MASTERARBEIT | MASTER'S THESIS

Titel | Title

The Long-Term Impact of Explicit Pronunciation Instruction on Austrian University Students' Production of Lingua Franca Core Features: Development and Learner-Reported Influences

verfasst von | submitted by

Marija Djenadic BEd

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Education (MEd)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 199 507 525 02
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Lehramt Sek (AB) Unterrichtsfach Englisch Unterrichtsfach Psychologie und Philosophie
Betreut von | Supervisor:	Mag. Dr. Karin Richter

Acknowledgements

As I bring my Master's studies at the University of Vienna to a close, a long chapter of learning and growth comes to an end, while another begins with new challenges ahead. At this turning point, I wish to express my heartfelt gratitude to those who have guided and supported me throughout this journey.

First and foremost, I would like to extend my deepest gratitude to my supervisor, Dr. Karin Richter, for her invaluable support and guidance throughout my studies. Thank you, Karin, for always believing in me when it mattered most and for showing genuine interest in my research ideas. I am deeply grateful for the many conversations we shared, which often sparked new perspectives and inspiring ideas. Your thoughtful feedback, academic advice, and trust in my work were essential to bringing this thesis to completion. Without your dedication and constant encouragement, my academic journey would hardly have taken shape in the way it did.

I am also profoundly grateful to every single participant who agreed to take part in this longitudinal study. Without your willingness to dedicate your time, this research project would not have been possible. I truly appreciate the contribution each of you made.

Special thanks go to my family, above all to my parents and my brother, who have always believed in me and supported every decision I have made. You have stood by me in all aspects of life and always found the right words at the right moments. Your love and support have played a vital role in enabling me to complete my studies, and I am endlessly thankful for that. I am equally grateful for my friends: my best friend Ruki, who knew when it was time to pause and simply sit down for a coffee as well as my university friends, Valerie, Sarah, and Verena, whose encouragement carried me through the ups and downs of student life. I am so glad we shared this journey together.

Above all, I would like to thank my fiancé, Filip. Your unwavering encouragement, mental support, and love have been the greatest gift throughout this process. You showed the deepest understanding for the many nights and weekends I devoted to this thesis, and you kept my motivation alive with coffee deliveries and silly jokes that made everything more bearable. I could not be more grateful to have you by my side.

Abstract

Pronunciation is a key component of communicative competence in English as a Foreign Language (EFL) teaching, particularly at the tertiary level, where students rely on English in academic and professional contexts. Research has shown that explicit pronunciation instruction (EPI) can play a crucial role in this respect, enhancing both phonological accuracy and overall communicative effectiveness. Yet, most studies have concentrated on short-term effects of EPI, with limited evidence on whether these gains are sustained over longer periods of time. Moreover, little is known about the extent to which advanced EFL learners with German as a first language (L1) retain or lose features that are critical to intelligibility, and which individual and contextual factors influence their pronunciation development. As a result, the long-term impact of EPI for this learner group remains largely unexplored. In response to these gaps, this thesis examines the long-term effects of EPI on advanced Austrian EFL learners' production of two selected segmental features of the Lingua Franca Core (LFC), namely vowel length in long monophthongs and word-initial aspiration of fortis plosives. Employing a longitudinal mixed-methods design, data were collected from 28 Bachelor students at the University of Vienna who were enrolled in the EPI course 'Practical Phonetics and Oral Communication Skills 1' (PPOCS 1) and its accompanying language lab. Participants' pronunciation development was assessed through pre- (T1), post- (T2), and delayed post-tests (T3), while a web-based survey at T3 elicited learner-reported influences on pronunciation changes over time. The findings reveal clear short-term gains in both vowel length and aspiration after one semester of EPI, with vowel length showing particularly strong improvement. At T3, much of this progress was retained. Although a modest decline in vowel length and aspiration accuracy was observed, mean scores remained well above pre-test (T1) levels. Survey responses further revealed that participants' pronunciation development was influenced less by deliberate practice and more by accessible forms of exposure, such as media input and occasional spoken interaction. Overall, the study demonstrates that EPI can foster durable gains in tertiary-level EFL learners' intelligibility. Together, the findings contribute new longitudinal evidence on pronunciation development and retention among advanced Austrian German-speaking EFL learners.

Table of contents

List of abbreviations

AE …… American English

BA …… Bachelor of Arts

BEd …… Bachelor of Education

CEFR …… Common European Framework of Reference for Languages

CLT …… Communicative Language Teaching

EFL …… English as a Foreign Language

ELF …… English as a Lingua Franca

EPI …… Explicit Pronunciation Instruction

GA …… General American

ID …… Individual Difference

IPA …… International Phonetic Alphabet

L1 …… First Language

L2 …… Second Language

LFC …… Lingua Franca Core

PPOCS 1 …… Practical Phonetics and Oral Communication Skills 1

RP …… Received Pronunciation

SLA …… Second Language Acquisition

T …… Time

List of figures

List of tables

1 Introduction

As English continues to function as a global lingua franca, the ability to communicate successfully has become a central skill for English Foreign Language (EFL) learners in higher education, many of whom rely on English in academic and professional contexts (cf. Seidlhofer, 2011; Jenkins, 2000). Yet, even advanced EFL learners may encounter communication difficulties if their pronunciation is not intelligible, since grammatical accuracy and lexical knowledge alone cannot guarantee successful interaction. Pronunciation, therefore, emerges as a decisive factor in developing communicative competence, with research stressing that intelligible speech is indispensable for effective communication in a foreign language (Dalton-Puffer et al., 1997, p. 116). This recognition has brough renewed attention to the role of explicit pronunciation instruction (EPI) which aims to enhance learners' intelligibility through targeted teaching of key pronunciation features. A major contribution in this regard is Jenkins' (2000) Lingua Franca Core (LFC) which identifies those features most crucial for mutual understanding in international communication, thereby shifting attention away from native-like norms toward communicative effectiveness. Yet, despite this growing recognition, important questions remain about whether the improvements fostered through EPI can be sustained over time.

Research to date largely confirms that EPI can lead to measurable improvements in learners' accuracy and intelligibility (Gordon & Darcy, 2022; Saito, 2012; Derwing & Munro, 2005). However, most of these investigations have relied on short-term post-tests, leaving open the question of how well learners retain the pronunciation skills they acquire and how their performance develops once instruction has ended. Understanding the long-term effects of EPI is particularly relevant, as communicative demands in real-world settings extend far beyond a single EPI course. Moreover, while research has examined EFL learners from a range of first language (L1) backgrounds, including Ukrainian (Halahan et al., 2023), Spanish (Gordon & Darcy, 2022), Japanese (Saito, 2012), and Persian (Yenkimaleki & van Heuven, 2019; Naeini & Adni, 2017; Ghorbani et al., 2016), comparatively little is known about German-speaking EFL learners. This gap is particularly pressing given that German L1 speakers, especially those from Austria, exhibit distinctive

pronunciation challenges that overlap directly with core features of the LFC (Berger, 2010; Swan & Smith, 2001).

Beyond instructional effects, increasing attention has also been paid to individual and contextual factors that influence how learners' pronunciation develops over time. In this context, research suggests that such factors can be as decisive as pronunciation instruction itself (cf. Gilakjani, 2012). These variables include learners' L1 background, their motivation, attitudes, sense of identity, as well as the extent to which they practise and are exposed to English in contexts outside the classroom. Despite the significance of these influences, many empirical studies on EPI have concentrated primarily on measurable changes in speech production, often overlooking the broader learner-related factors that contribute to pronunciation development and retention.

