the right the first sound is /w/. Can you hear the difference? T.: Listen again. Speak along with the CD.	Worksheet W1 CD Track W1 (1:06min) /w/ and /v/
	Task 2 T. reads some words from task 1 at random and asks students to listen for the correct word.	Worksheet W1

4. Production

a) Task W6

The instructions of this game may sound a bit difficult in the beginning, but once understood it is very easy and definitely worth playing the game.

Time		Media
20min	T. hands out worksheet W2 "Streetmap of Letterton" to the students. **N.B.: There is partner A and partner B** T. puts the transparency of the "Street Map of Letterton" on the OHP and explains the rules.	Worksheet W2/ Transparency Street Map of Letterton Partner A /PartnerB

b) Task W7: Tongue twisters

This is a fun activity but it if you don't have enough time it can easily be omitted.

Time		Media
	T. reads the tongue twisters on the transparency and asks students to repeat them in class. Then the students should come up with a tongue twister of their own. The class decides on the best/funniest, most interesting tongue twister.[4] Tongue Twisters Very well, Walter, You're welcome, Viviane What a very nice video! Vera's Wonderful Wedding Videos Willy Wellwood visits Vollowoolo in winter William Winter works for a video company	Transparency W3

[4] (Bowen & Marks, 1992, p. 31; Quetz, 1998, p. 42)

Worksheet W1[5]

Task 1: Listen to the pairs of words. In the words on the left the first sound is /v/. In the words on the right the first sound is /w/. Can you hear the difference?

In the original manual pictures of the following sounds were provided:

/v/ veils	/w/ wales	/v/ veal	/w/ wheel
/v/ vest	/w/ west	/v/ vine	/w/ wine
/v/ V	/w/ we	/v/ vet	/w/ wet

It doesn't matter if you don't understand all the vocabulary. Just try to hear the difference between /w/ and /v/.

Task 2: Listen to your teacher. Which word does she say?

Remember how to pronounce /w/ and /v/?

/v/
/v/ is a voiced consonant. Bite your bottom lip with your top teeth and push air out to make this sound.

/w/
/w/ is the short form of the long vowel /u:/. Put a pencil in your mouth and put your lips round it. Take the pencil out and keep your lips round. Push air out to make this sound. Example "wet" → "u:-et"

[5] Adapted from Bowler 2005: 55

Tongue Twisters:

Very well, Walter,

You're welcome, Viviane

What a very nice video!

Vera's Wonderful Wedding Videos

Willy Wellwood visits Vollowoolo in winter

William Winter works for a video company

Transparency W3

Worksheet W2
Adapted from Bowler 2005: 56

Street Map of Letterton[6] - Student A

Read the instructions and play the game. Work in pairs.

In the orginal manual a game similar to "battleship" is displayed. However, instead of finding ships, the students have to find houses in a grid of streets.
In order to find the correct house, the students have to correctly produce and identify the difference between /v/ and /w/ as the street names include minimal pairs such as "West Road" vs. "Vest Road".

/w/			
walk	/wʌn/		
one	/wɔ:k/		
swing	/swɪŋ/		
sweet	/swi:t/		
	W1		

Figure 10-11: Poster Approximant W1

10.4.6 Teacher's manual: Phonetic code chart

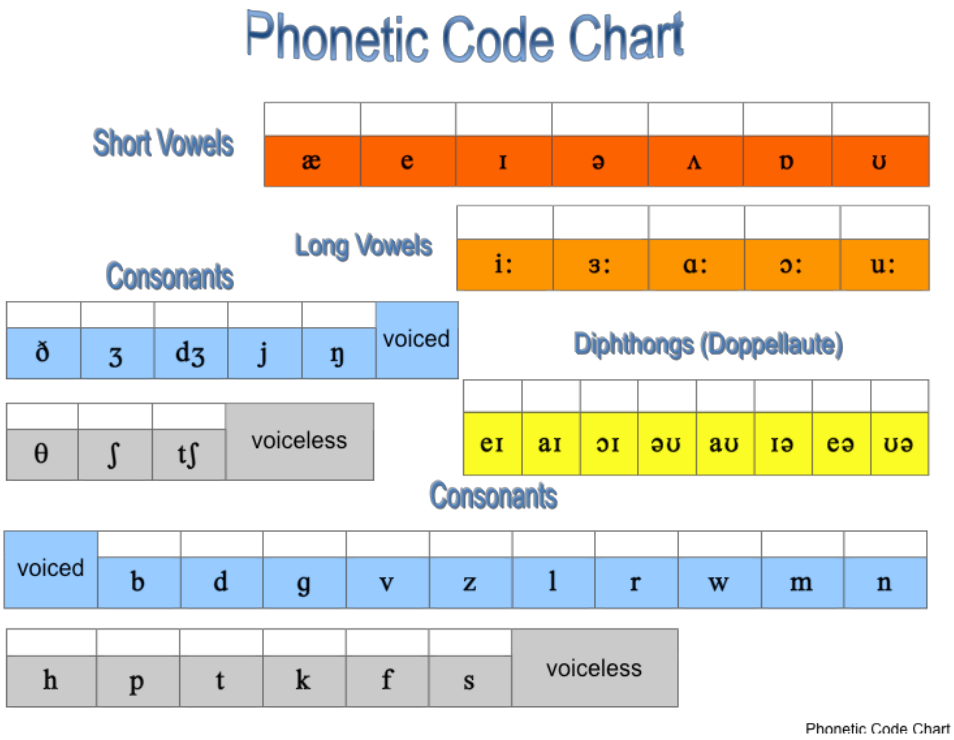


Figure 10-12: Phonetic code chart

10.4.7 Teacher's manual: Video files

The video files can be found online (see References).