Against this backdrop, the present thesis seeks to address two central research gaps. First, it investigates the long-term development and retention of pronunciation gains following EPI, thereby contributing to the relatively small body of longitudinal studies in pronunciation research (Derwing & Munro, 2013, p. 180). Second, it explores these effects in the under-researched context of advanced Austrian EFL learners, a group for whom pronunciation challenges have been documented in features such as vowel length and the realisation of consonant sounds that are central to the LFC (cf. Berger, 2010; Swan & Smith, 2001). To this end, the study adopts a longitudinal panel design involving 28 advanced EFL learners enrolled in an EPI course at the University of Vienna. Specifically, this thesis draws on data collected at three points: prior to EPI (T1), immediately after the EPI course (T2), and again sixteen months later (T3). In addition to auditory analyses of learners' production of vowel length and word-initial aspiration, the study incorporates data from a web-based survey, providing insight into learner-reported factors such as motivation, identity, and out-of-class exposure that may influence long-term pronunciation development. By combining a focus on segmental LFC features with an exploration of individual and contextual influences, this thesis aims to extend the current understanding of how EPI impacts not only immediate gains but also the durability of these pronunciation skills in advanced Austrian learners of English.

The thesis is organised into two main parts with Part I establishing the theoretical framework of the study. In this first part, Chapter 2 clarifies key terminology that lays the basis for understanding the subsequent chapters. Chapter 3, then, shifts the focus to the role of pronunciation in EFL teaching, reviewing the effectiveness of implicit and explicit instructional approaches, and discussing research on the short- and long-term outcomes of EPI. This is followed by an introduction to the concept of English as a Lingua Franca and Jenkins' Lingua Franca Core (Chapter 4), with particular attention to the intelligibility-related features most relevant for Austrian EFL learners. Building on this, Chapter 5 broadens the discussion by examining cognitive, social, and affective factors such as motivation, attitude, identity, and out-of-class exposure, all of which may shape both the development and long-term retention of pronunciation gains. Part II of the thesis focuses on the empirical study in which Chapter 6 details the research design, methodology, and data collection, including the pre-test, post-test, and delayed post-test. The same chapter also presents findings on the development and retention of the pronunciation features under investigation, alongside insights from questionnaires on individual and contextual influences. The results in this chapter are interpreted in light of the theoretical and empirical discussions in Part I. Finally, Chapter 7 summarises the main findings of the study, followed by the study's limitations directions for future research (Chapter 8).

2 Key Terminology

Before turning to the theoretical and methodological considerations of this study, it is first of all crucial to define the key terms that lay the basis for the analysis and discussion in the chapters that follow. Given that this thesis investigates the long-term development of selected segmental features among Austrian EFL learners, it is essential to clarify the phonological concepts that underpin this focus. This chapter, therefore, introduces the core terminology needed to contextualise the study. It begins by distinguishing between language perception and production, both of which play a central role in language learning and are particularly relevant given the study's focus on learners' oral production of segmental features. Section 2.2, then, outlines the difference between segmental and suprasegmental features of pronunciation, providing the necessary phonological framework for the analysis of the collected data. Finally, Section 2.3 discusses three key constructs in second language (L2) pronunciation research, namely intelligibility, comprehensibility, and accentedness, which are central to the study's focus on intelligibility-relevant features.

2.1 Perception vs. production

In order to understand the scope of this thesis, which centres on the production of selected segmental features, it is essential to first distinguish between the closely related concepts of language perception and language production. Both processes are fundamental to spoken communication and play distinct roles in the development of oral language skills.

Language perception refers to a listener's ability to hear, recognise, and categorise spoken sounds, often involving distinctions between phonemes that may not exist in the listener's L1 (Nagle, 2018, p. 235). In fact, L2 learners are frequently challenged in detecting unfamiliar or subtle phonetic contrasts, which may influence the precision of their production of similar sounds. Importantly, perception is not merely a passive act of decoding the acoustic signal. Rather, it is shaped by the listener's expectations, their degree of attention, and previous experiences with the target language (Schertz & Clare, 2019, p. 11). Together, these factors may help explain why L2 learners differ in how successfully they perceive and process English sounds.

Language production, on the other hand, involves the speaker's articulation of speech sounds and the generation of an acoustic signal from an intended message (Nagle, 2018, p. 234). In second language acquisition (SLA)[1], production is generally defined as the learner's ability to reproduce target-like sounds, including those that may not exist in their L1 phonological system. While it is often assumed that perception leads to improved production, this relationship is not always linear. Research has shown that learners can sometimes produce accurate speech sounds without reliably perceiving them, and vice versa (Kim & Clayards, 2019, p. 769).

Given that this thesis focuses on learners' spoken production of selected segmental features over time, a clear understanding of what production entails, and how it differs from perception, provides an essential foundation for interpreting the study's findings.

2.2 Segmental and suprasegmental features

Having clarified the distinction between perception and production, this section turns to the key phonological features that are realised in spoken language. These features are typically grouped into two broad categories, namely segmental and suprasegmental features, each of which plays a distinct role in spoken language. A comprehensive understanding of both domains is, thus, central to examining their relevance for successful oral communication.

To begin with, segmental features involve the articulation of individual sounds of speech (i.e., consonants and vowels), and are perceived as distinct units, so-called segments or phonemes, within a language (Kelly, 2000, p. 3). These phonemes form the basic building blocks of spoken language and are central to conveying meaning accurately. In fact, a small change in a single phoneme can result in a different meaning, as illustrated by the contrast between <fine> and <vine> (Plag et al., 2015, p. 37; Yule, 2010, p. 42; Kelly, 2000, p. 1). Such pairs, which differ in only one sound, are known as minimal pairs and play a crucial role in successful communication, due to their distinguishing meaning (Yule, 2010, p. 44; Kelly, 2000, p. 11). In order to represent these segmental features accurately

[1] Second language acquisition (SLA) refers to the process by which people learn a language other than their native language (cf. Larsen-Freeman, 2018)

across languages, the International Phonetic Alphabet (IPA) serves as a standardised system in both research and language instruction for the consistent transcription of sounds (Plag et al., 2015, p. 6).

Among segmental features, consonants can be classified according to how they are articulated. Specifically, they are produced by obstructing the airstream in various ways and at different points in the vocal tract, which results in a wide range of distinct sounds (Plag et al., 2015, p. 12). To describe these sounds, scholars typically draw on three main criteria: place, manner, and force of articulation (Plag et al., 2015, p. 17; Roach, 2009, pp. 51-52). Accordingly, Table 1 provides an overview of this classification, illustrating how individual English consonants are categorised based on these three dimensions.

Table 1. *English consonant chart*

MANNER OF ARTICULATION	PLACE OF ARTICULATION															
	Bilabial		**Labio-dental**		**Inter-dental**		**Alveolar**		**Palato-alveolar**		**Palatal**		**Velar**		**Glottal**	
Plosive	**p**						**t**	**d**					**k**	**g**	**ʔ**	
Fricative			**f**	**v**	**θ**	**ð**	**s**	**z**	**ʃ**	**ʒ**					**h**	
Affricate									**tʃ**	**dʒ**						
Nasal		**m**						**n**						**ŋ**		
Lateral liquid								**l**								
Retroflex liquid								**ɹ**								
Glide		**w**										**j**				

FORCE OF ARTICULATION

Voiceless	Voiced

As can be seen in Table 1, the consonant sounds are systematically organised according to how and where they are produced. The force of articulation (i.e., whether a sound is voiced or voiceless) adds a further layer of contrast among consonants, ensuring that no two sounds are articulated in exactly the same way. For instance, /t/ in *ten* and /d/ in *den* are identical in manner and place of articulation but differ in voicing, which makes them distinct sounds in English.

Vowel sounds, in contrast to consonants, are produced without complete obstruction of the airflow and are shaped primarily by the position of the tongue and the shape of the lips (Plag et al., 2015, p. 18; Kelly, 2000, p. 5). To classify vowel sounds in English, factors such as vowel height, vowel frontness, and lip rounding, which collectively determine vowel quality, are used as key descriptors (Plag et al., 2015, p. 18). The twelve vowels in English are shown in the vowel chart (Figure 1) below.

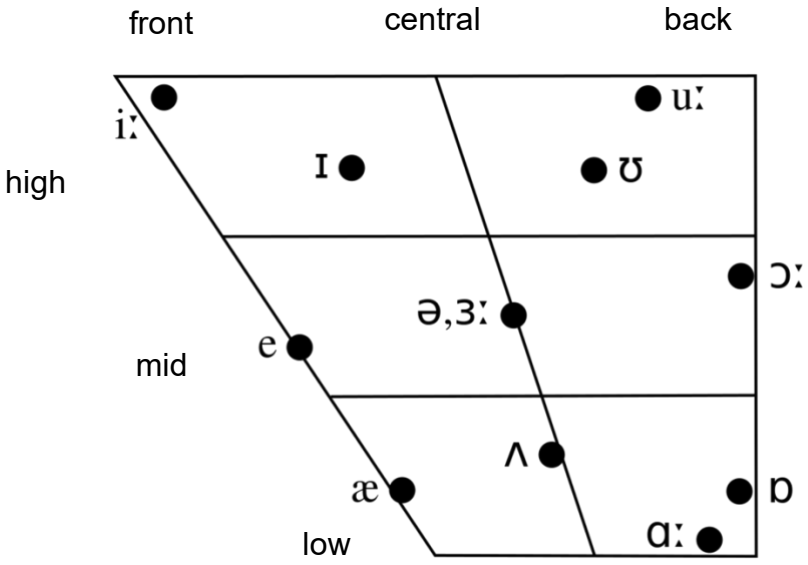


Figure 1. *English vowel chart (taken and adapted from Plag et al., 2015, p. 20)*

As can be seen in Figure 1, vowels are categorised according to the height and backness of the tongue during articulation. Tongue height refers to the vertical position of the tongue and includes high (e.g., /uː/), mid (e.g., /ɔː/), and low (e.g., /æ/) vowels. Tongue backness describes how far forward or back the tongue is and distinguishes between front (e.g., /iː/), central (e.g., /ɜː/), and back (e.g., /ɑː/) vowels. Together with lip rounding, these articulatory features contribute to vowel quality in English.

Although the primary focus of this thesis lies in the analysis of segmental features, it is still worthwhile to briefly address suprasegmental features to ensure a complete overview of the key features of pronunciation. In contrast to segmental features, which concern individual sounds, suprasegmental features operate over larger stretches of speech, including syllables, words, and sentences (Kelly, 2000, p. 3). These features involve, most notably, stress, intonation, and connected speech (Zielinski, 2015, p. 397; Kelly, 2000, p.

3). In this regard, suprasegmentals contribute to the overall prosodic shape of spoken language and are closely connected to a speaker's intention and discourse organisation (Zielinski, 2015, p. 398).

One such suprasegmental feature is word stress, which helps signal important information and indicate the structure of spoken discourse. In English, this involves producing stressed syllables with greater prominence, typically louder, longer, and higher in pitch than their unstressed counterparts. Since stress in English is contrastive, it can influence a word's meaning, as demonstrated in examples like <record> (noun) vs. <record> (verb) (Schmitt, 2016, pp. 137–138). Closely linked to stress is intonation which refers to the 'melody' of speech (Schmitt, 2016, p. 147). Intonation plays a key role in expressing attitudes, emotions, and pragmatic functions such as questioning, suggesting, or emphasising. Importantly, deviating from expected intonation patterns can lead to misunderstandings, especially for non-native speakers, who may be perceived as abrupt, disinterested, or impolite, even when their segmental production is accurate (Celce-Murcia et al., 2010, p. 247). Another key suprasegmental feature is rhythm which is also shaped by stress and intonation patterns. In fact, English is often described as a stress-timed language, meaning that the timing between stressed syllables tends to be relatively equal, regardless of the number of intervening unstressed syllables. This rhythmic structure contributes to the overall fluency and naturalness of speech (Roach, 2009, pp. 107-108). Finally, connected speech, including features such as linking, elision, and assimilation, further influences the flow and rhythm of spoken language. Collectively, these features aid speakers in producing speech more efficiently and fluently (Roach, 2009, pp. 111–116; Kelly, 2000, pp. 108–111).

As can be seen, suprasegmental features have a substantial impact on the perceived fluency and naturalness of speech. In some cases, particularly when stress or intonation patterns are used inappropriately, they may even affect intelligibility (Zhang & Yuan, 2020, p. 2; Field, 2005, p. 410). In fact, research has shown that inappropriate or absent stress patterns can significantly hinder comprehension, even when individual sounds are produced accurately (Hahn, 2004, p. 201). In any case, given the specific focus of this thesis on segmental features, suprasegmental features will not be examined in greater detail.

2.3 Intelligibility, comprehensibility and accentedness

Building on the discussion of segmental and suprasegmental features, it is equally important to consider how L2 speech is evaluated in terms of listener understanding and communicative success. In this context, three key concepts that are frequently referred to throughout this thesis are defined: intelligibility, comprehensibility, and accentedness. Although these terms are commonly used in pronunciation research, their definitions have often led to terminological confusion (Jenkins, 2000, p. 71). Given that this study adopts an intelligibility-oriented approach to examine the development of EFL learners' pronunciation over time, clarifying how these constructs are defined and differentiated is crucial for the analysis and discussion in subsequent chapters.

The first concept to be discussed is intelligibility, which is widely recognised as a central goal in pronunciation teaching, particularly in ELF contexts. ELF refers to the use of English as a common means of international communication among speakers from different L1 backgrounds, most of whom are non-native speakers (Jenkins, 2000, p. 11). Within this global context, the notion of international intelligibility is particularly relevant, referring to how well speakers from different linguistic backgrounds understand one another globally (Kaur, 2018, p. 542; Berns, 2008, p. 329). Despite its prominence in global communication, defining intelligibility itself often remains a challenge. As Field (2005, p. 400) notes, much of the discussion surrounding the construct lacks a clear consensus. In one of the most frequently cited definitions, Munro et al. (2006, p. 112) describe intelligibility as "the extent to which a speaker's utterance is actually understood by a listener", highlighting that intelligibility is best defined from the listener's perspective (Field, 2005, p. 400). Accordingly, in research, intelligibility is typically measured through tasks that require listeners to transcribe spoken utterances or identify individual words, as these methods provide concrete evidence of what was actually understood. However, such tasks have been criticised for reducing intelligibility to a question of linguistic accuracy, often overlooking the interactive and context-dependent nature of real-life communication. As Rajadurai (2007, p. 90) argues, many studies fail to acknowledge the active role of the listener *and* the speaker's efforts to accommodate, both of which crucially shape intelligibility in actual communicative exchanges. Further, contextual factors such as setting, topic, and the relationship between interlocutors, all of which influence how

speech is produced and processed, are frequently disregarded. As a result, the validity of assessing intelligibility using decontextualised word lists or isolated sentence recognition tasks is questionable. In response to this, Jenkins (2007, p. 85) calls for more valid research approaches and emphasises the need for naturalistic interactions between speakers that more accurately reflect how intelligibility operates in real-world exchanges. These observations suggest that intelligibility should not be viewed as a purely linguistic feature of speech but rather as a dynamic process that emerges through interaction and is influenced by a range of contextual and listener-related features, such as the communicative setting or the shared background knowledge.

Closely related to intelligibility is the concept of comprehensibility, which refers to a related but separate construct. While intelligibility measures actual understanding, comprehensibility reflects the perceived effort required to process spoken language. As Munro and Derwing (1995, p. 76) explain, “comprehensibility refers to listeners’ perceptions of the ease or difficulty involved in understanding an utterance.” This means that an utterance may be fully intelligible (i.e., all words are recognised correctly) yet still be perceived as highly difficult to follow (Levis, 2018, p. 17). Hence, this discrepancy suggests that comprehensibility involves not only transmitting the linguistic message but also drawing on listeners’ cognitive resources to interpret it (Derwing & Munro, 2005, pp. 385–386). Factors such as speech rate, fluency, and the listener’s familiarity with the speaker’s accent further shape how easily an utterance is processed. In this sense, comprehensibility plays a vital role in determining how well communication flows in real-time, especially in spontaneous interactions (Derwing & Munro, 2005, p. 385).