- V1_Long_Vowel_Sound_a.mp4 (BBC, 2010)
- V2_Long_Vowel_ɔː_04.mp4 (BBC, 2010)
- V3_Short_Vowel_Sound_æ.mp4 (BBC, 2010)
- V4_Short_Vowel_Sound_schwa.mp4 (BBC, 2010)
- D1_When 2 vowels go walking.mp4 (Between the Lions, 2009)
- F1_th.mp4 (Sconda, 2008)
- W1_w_v.mp4 (BBC, 2010)

10.4.8 Teacher's manual: Audio files

Due to copyright reasons the audio files couldn't be included in the thesis.

10.5 Results

Table 10-6: Mean approximant-vowel durations (in ms) and standard error by group, time and test item

	test item	Intervention		Control	
		Mean approximant-vowel duration (ms)	SE	Mean approximant-vowel duration (ms)	SE
pre-test	one	235.8	21.0	242.7	16.9
post-test	one	258.8	25.2	222.1	18.2
follow-up-test	one	251.5	20.6	231.3	22.1
pre-test	swing	187.1	11.3	198.8	17.7
post-test	swing	207.1	15.7	191.0	15.5
follow-up-test	swing	217.9	17.2	198.5	11.0
pre-test	wait	317.1	23.6	300.8	17.7
post-test	wait	345.4	37.7	286.7	20.5
follow-up-test	wait	318.8	23.8	295.0	11.6
pre-test	witch	160.0	11.1	137.1	8.3
post-test	witch	165.6	9.3	137.5	11.6
follow-up-test	witch	162.9	18.1	153.5	13.3
pre-test	total	225.0	34.5	220.0	34.7
post-test	total	244.3	38.6	209.3	31.3
follow-up-test	total	238.0	32.6	219.8	29.6

Table 10-7: Mean pre-plosive vowel durations by time group and plosive

	plosive	Intervention		Control	
		Mean vowel duration (ms)	SE	Mean vowel duration (ms)	SE
T_0 pre-test	/b/	182.4	11.8	168.6	9.6
	/d/	280.5	15.3	251.5	14.3
	/g/	216.8	14.0	188.9	11.0
T_1 post-test	/b/	215.0	13.4	174.0	10.4
	/d/	328.7	19.2	238.5	13.7
	/g/	242.1	14.9	187.8	9.9
T_2 follow-up-test	/b/	192.2	12.4	181.1	10.2
	/d/	307.6	16.3	269.3	13.6
	/g/	222.3	12.4	210.0	9.3

REFERENCES

Archibald, J., & Young-Scholten, M. (2003). The Second Language Segment Revisited. *Second Language Research*, *19*(3), 163–167. doi:10.1191/0267658303sr218ed

Asher, J. J., & Garcia, R. (1969). The Optimal Age to Learn a Foreign Language. *The Modern Language Journal*, *53*(5), 334. doi:10.2307/323026

Avery, P. W., & Ehrlich, S. (2002). *Teaching American English pronunciation*. Oxford: Oxford Univ. Press.

Banse, R., & Scherer, K. R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636.

Bauer, J. (2008). *Warum ich fühle, was du fühlst : intuitive Kommunikation und das Geheimnis der Spiegelneurone*. München: Heyne.

BBC. (2010). BBC Learning English | Pronunciation Tips. Retrieved September 25, 2013, from http://www.bbc.co.uk/worldservice/learningenglish/grammar/pron/sounds/

Best, C. T. (1993). *Emergence of language-specific constraints in perception of non-native speech: a window on early phonological development.* Dordrecht: Kluwer Academic Publishers. Retrieved from http://www.loc.gov/catdir/enhancements/fy0821/93009598-d.html

Best, C. T. (1995). A direct realist view of cross-language speech perception: New Directions in Research and Theory. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–204). Baltimore: York Press.

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Language Experience in Second Language Speech Learning. In Honor of James Emil Flege (pp. 13–34). https://doi.org/10.1075/lllt.17.07bes

Between the Lions. (2009). When Two Vowels Go Walking. Retrieved January 28, 2015, from https://www.youtube.com/watch?v=7fb3Pdt8kxg

Birdsong, D. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. Birdsong (Ed.), *Second Language Acquisition Research: Theory-Construction and Testing* (pp. 1–22). Mahwah N.J.: Laurence Erlbaum Associates.

Bishop, J., & Keating, P. (2012). Perception of pitch location within a speaker's range: Fundamental frequency, voice quality and speaker sex. *The Journal of the Acoustical Society of America*, *132*(2), 1100. doi:10.1121/1.4714351

Boersma, P., & Weenik, D. (2012). Praat: doing Phonetics by Computer. Retrieved January 25, 2011, from http://www.fon.hum.uva.nl/praat/

Bohn, O.-S., & Flege, J. E. (1992). The Production of New and Similar Vowels by Adult German Learners of English. *Studies in Second Language Acquisition*, *14*(02), 131–158. doi:10.1017/S0272263100010792

Bongaerts, T., Summeren, C. van, Planken, B., & Schils, E. (1997). Age and Ultimate Attainment in the Pronunciation of a Foreign Language. *Studies in Second Language Acquisition*, *19*(4), 447–465.

Bosch, L., Costa, A., & Sebastián-Gallés, N. (2000). First and second language vowel perception in early bilinguals. *European Journal of Cognitive Psychology*, *12*(2), 189–221.

Bowen, T., & Marks, J. (1992). *The Pronunciation Book: Student-centred Activities for Pronunciation Work*. London: Longman.

Bowler, B. (2005). *Pronunciation activities: elementary-intermediate*. London: Mary Glasgow.

Bowyer, T. (2006). Howjsay.com. Retrieved June 9, 2014, from http://www.howjsay.com/notes.html

Brown, H. D. (1980). The Optimal Distance Model of Second Language Acquisition. *TESOL Quarterly*, *14*(2), 157–164.