The third concept, accentedness, adds a further dimension to the evaluation of L2 speech. Accentedness refers to the degree to which a speaker’s pronunciation is perceived as deviating from a listener’s mental representation of a target or standard accent. In other words, accentedness is defined as “how much an L2 accent differs from the variety of English commonly spoken in the community.” (Derwing & Munro, 2005, p. 385). Unlike intelligibility and comprehensibility, accentedness does not relate to the success or ease of communication but rather to perceived nativeness. While accentedness often refers to a foreign accent in L2 speech, the concept can also describe regional or social varieties within an L1 (e.g., Scottish or Australian English). Importantly, a strong foreign accent

does not automatically imply low intelligibility or comprehensibility. To this, Levis (2018, p. 11) points out that speech can be heavily accented yet still be easily understood, especially when speaker and listener share the same L1, suggesting that phonological deviations do not necessarily hinder successful communication. This distinction is particularly relevant in ELF contexts, where a wide range of accents is common and intelligibility is often prioritised over native-likeness (Jenkins, 2000, pp. 69–70). It is also important to note that perceptions of accentedness are often influenced by listener bias, expectations, and attitudes, making it a more subjective and socially charged construct than intelligibility and comprehensibility (Derwing & Munro, 2009, p. 479).

To sum up, intelligibility, comprehensibility, and accentedness represent three interrelated but distinct constructs that affect how L2 pronunciation is perceived and understood. Whereas intelligibility concerns the successful transmission of meaning, comprehensibility relates to the listener's processing experience, and accentedness refers to the perceived deviation from a norm. The distinction between these terms will inform both the theoretical framework and the interpretation of the empirical data in the present study.

3 Pronunciation in EFL teaching

Having defined key terminology that lays the foundation for this study, this chapter shifts the focus to the role of pronunciation in EFL teaching. Although often overshadowed by grammar and vocabulary, pronunciation is a core component of communicative competence (i.e., the ability to use language effectively and appropriately in various social interactions) (Walker, 2010, p. 147). Given that developing communicative competence is the central aim of Communicative Language Teaching (CLT)[2], this approach prioritises intelligibility and spoken interaction. Within this framework, pronunciation teaching becomes essential for enabling learners to practise spoken interaction. Against this backdrop, this chapter explores the relevance of pronunciation instruction in greater detail. The chapter begins by outlining the importance of pronunciation, before examining two

[2] CLT is a widely used approach in language education that prioritises meaningful communication and the development of communicative competence. Rather than focusing on the explicit teaching of grammar rules or isolated linguistic forms, CLT emphasises interaction, fluency, and the use of authentic language in context (Richards & Rodgers, 2001, pp. 156-157).

instructional approaches, namely implicit and explicit pronunciation instruction. It then reviews empirical findings on the effectiveness of EPI, with particular attention to its long-term impact. Finally, the chapter considers the overarching goals that pronunciation instruction should pursue, including a discussion of models and norms, the prominence of General American (GA) and Received Pronunciation (RP), and the growing emphasis on intelligibility as a goal in pronunciation instruction.

3.1 Importance of pronunciation instruction

Pronunciation forms a core component of foreign language competence and is widely acknowledged as essential for successful oral communication. In fact, its importance often becomes particularly evident when learners experience communication breakdowns due to imprecise pronunciation. Even highly proficient learners who demonstrate strong L2 lexical and grammatical control may encounter difficulties in being understood if their pronunciation, at both the segmental and suprasegmental level (see Section 2.2), lacks clarity (Nguyen et al., 2021, p. 2; Fraser, 2000, p. 7). Despite its crucial role in communication, pronunciation is frequently overshadowed by grammar and vocabulary in classroom practice. Yet, research consistently identifies pronunciation as the most immediately noticeable and influential feature of spoken interaction, directly shaping how successfully a speaker is understood (Almusharraf, 2024, p. 2; Uzun, 2022, p. 23; Nguyen et al., 2021, p. 2; Atar, 2018, p. 97; Derwing & Munro, 2005, p. 379). As Dalton-Puffer et al. (1997, p. 116) aptly put it, “good pronunciation is […] indispensable for adequate communication in a foreign language.”

Extending the discussion to segmental features, even minor inaccuracies can lead to considerable misunderstandings. In this regard, Jenkins (2000, p. 83) argues that pronunciation is the linguistic domain most likely to impede successful communication, as subtle phonetic variations can significantly alter the intended meaning. Along similar lines, Kelly (2000, p. 11) highlights that a single mispronounced sound may entirely alter the intended message and cause misunderstandings. The impact of segmental inconsistencies often becomes evident in everyday situations, such as ordering in a restaurant. For example, a speaker who mispronounces a single phoneme (e.g., a vowel sound) by requesting soap (i.e., /səʊp/) instead of soup (i.e., /su:p/) is highly likely to cause

misunderstandings, which may eventually result in unintended confusion (Kelly, 2000, p. 11). In contexts with higher communicative risk, such mispronunciations can have far-reaching consequences, as demonstrated by the 1977 Tenerife airplane disaster, in which the imprecise articulation of standard aviation phraseology contributed to a fatal miscommunication between the pilot and air traffic control, both of whom were non-native speakers of English (Szpyra-Kozłowska, 2015, p. 2). These examples illustrate that even advanced EFL learners with a solid command of syntax and lexis may fail to communicate effectively if their pronunciation is imprecise at the segmental level. Given the potential severity of such misunderstandings, Munro and Derwing (2011, pp. 316-317) stress the importance of intelligible pronunciation for successful oral interaction, highlighting the need for pronunciation instruction to explicitly address segmental features that may hinder intelligibility.

While segmental accuracy is undeniably crucial, suprasegmental aspects of speech, such as word stress, rhythm and intonation, also play a vital role in ensuring successful communication. When used effectively, suprasegmental features can help listeners process information in meaningful units. However, a misplaced stress, for example, can impose strain on the listener and hinder the processing of entire chunks of speech (Jenkins, 2000, p. 42). Such issues may arise even when segmental articulation is accurate, suggesting that intelligibility depends on more than just the precise production of individual sounds. Yet, as Darcy (2018, p. 30) cautions, accurate suprasegmental control alone cannot compensate for unintelligible segmental pronunciation. This reciprocal relationship highlights the need to consider both aspects when addressing pronunciation in the classroom. Indeed, rather than prioritising one feature over the other, Gordon and Darcy (2022, p. 190) propose that pronunciation instruction should take an integrated approach, targeting both individual sounds and prosodic features to equip learners with the necessary tools for effective oral communication.

Beyond supporting phonological accuracy, pronunciation instruction also plays a central role in shaping learners' confidence and willingness to engage in spoken interaction, particularly at the tertiary level. As Sharma (2021, p. 65) notes, targeted pronunciation instruction can boost speaking confidence and, in turn, increase learners' readiness to participate in spoken tasks. This is especially relevant for advanced EFL learners in

academic or professional settings, where high levels of spoken proficiency are often expected. However, if existing pronunciation difficulties remain unaddressed, they may undermine learners' confidence and lead to increased anxiety or reluctance to communicate (Baran-Łucarz, 2014, p. 445). Consistent pronunciation training can, therefore, serve not only to improve production skills but also support communicative engagement (Baran-Łucarz, 2014, p. 467; Clément et al., 2003, p. 191). This is echoed in Nguyen et al.'s (2021, p. 8) study in the Vietnamese university context, where both teachers and students identified pronunciation instruction as essential for developing clarity and confidence needed for professional communication.