Cardoso, Walcir. 2010. 'Teaching Foreign SC Onset Clusters: Comparing the Effects of Three Types of Instruction'. Pp. 61–66 in Selected Papers from the Proceedings of the 14th International Conference of the Greek Applied Linguistics Association. GALA.

Carr, P. (2008). *A glossary of phonology*. Edinburgh: Edinburgh University Press.

Catford, J. C. (1988). *A practical introduction to phonetics*. Oxford: Clarendon.

Celce-Murcia, M. (1997). *Teaching Pronunciation* (Vol. Buch). Cambridge: Cambridge Univ. Press.

Chen, Hsuan-Yu, and Jaya Goswami. 2011. 'Structuring Cooperative Learning in Teaching English Pronunciation'. English Language Teaching 4(3).

Cholin, J. (2008). The mental syllabary in speech production: An integration of different approaches and domains. *Aphasiology*, *22*(11), 1127–1141. doi:10.1080/02687030701820352

Chujo, K. (2004). Measuring Vocabulary Levels of English Textbooks and Tests Using a BNC Lemmatised High Frequency Word List. *Language and Computers*, *51*, 231–249.

Clark, J. (2007). *An introduction to phonetics and phonology* (3rd ed). Malden, MA ; Oxford: Blackwell Pub.

Cleary, M., & Pisoni, D. B. (2001). Speech Perception and Spoken Word Recognition: Research and Theory. In E. B. Goldstein (Ed.), *Blackwell handbook of perception* (pp. 499–534). Oxford, UK; Malden, Mass., USA: Blackwell.

Council of Europe. (2001). The Common European Framework of References for Languages (CEFR). Retrieved March 31, 2012, from http://www.coe.int/t/dg4/linguistic/cadre_en.asp

Counselman, D. (2010). *Improving pronunciation instruction in the second language classroom*. The Pennsylvania State University. Retrieved from https://etda.libraries.psu.edu/paper/11030/6503

Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: socially constructed metalanguage and critical listening. *Language Awareness*, *20*(3), 159–182. https://doi.org/10.1080/09658416.2011.570347

Cox, F. (2008). Consonant Acoustics: Fricatives. Retrieved February 24, 2015, from http://clas.mq.edu.au/speech/acoustics/consonants/fricweb.html

Cruttenden, A. (2008). *Gimson's Pronunciation of English*. London: Hodder Education.

Cruz-Ferreira, M. (2009). Multilingual Accents. *Speak Out! The Newsletter of the IATEFL Pronunciation Special Interest Group*, (40), 22–26.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford Univ. Press.

Darcy, Isabelle, Doreen Ewert, and Ryan Lidster. 2012. 'Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox". In Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference, edited by J. Levis and K. LeVelle.

DeCoster, J. (2001). Psychological Research Methods. Retrieved from http://www.stat-help.com/notes.html

DeKeyser, R. (2012). Age effects in second langauge learning. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 442–460). London; New York, NY: Routledge.

Denes, P. B., & Pinson, E. N. (1993). *The speech Chain: The Physics and Biology of Spoken Language*. New York: W. H. Freeman and Company.

Derwing, T. M., & Munro, M. J. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Derwing, T. M., & Rossiter, M. J. (2003). The Effects of Pronunciation Instruction on the Accuracy, Fluency, and Complexity of L2 Accented Speech. *Applied Language Learning*, *13*(1), 1–17.

Derwing, Tracey M., and Murray J. Munro. 2015. Pronunciation Fundamentals: Evidence-Based Perspectives for L2 Teaching and Research. Amsterdam ; Philadelphia: John Benjamins Publishing Company.

Di Pellegrino, Giuseppe, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, and Giacomo Rizzolatti. 1992. 'Understanding Motor Events: A Neurophysiological Study'. Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale 91:176–80.

Eckert, H., & Barry, W. (2005). *The phonetics and phonology of English pronunciation*. Trier: WVT, Wiss. Verl. Trier.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*, 315–330.

Eckman, F. (2008). Typological markedness and second language phonology. *Phonology and Second Language Acquisition, Philadelphia: John Benjamins*, 1–31.

Eckman, F. (2012). Second language phonology. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 91–105). London; New York, NY: Routledge.

Edelhoff, C. (2007b). *Notting Hill Gate* (Dr. A,1). Braunschweig: Diesterweg.

Edelhoff, C. (Ed.). (2007a). *Für Klasse 5: Unterrichtsvorschläge mit Lösungen.* Braunschweig: Diesterweg.

Edmondson, W. (1999). *Twelve Lectures on Second Language Acquisition: Foreign Language Teaching and Learning Perspectives*. Tübingen: Narr Verlag.

Elliott, A. R. (1995). Field Independence/Dependence, Hemispheric Specialization, and Attitude in Relation to Pronunciation Accuracy in Spanish as a Foreign Language. *The Modern Language Journal*, *79*(3), 356–371.

Elliott, A. R. (1995). Field Independence/Dependence, Hemispheric Specialization, and Attitude in Relation to Pronunciation Accuracy in Spanish as a Foreign Language. *The Modern Language Journal*, *79*(3), 356–371.

Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton: The Hague.

Fehling, S. (2008). *Language Awareness und bilingualer Unterricht : eine komparative Studie* (2., überarb. Aufl.). Frankfurt am Main: Lang.

Feilding, C. (2013). Sound source in singing. Retrieved April 18, 2014, from http://www.google.de/imgres?imgurl=http%3A%2F%2Fwww.feilding.net%2Fsfuad%2Fmusi3012-01%2Fimages%2Flectures%2Fvocal_fold_cycles.gif&imgrefurl=http%3A%2F%2Fwww.feilding.net%2Fsfuad%2Fmusi3012-01%2Fhtml%2Flectures%2F017_instruments_IV.htm&h=231&w=614&tbnid=3LCyJ_M8F3uL0M%3A&zoom=1&docid=0u-f5hhYnjjKaM&ei=aDdRU_LJKcOytAbj04GgBA&tbm=isch&client=firefox-a&iact=rc&uact=3&dur=4217&page=1&start=0&ndsp=42&ved=0CI4BEK0DMBI

Field, A. P. (2009). *Discovering statistics using SPSS: (and sex, drugs and rock "n" roll)* (3rd ed). Los Angeles: SAGE Publications.

Finkbeiner, C. (1996). Zugänge und Grundlagen zur Erforschung des Fremdsprachenunterrichts. *Fremdsprachen Und Hochschule*, (48), 36–66.

Flege, J. E. (1981). The Phonological Basis of Foreign Accent: A Hypothesis. *TESOL Quarterly*, *15*(4), 443–455.

Flege, J. E. (1987). The Production of "New" and "Similar" Phones in a Foreign Language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics*, *15*(1), 47–65.

Flege, J. E. (1995). Speech perception and linguistic experience:Theoretical and methodological issues. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233–277). Baltimore: York Press.

Flege, J. E. (1997). English vowel productions by Dutch talkers: more evidence for the "similar" vs "new" distinction. In A. R. James & J. Leather (Eds.), *Second-Language Speech: Structure and Process* (1st ed., pp. 11–52). Berlin: Mouton de Gruyter.

Flege, J. E., & Fletcher, K. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustic Society of America*, *91*(1), 370–389.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, *25*(2), 169–186. doi:10.1006/jpho.1996.0040

Flege, J. E., Munro, M. J., & Mackay, I. R. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, *16*(1), 1–26.

Flege, J. E., Schirru, C., & Mackay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, *40*(4), 467–491.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, *41*(1), 78–104. doi:10.1006/jmla.1999.2638

Fogassi, Leonardo, and Pier Francesco Ferrari. 2009. 'Mirror Neurons, Gestures and Language Evolution'. Pp. 29–45 in 119304, Benjamins Current Topics: 13, edited by C. Abry, A. Vilain, and J.-L. Schwartz. John Benjamins Publishing Company.

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, *84*(1), 115–123.

Fowler, C. A., & Dekle, D. J. (1991). Listening with eye and hand: Cross-modal contributions to speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *17*(3), 816–828. doi:10.1037/0096-1523.17.3.816

Fox, A. V. (2000). *The acquisition of phonology and the classification of speech disorders in German-speaking children* (unpublished). Newcastle upon Tyne.

Garcia, N. (2005). *Foreign Language Phonology: Field Independence, Attitude, and the Success of Formal Instruction in Spanish Pronunciation*. Retrieved from http://d-scholarship.pitt.edu/9008/1/GarciaMAthesis2005.pdf

Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Gnutzmann, C. (1999). *Teaching and learning English as a global language : native and non-native perspectives*. Tübingen: Stauffenburg.

Goethe Gymnasium. (2014). Goethe-Gymnasium Frankfurt am Main. Retrieved June 5, 2014, from http://www.gg-ffm.de/

Gopnik, A., Kuhl, P., & Meltzoff, A. (2001). *Forschergeist in Windeln. Wie Ihr Kind die Welt begreift* (2nd ed.). München: Hugendubel.

Götz, S. (2011). *Fluency in Native and Nonnative English Speech: Theory, description, implications* (unpublished). Justus-Liebig University, Giessen & Macquarie University Sydney.

Grotjahn, R. (1998). Ausspracheunterricht: Ausgewählte Befunde aus der Grundlagenforschung und didaktisch-methodische Implikationen, *9*(1), 35–83.

Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., Dull, C. Y., & Scovel, T. (1972). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry*, *13*(5), 421–428. doi:10.1016/0010-440X(72)90083-1

Gut, U. (2009). *Non-native speech : a corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt am Main; Oxford: Peter Lang.

Halibisky, B. (1999). Measuring Dis/Similarities Between Objects (Cells) In `n´- Dimensional Space. Retrieved August 20, 2014, from http://hlab.stanford.edu/~brian/making_measurements.html

Hardison, D. M. (2012). Second language speech perception: a cross-disciplinary perspective on challenges and accomplishements. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 349–363). London; New York, NY: Routledge.

Harrington, J. (2012). Acoustic Phonetics. *The Handbook of Phonetic Sciences*, *79*, 81.

Harrington, J. (2013). Acoustic Phonetics. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (pp. 81–129). Chichester, U.K.; Malden, MA: Wiley-Blackwell. Retrieved from http://dx.doi.org/10.1002/9781444317251

Haß, F. (2006). *Fachdidaktik Englisch. Tradition - Innovation - Praxis (Lernmaterialien) mit DVD-ROM* (1. ed.). Stuttgart: Klett.

Haskins Laboratories. (2008). Measuring and Modeling Speech Production. Retrieved April 18, 2014, from http://www.haskins.yale.edu/featured/heads/mmsp/intro.html

Hayes-Harb, R. (2007). Lexical and Statistical Evidence in the Acquisition of Second Language Phonemes. *Second Language Research*, *23*(1), 65–94. doi:10.1177/0267658307071601

Hayward, K. (2000). *Experimental phonetics*. Harlow, Eng.; New York: Longman.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.

Herbst, T. (1992). Pro-Nunciation: Zur Bedeutung einer guten Aussprache in der Fremdsprache. *Die Neueren Sprachen*, *91*(1), 2–18.