Despite these well-documented benefits, pronunciation instruction continues to receive limited attention in many EFL teaching contexts. This marginalisation has been linked to several interrelated factors (Nguyen & Newton, 2020, p. 12; Murphy, 2018, p. 301; Lightbown & Spada, 2013, p. 68). On the institutional level, pronunciation is frequently deprioritised in curricula, with grammar, reading, and writing often receiving comparatively more attention (Sesma et al., 2022, p. 8; Ghorbani et al., 2016, p. 60). From a pedagogical perspective, many teachers report a lack of adequate training in phonetics and phonology (Bai & Yuan, 2018, p. 134) and feel insufficiently prepared to teach pronunciation effectively (Richter & Schölzky, 2023, p. 5). Even experienced instructors may lack confidence in this area, which further discourages the inclusion of pronunciation teaching in the classroom. Murphy (2018, p. 301) and Lightbown and Spada (2013, p. 68) similarly note that pronunciation often receives less classroom time than other skills (e.g., reading, writing), despite its demonstrable impact on communicative competence. While these studies differ in emphasis, some focusing on curricular constraints, others on teacher-related challenges, they collectively point to a persistent disconnect between the communicative importance of pronunciation and its practical implementation in EFL instruction across different educational levels.

Taken together, these insights highlight the multifaceted role of pronunciation teaching, most notably in developing EFL learners' communicative competence. Yet, beyond its importance in EFL instruction, it is also essential to consider how pronunciation can be taught and which instructional approaches best support EFL learners' long-term development - issues that will be examined in the subsequent sections.

3.2 Implicit vs. explicit pronunciation instruction

Having established the communicative importance of pronunciation, the next step is to consider how phonological features can be effectively addressed in instructional settings. A central distinction in pronunciation teaching lied between implicit and explicit approaches. While this section predominantly focuses on instructional approaches, it is useful to briefly distinguish them from methods, as the two terms are sometimes used interchangeably in pedagogical discussions. An approach generally refers to a theoretical orientation toward language learning, whereas a method is the practical realisation of that approach, including specific procedures, teacher and learner roles, and materials used (Harmer, 2011, p. 62). Although some teaching methods are mentioned for illustrative purposes, this section is concerned primarily with approaches rather than methods.

To begin with, implicit pronunciation instruction is characterised by a learning process in which phonological features are acquired gradually and without overt explanation. Rather than drawing learners' attention to specific rules, this approach relies on rich auditory input and repeated exposure to target-language speech (Yenkimaleki & van Heuven, 2019, p. 117). Thus, learners are assumed to absorb pronunciation features naturally, often without conscious awareness of the underlying phonological forms (Ellis, 2009, p. 16). For example, students might listen to story episodes or participate in communicative tasks such as role-plays, discussions, or listening activities, through which they gradually develop sensitivity to segmental and suprasegmental patterns (Norris & Ortega, 2000, p. 437; Derwing & Munro, 2015, p. 386). The aim of this approach is to support learners in internalising pronunciation features (e.g., intonation, segmental sounds) through repeated exposure to intelligible input and communicative use, rather than through explicit instruction. While this approach aligns with the principles of CLT, it may provide limited support in helping learners notice and correct persistent pronunciation inconsistencies, especially when those challenges are not salient in the input they receive.

In contrast, EPI involves a structured and intentional focus on phonological form, targeting both segmental and suprasegmental features (Saito, 2011, p. 46). EPI typically incorporates metalinguistic explanation, rule clarification, and targeted practice. Instruction may be delivered deductively (i.e., through direct explanation) or inductively

(i.e., through guided discovery), enabling learners to notice specific features and patterns in the target language (DeKeyser, 2003, p. 321). This might involve, for instance, contrasting voiced and voiceless consonants or raising awareness of stress placement. According to Derwing and Munro (2005, p. 388), EPI can aid learners in identifying mismatches between their own productions and more intelligible target forms, thereby promoting phonological awareness. This increased awareness plays a crucial role in helping learners gain more control over their speech output, which supports the development of intelligible communication (Field, 2005, pp. 401–402). What is more, EPI typically engages learners in deliberate and focused practice activities, often supported by visual models or auditory input, which has been described as a particularly effective in developing learners' pronunciation skills (Archer & Hughes, 2011, p. 1). Yenkimaleki and van Heuven (2019, pp. 117–118) further emphasise that learners receiving EPI are generally able to verbalise what they have learned, which indicates deeper cognitive processing.

The debate between implicit and explicit pronunciation instruction has given rise to divergent theoretical perspectives on how foreign language learners best acquire phonological features. Advocates of implicit pronunciation instruction often draw on Krashen's (1981) Input Hypothesis, which holds that language acquisition occurs most effectively through exposure to comprehensible input rather than explicit learning. From this perspective, applying learned phonological rules in real-time communication is considered impractical due to learners' limited cognitive resources. In contrast, other scholars (cf. Yenkimaleki & Heuven, 2019; Saito, 2011; Derwing & Munro, 2005) emphasise the value of EPI, especially in its potential to improve learners' intelligibility and communicative competence. In response to such claims, an increasing number of studies have investigated the effectiveness of EPI in promoting intelligible L2 pronunciation. Section 3.2.1, therefore, reviews a number of empirical findings on EPI in greater detail.

3.2.1 Empirical findings on EPI

As outlined in the previous section, EPI is designed to raise learners' awareness of phonological features and engage them in focused practice to improve intelligibility. In fact, a substantial body of research has explored the impact of EPI, particularly its

effectiveness in promoting both suprasegmental and segmental gains, enhancing learners' performance at tertiary level, and reducing affective barriers to oral communication. This section reviews selected empirical studies that provide insight into the short-term benefits of EPI across different learner groups and instructional settings.

One area that has received considerable attention in research is the impact of EPI on suprasegmental features. In a ten-week intervention, Gordon and Darcy (2022, p. 189) found that university students receiving EPI on prosody, rhythm, and stress outperformed groups who received segmental-focused or combined instruction, particularly in terms of fluency and comprehensibility in spontaneous speech. Comparable findings have been observed by Yenkimaleki and van Heuven (2019, p. 130). Learners who received explicit prosody instruction, supported by theoretical explanations, outperformed those who were exposed to implicit pronunciation instruction through recasts (i.e., reformulations of learners' errors that preserve the intended meaning). These findings suggest that the focused and explanatory nature of EPI can facilitate greater phonological awareness of suprasegmental features, leading to improved spontaneous oral production. However, in both studies, the reported gains were only measured immediately after instruction, raising questions about the long-term durability of such improvements.

Similar short-term benefits have also been observed for segmental features. One notable contribution by Naeini and Adni (2017, pp. 116–117) demonstrated that explicit multisensory training on consonants and consonant clusters significantly improved segmental accuracy and fluency among Iranian EFL learners. Ghorbani et al. (2016, pp. 63-66) reached a comparable conclusion regarding vowel perception, showing that raising learners' phonological awareness through EPI over eight weeks was more effective than relying on imitation and repetitive input alone. In their study, participants received explicit instruction on vowel sounds and completed a series of minimal pair discrimination activities. To measure improvement, a 40-item vowel identification test was administered before and after the treatment. The findings suggest that explicit awareness-raising and targeted perceptual training can lead to significant gains in segmental perception. Further support for explicit segmental-based instruction stems from Saito's (2012) comprehensive review of 15 L2 pronunciation intervention studies, five of which explicitly targeted segmental features. All five studies demonstrated significant improvement at a controlled

level, and one study (i.e., Saito & Leyster, 2012) even reported gains in spontaneous production. The latter investigated the production of the English /ɹ/ among 64 Japanese learners of English following only four hours of EPI, showing that even short-term segmental pronunciation instruction can lead to measurable outcomes (Saito & Leyster, 2012, p. 23). Similarly, Saito's (2011, p. 45-57) earlier study found measurable segmental improvements in sentence-reading tasks, although effects on accentedness remained limited. As with much of the suprasegmental research, however, the majority of these studies only examined short-term instructional outcomes. By contrast, the present study investigates whether such segmental improvements, particularly in vowel length and word-initial aspiration, are retained over time among tertiary-level EFL learners.