Herrmann, F. (2010). *Syllable Frequency and Speaker Sex Effects on Coarticulation: An Acoustic Investigation*. University of Sheffield. Retrieved from about:blank

Hessisches Kultusministerium. (2010a). Landesportal Hessen - G8. Retrieved November 19, 2011, from http://www.hessen.de/irj/hessen_Internet?cid=43b5544fa17727e549121576c0afe7fc

Hessisches Kultusministerium. (2010b). Lehrplan Englisch (G8) Gymnasialer Bildungsgang Jahrgangsstufen 5G bis 9G und gymnasiale Oberstufe. Retrieved from http://www.hessen.de/irj/HKM_Internet?cid=ac9f301df54d1fbfab83dd3a6449af60

Hessisches Kultusministerium. (2011a). Bildungsstandards und Inhaltsfelder. Das neue Kerncurriculum für Hessen. Sekundarstufe I - Gymnasium. Moderne Fremdsprachen. Retrieved March 29, 2012, from http://www.iq.hessen.de/irj/IQ_Internet?uid=44540e7a-7f32-7821-f012-f31e2389e481

Hessisches Kultusministerium. (2011b). Handreichung zur Arbeit mit den Lehrplänen Lehrplänen der Bildungsgänge Hauptschule, Realschule und Gymnasium ENGLISCH an schulformübergreifenden (integrierten) Gesamtschulen und Förderstufen. Retrieved from http://www.hessen.de/irj/HKM_Internet?cid=e9791809d4e92c28a030b4f9340b0d5e

Hessisches Kultusministerium. (2011c). Lehrplan Englisch Bildungsgang Hauptschule Jahrgangsstufen 5 bis 9/10. Retrieved from http://www.hessen.de/irj/HKM_Internet?cid=770244b3f3f61faf79f08f0f0db32a30

Hessisches Kultusministerium. (2011d). Lehrplan Englisch Bildungsgang Realschule Jahrgangsstufen 5 bis 10. Retrieved November 20, 2011, from http://www.hessen.de/irj/HKM_Internet?cid=f1e079cc428af80d07f4fe2db20fe301

Hillenbrand, James, Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *J. Acoustical Society of America*, *97*(5), 3099–3111.

Hillenbrand, James. (2012). Alvin experiment-control software. Retrieved January 25, 2011, from http://homepages.wmich.edu/~hillenbr/

Hincks, R., & Edlund, J. (2009). Promoting increased pitch variation in oral presentations with transient visual feedback. *Language Learning & Technology*, *13*(3), 32–50.

Hirschfeld, U. (1997). Welche Aussprache lehren wir? In A. Wierlacher (Ed.), *Jahrbuch Deutsch als Fremdsprache : intercultural German studies* (Vol. 23, pp. 175–188). München: Iudicium.

Hirschfeld, U. (2007). In K.-R. Bausch, H. Christ, & H.-J. Krumm (Eds.), *Handbuch Fremdsprachenunterricht* (5., unveränderte Auflage.). UTB, Stuttgart.

Hooke, R., & Rowell, J. (1982). *A handbook of English pronunciation: [featuring notes in French, German, Greek, Italian, Portugese, Spanish]*. London: Edward Arnold.

Huthaily, K. Y. (2008). *Second language instruction with phonological knowledge: Teaching Arabic to speakers of English*. ProQuest. Retrieved from https://books.google.com/books?hl=en&lr=&id=EJNLdtlA6jQC&oi=fnd&pg=PA1&dq=%22For+the+most+part,+students+in+the+experimental+group+stated+that+it+is+important%22+%22production+and+sound+recognition+skills+of+adults+learning+a+second%22+%22results+of+this+study+strongly+suggest+that+including+an+introductory+component%22+&ots=etvxQA7La9&sig=-G1L7hlbFlSHutTDmFDuXS0zBpg

IBM Corp. (2013). IBM SPSS Statistics (Version 21). Armonk, NY.

Ingels, S. A. (2011). *The Effects of Self-Monitoring Strategy Use on the Pronunciation of Learners of English*. University of Montana.

International Phonetic Association. (1999). *Handbook of the International Phonetic Association : a guide to the use of the International Phonetic Alphabet*. Cambridge, U.K.; New York, NY: Cambridge University Press.

Isaacs, T. (2008). Towards Defining a Valid Assessment Criterion of Pronunciation Proficency in Non-Native English Speaking Graduate Students. *The Canadian Modern Language Review*, (64), 555–580.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y. I., Kettermann, A., & Siebert, Claudia. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*(1), B47–B57.

Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *Journal of the Acoustical Society of America*, *126*(5), 2603–2618.

Jenkins, J. (2000). *The phonology of English as an international language : new models, new norms, new goals*. Oxford: Oxford University Press.

Jescheniak, J. D., & Levelt, W. J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824.

Johnson, K. (2005). *Acoustic and auditory phonetics*. Malden, Mass.: Blackwell. Retrieved from http://www.gbv.de/dms/bs/toc/354883518.pdf

Johnson, K. (2013). Linguistics 110 Introduction to Phonetics and Phonology. Retrieved April 18, 2014, from http://linguistics.berkeley.edu/~kjohnson/ling110/syllabus13.html

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the World: Teaching and Learning the Language and Literatures* (1st ed., pp. 11–36). Cambridge: Cambridge University Press.

Keller, E. (1987). The cortical representations of motor processes in speech. In E. Keller & M. Gopnik (Eds.), *Motor and Sensory Processes of Languages*. Hillsdale, New Jersey: Laurence Erlbaum Associates.

Kelly, G. (2000). *How To Teach Pronunciation* (Pap/Com). Essex: Longman.

Kennedy, S. (2003). Oral peer corrective feedback in an ESL classroom: Training Quebec francophone learners in the prounciation of 'th'. unpublished Master Thesis. Retrieved from http://spectrum.library.concordia.ca/2145/1/MQ77634.pdf

Kent, R. D. (2000). RESEARCH ON SPEECH MOTOR CONTROL AND ITS DISORDERS: A REVIEW AND PROSPECTIVE. *Journal of Communication Disorders*, *33*, 391–428.

Kenworthy, J. (1987). *Teaching English Pronunciation*. Longman.