Focusing on university students is especially relevant given that these learners often exhibit fossilised[3] pronunciation patterns influenced by their L1, particularly when they have received only little phonological training. A recent study by Halahan et al. (2023, p. 134) observed that many Ukrainian university students were unable to reliably identify and produce key phonological features of English, including vowel sounds, rhythm and intonation patterns. These learners frequently reported self-doubt and a tendency to avoid communicative situations due to their phonological difficulties (Halahan et al., 2023, p. 135). The authors argue that EPI, such as training in segmental distinctions and vowel length, as well as contrasts between L1 and L2 phonological systems, can help overcome these difficulties and prevent phonological fossilisation (Halahan et al., 2023, p. 139). Hence, EPI at the tertiary level can be particularly helpful in refining advanced EFL learners' pronunciation, enhancing their intelligibility, and overcoming challenges associated with L1 influence. This insight is directly relevant to the present study, which focuses on advanced EFL learners in a university context.

Beyond phonological gains, EPI may also contribute to reducing pronunciation-related anxiety and eliciting more active oral participation. By definition, pronunciation anxiety refers to learners' fear of speaking caused by concerns about making phonological errors, an affective barrier that is particularly pronounced among post-puberty learners who tend to be more self-conscious about their speech (Baran-Łucarz, 2014, p. 445). Research

[3] Fossilisation in pronunciation research refers to the persistent inability of an L2 learner to fully acquire a specific target language feature (Nakuma, 1998, p. 247).

suggests that EPI can help mitigate this form of anxiety by building learners' confidence and promoting a sense of control over their speech production. Sharma (2021, p. 64), for example, argues that EPI enables learners to gain confidence, which in turn elicits greater engagement in oral tasks. To address pronunciation anxiety more directly, Clément et al. (2003, p. 191) and Baran-Łucarz (2014, p. 467) recommend introducing pronunciation training gradually and systematically, allowing learners to develop their pronunciation skills through repeated and guided practice. These findings indicate that EPI not only fosters learners' phonological development but may also help create a more emotionally supportive and communicative EFL classroom environment.

In examining the impact of EPI, it is also important to account for contextual factors, such as the instructional setting. As Zhang and Yuan (2020, p. 2) argue, the setting of the intervention, whether a classroom or a more controlled laboratory environment, can influence learning outcomes, with laboratory settings often yielding greater phonological gains due to reduced distractions and higher treatment quality. In a similar vein, Kissling (2013, p. 720) found that it may not be the instructional approach alone, but rather the combination of input, practice, and feedback that determines learners' success in pronunciation development. Accordingly, both the instructional content and the setting in which EPI takes place should be considered when evaluating its impact.

Despite strong empirical support for EPI, some researchers continue to question whether EPI is necessary. Studies suggest that exposure to the target language alone may also yield short-term improvements in pronunciation. To exemplify, Uchihara et al. (2022, pp. 21, 33–34) observed gains in university students' comprehensibility following repeated exposure to isolated English words paired with pictures. However, the authors report that effects dissipated within one week, suggesting that input alone may not produce lasting change in pronunciation. Similarly, Derwing et al. (2008, p. 359) found that significant pronunciation improvement typically requires sustained and explicit instruction. These findings suggest that while auditory input can support development to some extent, it is unlikely to produce long-term gains without the structured training offered by EPI. In essence, this highlights the need for more longitudinal research to investigate the lasting impact of EPI, an underexplored research area which this thesis seeks to address in the following section.

3.2.2 Long-term impact of EPI

As established in the previous Section 3.2.1, EPI has shown to produce clear short-term benefits across a range of phonological features. However, most of these findings are based on immediate post-test results, offering limited insight into whether learners retain the acquired pronunciation features over time. Drawing on Algethami (2016, p. 297), instructional effectiveness is rarely evaluated beyond the end of instruction, leaving the durability of such gains largely underexplored. Although short-term improvements in intelligibility and accuracy have been widely reported (e.g., Sardegna & Dickerson, 2023; Gordon & Darcy, 2022; Yenkimaleki & van Heuven, 2019), the question of how stable these gains are in the long run remains unanswered in much of the literature.

Although longitudinal studies in L2 pronunciation remain relatively limited mainly due to the difficulty of tracking learners over time (Nagle, 2021, p. 154; Derwing & Munro, 2013, p. 180), several investigations have begun to explore whether the benefits of EPI are retained over time. As Nagle (2021, p. 156) explains, longitudinal research in this field typically involves tracking the same group of learners across extended periods, often through delayed post-tests or follow-up phases. These studies often provide valuable insights into the durability of instructional effects that cannot be captured through immediate post-tests alone. For example, Sardegna and Jarosz (2023, p. 23) report that highly motivated learners who continued with pronunciation practice beyond EPI retained considerable improvements in suprasegmental features such as linking and word stress. Similarly, Saito and Hanzawa (2018, pp. 21–22) observed continued development in prosody among advanced Japanese EFL learners over two academic semesters, although segmental improvements proved less consistent. These findings suggest that the long-term retention of pronunciation gains may depend both on the type of feature and learners' post-instructional practice habits.

Further longitudinal studies offer more differentiated findings into how individual pronunciation features develop over time. To exemplify, Brekelmans (2017) examined whether the RP-like pronunciation of advanced Dutch EFL learners deteriorated, improved or remained stable after EPI had ended. To track changes in pronunciation, recordings from participants' third academic year were compared to those collected during

their first- and second-year pronunciation exams (Brekelmans, 2017, p. 47). Their findings reveal that while learners' general pronunciation proficiency remained stable, not all features developed in the same way. Some features, such as stress patterns and rhoticity, were acquired relatively easily and retained over time. In contrast, features known to be difficult for Durch learners, such as the dental fricatives /θ/ and /ð/, the vowel /æ/, continued to deviate from RP-like pronunciation. These results suggest that certain features may be more robust once learned, while others are more prone to fossilisation. Interestingly, retention was influenced by post-instruction exposure (e.g., the number of English-taught courses at university) more than external experiences like study abroad (Brekelmans, 2017, pp. 45–54).

The variability in long-term findings is further illustrated by Algethami's (2016, pp. 300–302) study, which tracked nine Arab EFL learners' production of the /p/ sound and showed limited long-term success, with only one participant demonstrating measurable improvement after 11 weeks of EPI. What is worth noting is that Algethami (2016) lacked immediate post-test data and focused on a single phoneme, which limits broader generalisability. Such limitations are, in fact, not exceptional as longitudinal pronunciation studies often vary in sample size, target features, and analysis procedures, complicating efforts to compare findings across different instructional settings (Thomson & Derwing, 2015, p. 326).

Given the complexity of tracking pronunciation retention, it comes as little surprise that researchers have called for more comprehensive long-term investigations. Derwing (2010, p. 27) argues that identifying which features are likely to stabilise without further instruction could help teachers prioritise persistent learner difficulties more effectively. Beyond evaluating such pronunciation retention, longitudinal studies may also serve well to reveal patterns of fossilisation, offering insights into which features are more susceptible to deteriorate, and which are likely to remain stable (Thomson & Derwing, 2015, p. 339).

Building on this call for more longitudinal designs, Derwing and Munro (2013, p. 180) argue that it is only through long-term observation that researchers can determine the lasting effects of EPI. Yet to date, very few studies have extended their investigations beyond a single academic year. Addressing this gap, the present study adopts a

longitudinal research design. By tracking advanced Austrian EFL learners who completed a semester-long EPI course, the study investigates whether gains in selected segmental features are retained across three testing phases spanning a total of one and a half years. Such an extended observation period is exceptionally rare in L2 pronunciation research and allows this study to provide new valuable insights into the long-term durability of EPI.