Kesselmann, H., Rogan, J. C., Medoza, J. L., & Breen, L. L. (1980). Testing the validity conditions of repeated measures F tests. *Psychological Bulletin*, (87), 479–481.

Kieweg, W. (1998). Lernprozessorientierte Kriterien zur Evaluierung von Englisch-Lehrwerken. *Der Fremdsprachliche Unterricht Englisch*, (4), 27–38.

Kilgarriff, A. (1997). Putting frequencies in the dictionary. *Int J Lexicography*, *10*(2), 135–155.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., … Vollmer, H. J. (2004). The Development of the National Educational Standards. Bundesministerium für Bildung und Forschung (BMBF).

König, E., & Gast, V. (2007). *Understanding English-German contrasts*. Berlin: Schmidt. Retrieved from http://katalog.ub.uni-heidelberg.de/titel/66408843

Krashen, S. D., Long, M. A., & Scarcella, Robin. (1979). Age, Rate and Eventual Attainment in Second Language Acquisition. *TESOL Quaterly*, (13), 573–582.

Kuhl, P. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 11850–11857.

Kuhl, P., & Iverson, P. (1995). Linguistic Experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 121–154). Baltimore: York Press.

Ladefoged, P. (1996). *Elements of acoustic phonetics* (2nd ed). Chicago: University of Chicago Press.

Ladefoged, P. (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell Pub.

Ladefoged, P. (2006). *A course in phonetics*. Southbank, Vic.: [Thomson Wadsworth].

Ladefoged, P. (2007). *Vowels and consonants*. Malden, Mass. [u.a.]: Blackwell.

Lado, R. (1957). *Linguistics across cultures : applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.

Language Avenue. (2009). Sound TH. Retrieved January 28, 2015, from http://languageavenue.com/teaching-ideas/english-tongue-twisters/sound-th

Leech, G. (2001). *Word frequencies in written and spoken English*. Harlow: Longman.

Lenneberg, E. H. (1967). *Biological Foundations of Language* (1st Edition). New York, Wiley.

Liew, Sook-Lei, and Lisa Aziz-Zadeh. 2013. 'The Human Mirror Neuron System, Social Control, and Language'. Pp. 183–205 in.

Lima, Ronaldo Mangueira. 2010. 'Pronunciation for Communication: An Investigation of the Effects of Explicit and Selected Teaching of Pronunciation in the Brazilian EFL Classroom'. Pp. 299–304 in Selected Papers from the Proceedings of the 14th International Conference of the Greek Applied Linguistics Association. GALA.

Limesurvey. (2012). LimeSurvey. Retrieved February 2, 2011, from http://limesurvey.uni-frankfurt.de/limesurvey/admin/admin.php?sid=48615

Lindblom, B. (1996). Role of articulation in speech perception: Clues from production. *Journal of the Acoustical Society of America*, *99*, 1683–1692.

Liu, Q., & Fu, Z. (2011). The Combined Effect of Instruction and Monitor in Improving Pronunciation of Potential English Teachers. *English Language Teaching*, *4*(3). https://doi.org/10.5539/elt.v4n3p164

Long, M. H. (1990). Maturational Constraints on Language Development. *Studies in Second Language Acquisition*, *12*(3), 251–285.

Lund Research Ltd. (2013). Leard Statistics. Retrieved August 7, 2014, from https://statistics.laerd.com/premium/ma/mixed-anova-in-spss.php

Machač, P., & Skarnitzl, R. (2009). *Principles of phonetic segmentation* (1st ed.). Prague: Epocha.

Mackey, A., & Gass, S. M. (2005). *Second language research : methodology and design*. Mahwah, NJ: Lawrence Erlbaum.

MacLeod, A., & Summerfield, Q. A. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiologyy*, *21*(2), 131–139.

Major, R. C. (2001). *Foreign accent : the ontogeny and phylogeny of second language phonology*. Mahwah, NJ [u.a.]: L. Erlbaum.

Mangold, M. (2005). *Das Aussprachewörterbuch: Duden* (Auflage: 6., überarbeitete und aktualisierte Auflage). Mannheim: Bibliographisches Institut.

Maxwell, S. E. (2003). *Designing Experiments and Analyzing Data: A Model Comparison Perspective* (Auflage: 2). Routledge.

Mayr, R. (2005). *The perception and production of German monophthongs by British learners of German*. University of Sheffield, Sheffield. Retrieved from http://robertmayr.weebly.com/publications.html

McGurk, H., & MacDonald, J. W. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748.

Mehlhorn, G. (2005). Learner Autonomy and Pronunciation Coaching. *Proceedings of Phonetics Teaching and Learning Conference, University College of London*. Retrieved from http://www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp13.pdf

Missaglia, F. (1999). Contrastive Prosody in SLA - an empirical study with adult Italian learners of German. In *Proceedings of the 14th International Congress of Phonetic Science* (Vol. 1, pp. 551–554).

Missaglia, F., & Sendelmeier, W. (1999). Die Realisierung deutscher Vokale durch italienische Muttersprachler: Eine experimentalphonetische Untersuchung., *10*(1), 73–95.

Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context and speaking style in English stressed vowels. *Journal of the Acoustical Society of America*, *96*, 40–55.

Morley, J. (1991). The Pronunciation Component in Teaching English to Speakers of Other Languages. *TESOL Quarterly*, *25*(3), 481–520.

Moyer, A. (1999). Ultimate Attainment in L2 Phonology: The Critical Factors of Age, Motivation, and Instruction. *Studies in Second Language Acquisition*, *21*(1), 81–108.

Moyer, A. (2007). Empirical Considerations on the Age Factor in L2 phonology. *Issues in Applied Linguistics*, 109–127.

Moyer, A. (2013). *Foreign accent: the phenomenon of non-native speech*. Cambridge: Cambridge University Press.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, *58*(3), 479–502.

Nagamine, T. (2011). Effects of hyper pronunciation training methodon Japanese Unviersity Students' pronunciation. *Asian EFL Journal Teaching Articles*, *53*, 35–50.

Neri, A., Cucchiarini, C., & Strik, H. (2006). Selecting segmental errors in non-native Dutch for optimal pronunciation training. *IRAL - International Review of Applied Linguistics in Language Teaching*, *44*(4), 357–404. doi:10.1515/IRAL.2006.016

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects, *4*, 272–281.

Oxford University Press. (2014). Oxford Dictionaries. Retrieved November 1, 2014, from http://www.oxforddictionaries.com/words/privacy-policy-and-legal-notice

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, *5*(3), 261–283. doi:10.1007/BF01067377

Park, H. (2009). *The effect of gender and length of English study on the pronunciation of certain English consonants: A case study of Korean EFL learners*. Southern Illinois University at Carbondale.

Patkowski, M. S. (1990). Age and Accent in a Second Language: A Reply to James Emil Flege. *Applied Linguistics*, *11*(1), 73–89. doi:10.1093/applin/11.1.73

Perkell, J. S. (2013). Five Decades of Research in Speech Motor Control: What Have We Learned, and Where Should We Go From Here? *Journal of Speech, Language, and Hearing Research*, *56*(6), 1857–1874. https://doi.org/10.1044/1092-4388(2013/12-0382)

Perlmutter, M. (1989). Intellibility rating of L2 speech pre an d post intervention. *Perceptual and Motor Skills*, *68*(2), 515–521. https://doi.org/10.2466/pms.1989.68.2.515

Pickering, L. (2012). Second language speech perception: a cross-disciplinary perspective on challenges and accomplishements. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 349–363). London; New York, NY: Routledge.

Pickett, (1999). *The acoustics of speech communication: fundamentals, speech perception theory, and technology*. Boston: Allyn and Bacon.

Pickett, (2001). *The acoustics of speech communication: fundamentals, speech perception theory, and technology*. Allyn and Bacon.

Pienemann, Manfred. 1984. 'Psychological Constraints on the Teachability of Languages*'. Studies in Second Language Acquisition 6(2):186–214.

Piepho, H.-E. (1974). *Kommunikative Kompetenz als übergeordnetes Lernziel im Englischunterricht*. Limburg: Frankonius-Verlag.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001a). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, *29*(2), 191–215. doi:10.1006/jpho.2001.0134

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001b). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, *29*(2), 191–215. doi:10.1006/jpho.2001.0134

Podesva, R., & Sharma, D. (Eds.). (2013). *Research methods in linguistics*. Cambridge ; New York: Cambridge University Press.

Prince, A., & Smolensky, P. (2002). *Optimality Theory Constraint Interaction in Generative Grammar* (First circulated: April, 1993 No. ROA Version: August, 2002) (pp. 1–256). Rutgers University Centre for Cognitive Science. Retrieved from roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF

Purcell, E. T., & Suter, R. W. (1980). Predictors of Pronunciation Accuracy: A Reexamination. *Language Learning*, *30*(2), 271–287.

Quetz, J. (1998). Aussprache und Hörwahrnehmung, Fernstudienprojekt "Fremdsprachen in Grund- und Hauptschulen." Universität Koblenz-Landau.

Richardson, B. (2008). Negative Attitudes to English Teachers' "Non-Standard" Accents and Solutions. *Speak Out! The Newsletter of the IATEFL Pronunciation Special Interest Group*, (39), 22–25.

Roach, P. (2009). A little encyclopaedia of phonetics. *University of Reading, UK*. Retrieved from http://www.freewebs.com/raulillox/Enciclopedia.pdf

Sakai, Mari, and Colleen Moorman. 2018. 'Can Perception Training Improve the Production of Second Language Phonemes? A Meta-Analytic Review of 25 Years of Perception Training Research'. Applied Psycholinguistics 39(01):187–224.

Sardegna, V. G. (2011). Pronunciation learning strategies that improve ESL learners' linking. In *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*.

Scarborough, R. (2005). Linguistics 105/205: Phonetics. Retrieved April 18, 2014, from http://www.stanford.edu/class/linguist205/