3.3 Goals of pronunciation instruction

While the previous sections highlighted the effectiveness and long-term impact of EPI, it is equally important to consider the goals that EPI should ultimately pursue. In fact, defining the goals of pronunciation teaching has long been a central issue in language education, particularly in higher education contexts, where learners are expected to achieve high levels of spoken proficiency for academic or professional purposes. In the following sub-sections, the discussion turns to key considerations that shape these instructional goals, including the distinction between models and norms, the prominence of native-like varieties and the growing emphasis on intelligibility as a goal in pronunciation instruction.

3.3.1 Models and norms

Before examining the goals of pronunciation teaching in greater detail, it is helpful to distinguish between the notions *model* and *norm*, as they reflect different orientations. In pronunciation instruction, a model generally refers to the spoken variety or accent presented to learners as a reference for imitation (Brown, 1991, p. 39), usually the teacher or other target language input that students are regularly exposed to (e.g., audio materials). Such models are often based on established native-speaker varieties, most commonly Received Pronunciation (RP) and General American (GA)[4]. Historically, these native-speaker varieties reflect what Kachru (1985) termed *Inner Circle Englishes*, referring to English as a native language in countries like the UK or USA. To better understand the sociolinguistic positioning of these and other varieties, Figure 2 presents

[4] A more detailed discussion of the two commonly referenced native-speaker varieties (i.e. RP and GA) is provided in the subsequent Section 3.3.2.

Kachru's (1985) Three Circles Model, which has been widely used to conceptualise the global use of English.

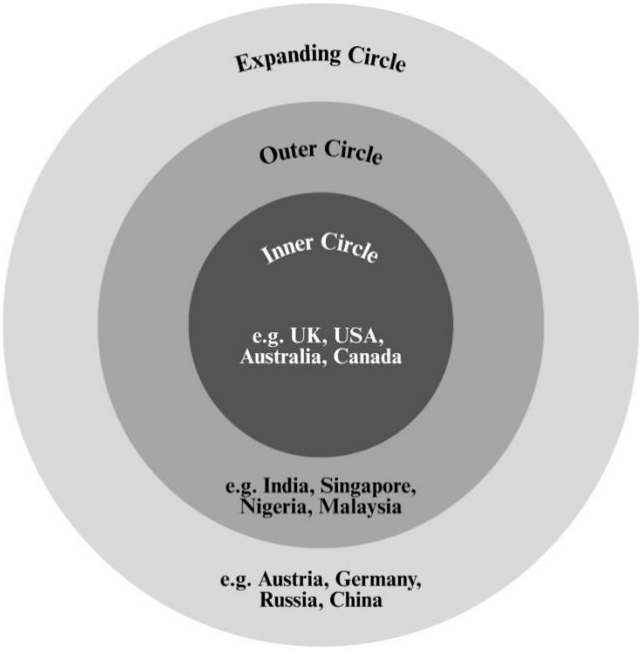


Figure 2. *Kachru's (1985) Three Circle Model*

As shown in Figure 2, Kachru's model categorises the use of English into three sociolinguistic contexts: the Inner Circle, where English is spoken as a native language (e.g., UK, USA, Australia); the Outer Circle, which includes countries with a colonial history where English functions as a L2 (e.g., India, Nigeria); and the Expanding Circle, encompassing countries where English is learned and used as a foreign language (e.g., Austria, Germany) (Jenkins, 2000, p. 12).

The continued reliance on Inner Circle models in pronunciation instruction, especially in Expanding Circle contexts like Austria or Germany, reflects a long-standing preference for native-speaker varieties as idealised standards for pronunciation learning (Jenkins, 2007, p. 9). However, using RP or GA as models does not necessarily imply that learners are expected to fully replicate them. According to Dalton and Seidlhofer (1994, p. 27), using RP or GA as a model allows for flexible approximation, depending on situational needs and instructional goals. In this view, the model serves as a guide rather than a fixed target, so that "with one accent model different goals can be pursued, e.g. minimal intelligibility, comfortable intelligibility or native-like pronunciation." (Szpyra-Kozłowska, 2015, p. 8).

In addition, many tertiary-level learners aspire to approximate native-like pronunciation for personal, academic, or professional reasons (Lightbown & Spada, 2013, p. 71; Scales et al., 2006, p. 715). In this regard, Schmitt (2016, pp. 24-25) describes the pursuit of native-like pronunciation as a form of linguistic art, in which pronunciation extends beyond its communicative purpose and becomes a matter of identity, self-expression, and linguistic pride. For advanced learners, aiming to sound like a native speaker may reflect a desire to demonstrate linguistic mastery and high-level competence. In such a case, native-like pronunciation becomes a symbol of skill and dedication that might help learners avoid negative social evaluations, such as being perceived as less competent, less educated, or less fluent. Thus, for some learners, the goal of approximating a native-speaker model accent is not imposed externally but chosen deliberately, based on personal motivation that goes beyond mere intelligibility.

In contrast to the notion of a model, norms carry stronger associations with linguistic correctness and prescriptive expectations. While a model serves as a point of reference, a norm is typically viewed as a fixed standard that learners are expected to fully attain, often regardless of context or communicative needs. As Dalton and Seidlhofer (1994, p. 27) argue, treating RP or GA as a norm implies that these varieties represent the only correct forms of pronunciation, with the ultimate goal of imitating them fully. This perspective disregards the practical realities of language use by assuming that native-like accuracy is attainable, despite empirical evidence showing that very few learners ever reach such proficiency (Saito, 2021, pp. 2-4). It further overlooks the diversity of learners' linguistic backgrounds, personal goals, and ultimately their identities (Walker, 2010, pp. 20-21).

Having defined the two terms, it becomes evident that a model is primarily pedagogical and practical, whereas a norm is more ideological or aspirational. At this point, it is crucial to state that native-speaker models such as RP or GA can serve as useful points of reference for both pronunciation teachers and EFL learners but should not automatically be treated as fixed targets.

3.3.2 GA and RP

Building on the distinction between model and norm, it is worth taking a closer look at the two previously addressed pronunciation models that continue to dominate as models in English pronunciation instruction (i.e., GA and RP). Despite increasing interest in alternative approaches to teaching pronunciation (e.g., intelligibility-based pronunciation teaching, diverse accent exposure), native-speaker models remain dominant in many EFL contexts. Most commonly, learners are exposed to RP and GA, both of which are widely regarded as prestige varieties and continue to feature prominently in teaching materials and curricula (Schmitt, 2016, pp. 27–28; Walker, 2010, p. 5).

Among these varieties, RP remains the most thoroughly described and historically prestigious accent associated with British English (BE). Often referred to as the ‘Queen’s English’, ‘Oxford English’ or ‘BBC English’ (Schmitt, 2016, p. 33, Rogerson-Revell, 2011, p. 7), RP has traditionally been linked to the British upper class and is characterised as being non-regional, a feature that has contributed to its reputation as a ‘neutral’ standard. According to Schmitt (2016, p. 30), RP is frequently used as a reference point due to its perceived clarity and status, even beyond the borders of the UK. However, given that RP is spoken by fewer than 3% of the British population (Jenkins, 2000, p. 14), its status as a dominant pronunciation model is increasingly debatable. Jenkins (2000, p. 204) even argues that adopting such a minority accent as a model may be problematic, particularly in international contexts where learners are rather unlikely to encounter RP speakers. Kelly (2000, p. 14) reinforces this view by stating that RP reflects social prestige more than communicative practicality. Despite these concerns, RP remains widely used in EPI across many European contexts, including Austria.