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, *9*(3), 235–248.
Schmidt, R. A. (1988). *Motor control and learning*. Champaign, Ill.: Human Kinetics Publ.
Schneider, G. (2014). Ernst-Reuter-Schule II. Retrieved June 5, 2014, from http://www.ersii.de/
Schuhmann, J. (1976). Social distance as a factor in second language acquisition. *Language Learning*, *26*(1), 135–143.
Schwarz, H. (2006). *English G 21 - Ausgabe A 1. 5. Schuljahr - Schülerbuch*. Cornelsen.
Schwarz, H. (Ed.). (2007). *English G 21 Handreichungen für den Unterricht.* Berlin: Cornelsen.
Sconda, D. (2008). ► "TH" Pronunciation English Meeting ESL Lesson. Retrieved January 28, 2015, from https://www.youtube.com/watch?v=Ag4qoNzEH4w
Scovel, T. (2000). A Critical Review of the Critical Period Research. *Annual Review of Applied Linguistics*, *20*, 213–223.
Selinker, L. (1972). Interlanguage. *IRAL*, *10*(3), 209–231.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (0002 ed.). Cengage Learning.
Singleton, D. (2001). Age and Second Language Acquisition. *Annual Review of Applied Linguistics*, *21*, 77–89. doi:10.1017/S0267190501000058
Smith, A. (2013). Development of neural control of orofacial movements of speech. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2. ed, pp. 251–296). Chichester, Malden, Mass: Wiley-Blackwell.
Smith, B. (2013). An acoustic analysis of voicing in American English dental fricatives. *OSUWPL*, *60*, 117–128.
Smith, B. (2013). An acoustic analysis of voicing in American English dental fricatives. *OSUWPL*, *60*, 117–128.
Smith, B. L., & Peterson, E. A. (2012). Native English speakers learning German as a second language: Devoicing of final voiced stop targets. *Journal of Phonetics*, *40*(1), 129–140. doi:10.1016/j.wocn.2011.09.004
Sobkowiak, W. (2009). Perceived Phonetic difficulty of English words. Retrieved March 4, 2012, from http://ifa.amu.edu.pl/~swlodek/diffind2.pdf
Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, *6*, 111–128.
Strange, W. (1995). *Speech perception and linguistic experience*. Baltimore: York Press.
Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, *26*(4), 311–344. doi:10.1006/jpho.1998.0078
Suarez, P. (2013). The pronunciation of the aspirated consonants p, t, and k in English by native speakers of Spanish and French. Retrieved from http://digitalcommons.fiu.edu/cgi/viewcontent.cgi?article=1133&context=sferc
Thompson, I. (1991). Foreign Accents Revisited: The English Pronunciation of Russian Immigrants. *Language Learning*, *41*(2), 177–204.
Thomson, R. I., and T. M. Derwing. 2014. 'The Effectiveness of L2 Pronunciation Instruction: A Narrative Review'. Applied Linguistics.
Thornbury, Scott. 2002. How to Teach Vocabulary. 13. impr. Harlow: Longman, Pearson Education.
Trofimovich, Pavel, Patsy M. Lightbown, Randall H. Halter, and Hyojin Song. 2009. 'COMPREHENSION-BASED PRACTICE'. Studies in Second Language Acquisition 31(04):609.
Trubetzkoy, N. S. (1969). *Principles of Phonology*. University of California Press.

Tsiartsioni, Eleni. 2010. 'The Effectiveness of Pronunciation Teaching to Greek State School Students'. Pp. 429–46 in Selected Papers from the Proceedings of the 14th International Conference of the Greek Applied Linguistics Association. GALA.

UCL. (2014). Vowels. Retrieved April 19, 2014, from http://www.phon.ucl.ac.uk/courses/spsci/iss/week5.php

Ultrax. (2012). MRI Schwa vowel - YouTube. Retrieved April 19, 2014, from http://www.youtube.com/watch?v=2BKA2sdr17Q

Varley, R., & Whiteside, S. (1998). Evidence of dual-route phonetic encoding from appraxia of speech: implications for phonetic encoding models. Presented at the The 5th International Conference on Spoken Language Processing, incorporating the 7th Australian International speech Science and Technology Conference, Sydney.

Varley, R., Whiteside, S., Windsor, F., & Fisher, H. (2006). Moving up from the segment: A comment on Aichert and Ziegler's Syllable frequency and syllable structure in apraxia of speech, Brain and Language, 88, 148–159, 2004. *Brain and Language*, *96*(2), 235–239. doi:10.1016/j.bandl.2005.04.008

Warsi, J. (2001). *Effects of Visual Feedback on Second Language Productive Phonology*. Retrieved from http://jilaniwarsi.tripod.com/first_comp.pdf

Weinberger, S. H. (1997). Minimal segements in second language phonology. In A. R. James & J. Leather (Eds.), *Second-Language Speech: Structure and Process* (1st ed., pp. 263–312). Berlin: Mouton de Gruyter.

Wellenreuther, M. (2000). *Quantitative Forschungsmethoden in der Erziehungswissenschaft*. Weinheim: Juventa-Verl.

Whiteside, S. P., Grobler, S., Windsor, F., & Varley, R. (2010). An acoustic study of vowels and coarticulation as a function of utterance type: A case of acquired apraxia of speech. *Journal of Neurolinguistics*, *23*(2), 145–161.

Whiteside, S., & Varley, R. (1998a). A reconceptualisation of apraxia of speech: a synthesis of evidence. *Cortex*, *34*, 221–231.

Whiteside, S., & Varley, R. A. (1998b). Coarticulation in apraxia of speech: an acoustic study of non-words. *Logopedics Phonatrics Vocology*, *23*(4), 155–163.

Wieden, W., & Nemser, W. (1991). *The pronunciation of English in Austria : a developmental and regional study*. Tübingen: G. Narr.

Wikimedia Commons. (2008a). *IPA vowel chart for Received Pronunciation diphthongs*. Retrieved from http://commons.wikimedia.org/wiki/File:RP_vowel_chart_(diphthongs).gif

Wikimedia Commons. (2008b). *IPA vowel chart for Received Pronunciation monophthongs*. Retrieved from http://commons.wikimedia.org/wiki/File:RP_vowel_chart_(monophthongs).gif

Wode, H. (1993). *Einführung in die Psycholinguistik. Theorien, Methoden, Ergebnisse*. Ismaning: Hueber Verlag.

Wrench, A. A. (1995). Analysis of fricatives using multiple centres of gravity. In *Proc. International Congress of Phonetic Sciences* (Vol. 4, pp. 460–463). Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.3.5666&rep=rep1&type=pdf

Wright, R. (2003). Factors of lexical competition in vowel articulation. In *Papers in Laboratory Phonology VI: Phonetic interpretation* (pp. 45–67). Cambridge: Cambridge University Press.

Young, R. F. (2011). English 333: Second Language Acquisition. Retrieved May 15, 2012, from http://www.english.wisc.edu/rfyoung/eng-333.fall2011.syllabus.html

ZFL. (2011, January 20). Zentrum für Lehrerbildung und Schul- und Unterrichtsforschung. Retrieved January 30, 2011, from http://www.zlf.uni-frankfurt.de/

Zhang, M. (2006). Topics in Chinese Grammar. Retrieved October 29, 2014, from http://teaching.ust.hk/~huma300k/suppl2.htm

Zobl, Helmut. 1983. 'Markedness and the Projection Problem'. Language Learning 33(3):293–313.