Alongside RP, GA has emerged as the second most prominent native model in European EFL contexts. Henderson et al. (2012, p. 21) already noted more than a decade ago that “although RP is still the dominant form for both reception and production work in English, GA seems to be making inroads”. Commonly defined as the accent that most Americans perceive as most neutral or unmarked, GA is similarly not tied to a specific region or socio-economic background and is, therefore, considered more inclusive that RP (Newbrook, 1986, p. 77). Notably, GA has gained traction among EFL students, especially in

European contexts, where learner preferences increasingly reflect the influence of global media. For instance, Henderson et al. (2012, p. 20) found that while teachers in countries such as Finland, France, and Germany tended to favour RP, their students were more inclined toward GA. A similar trend was observed among Austrian university students who expressed a growing preference for American English (AE) (Richter & Weissenbäck, 2021, pp. 7-8). The authors link this shift to the widespread influence of digital media and streaming platforms such as Netflix and Amazon Prime, where AE clearly dominates and shapes learners' perceptions of what is desirable (Richter & Weissenbäck, 2021, p. 2).

In addition to this increased media exposure, GA may also be regarded as more accessible from a phonological perspective. As Brown (1991, p. 36) and Newbrook (1986, p. 78) point out, GA's rhotic nature and relatively less complex vowel system may reduce the cognitive load for learners, making it easier to acquire. Its greater consistency with English spelling conventions, particularly compared to RP, may also support more intuitive pronunciation and spelling associations, especially for learners without extensive phonological training.

While RP and GA continue to be the two most commonly adopted native-speaker models in pronunciation instruction, it is important to acknowledge the growing recognition of other native and non-native models, particularly in response to the globalisation of English. As Murphy (2014, p. 258) argues, the goal of pronunciation instruction has never been to promote one variety, native or non-native, as the one and only ideal. Instead, a model should primarily serve as a practical and intelligible reference that supports students' learning. In this regard, exposure to a range of English accents, beyond RP and GA, can be beneficial, as it may enhance learners' receptive pronunciation skills and prepare them to engage with the diversity of real-world English use (Charpentier-Jiménez, 2019, pp. 15-16). The rise of ELF, in particular, has further questioned the dominance of native-speaker models by shifting the focus toward promoting intelligibility across diverse teaching contexts over strict adherence to a single variety. Overall, these developments suggest that while native-speaker models like RP and GA can serve as useful guides, especially for learners who consciously choose to approximate them, their use should ultimately support learners in developing intelligible speech.

3.3.3 Intelligibility as a goal

Having explored which pronunciation models are commonly used in EFL teaching, attention must also be given to the goal that pronunciation instruction is generally understood to pursue. In light of the international role of English and its frequent use as a lingua franca, many scholars argue that the overarching aim of pronunciation teaching should be international intelligibility rather than adherence to a native-speaker standard (Wang & Wen, 2023, p. 2; Jenkins, 2000, p. 16). Especially in Expanding Circle contexts, where learners are more likely to interact with other non-native speakers, the ability to be understood has become more relevant than sounding like a native speaker (Wang & Wen, 2023, p. 20; Jenkins, 2002, p. 83).

In light of these considerations, it appears that achieving native-like articulation is not an overarching goal in pronunciation instruction. As established earlier in this thesis (see Section 2.3), speech may be heavily accented yet remain fully intelligible (Pennington & Rogerson-Revell, 2019, p. 133; Derwing & Munro, 2015, p. 6). The distinction between accentedness and comprehensibility illustrates that accentedness refers to perceived deviation from a native norm, whereas comprehensibility concerns how easily a message can be processed by a listener. Hence, interlocuters may perceive a strong foreign accent yet have no difficulty comprehending the message.

Recognising that intelligibility does not depend on native-like pronunciation has led to two divergent orientations in pronunciation instruction: the so-called nativeness principle and the intelligibility principle. While the former promotes native-like pronunciation as both possible and desirable, the latter holds that being understood should be the primary aim of pronunciation instruction (Levis, 2005, p. 370). Importantly, the nativeness principle has been criticised for setting unrealistic expectations, especially for adult learners, which may lead to decreased motivation or linguistic insecurity (Hodgetts, 2020, p. 59; Wagner & Toth, 2017, p. 87). In contrast, the intelligibility principle offers a more inclusive, achievable, and pedagogically sound alternative, especially for learners who primarily use English in international communication (Wang & Wen, 2023, p. 20).

From an instructional perspective, adopting the intelligibility principle requires a context-sensitive approach. To this, Levis (2005, pp. 372–373) explains that intelligibility is shaped not only by the speaker's phonological production but also by the communicative context, including the listener's background, the interactional setting, and the communicative purpose. In other words, what counts as intelligible may vary across contexts which reinforces the need to adapt pronunciation instruction to the specific needs of the learners. Against this backdrop, researchers, such as Jenkins (2002) and Murphy (2014), have called for a reassessment of which pronunciation models should be presented in classrooms. Jenkins (2002, p. 83) proposes that students should be exposed to a range of intelligible non-native accents as part of their learning experience. This can be done through carefully selected listening materials or classroom activities that feature proficient non-native speakers engaged in authentic communication. Murphy (2014, p. 259) supports this view, arguing that such exposure can help learners set more realistic and attainable pronunciation goals, especially in settings where English is used predominantly among non-native users. This approach would not only better reflect the reality of global English use but also provide students with more achievable points of reference.

Although intelligibility is increasingly acknowledged as the overarching goal of pronunciation instruction, both the nativeness and intelligibility principles continue to shape pronunciation curricula and teaching practices (Levis, 2015, p. 371). The persistent dominance of prestige varieties, such as RP and GA, in teaching materials suggests that the shift toward intelligibility-oriented practices has yet to be fully realised. Nonetheless, the growing emphasis on intelligibility reflects a broader reorientation in pronunciation pedagogy, from unattainable native-like targets toward more realistic and communicatively oriented goals. This orientation aligns closely with the realities of English use in international communication, particularly in contexts where English serves as a lingua franca. The following chapter therefore shifts to the concept of ELF, with a particular focus on the Lingua Franca Core, a set of pronunciation features that support international intelligibility.

4 English as a Lingua Franca and the Lingua Franca Core

Having established intelligibility as a guiding principle in pronunciation teaching, this chapter now turns to ELF and its pedagogical implications. As English is increasingly used as a lingua franca among speakers of different L1s, the question arises which pronunciation features are essential for maintaining international intelligibility in such interactions. In response to this, Jenkins (2000) proposed the Lingua Franca Core, a set of empirically grounded features shown to support intelligibility in ELF communication. Jenkins' insights have important implications for pronunciation instruction, especially for learners whose L1 may present specific challenges.

This chapter begins by outlining the concept of ELF. It then introduces the LFC and explores the principles and pedagogical implications of ELF-informed pronunciation teaching, focusing on intelligibility, teachability, and learner identity. Finally, Austrian German is examined in relation to the LFC to identify features that should be prioritised in pronunciation instruction to ensure successful communication.

4.1 Conceptualising ELF

Over the past decades, English has undergone a significant transformation from being primarily the language of its native speakers to serving as a global means of international communication among individuals from a wide range of linguistic backgrounds. This global spread has been driven by historical, economic, and cultural forces, resulting in English being used in increasingly diverse contexts across the world. As a consequence, communication in English now occurs more frequently between non-native speakers than between native and non-native interlocutors (Walker, 2010, p. 6).

This development has given rise to the concept of English as a Lingua Franca, which describes instances where English is used as a shared communicative medium between speakers who do not share the same L1. Notably, ELF interactions are not limited to a particular geographical region or user group (Jenkins, 2007, p. 5) but occur both within and across Kachru's previously discussed Inner, Outer, and Expanding Circles (see Section 3.3. 1) (Seidlhofer, 2011, p. 3). Accordingly, the early view that ELF excludes

native speakers altogether (House, 1999, p. 74; Firth, 1996, p. 240) is no longer widely accepted. Instead, ELF is defined as any use of English among speakers of different L1s, regardless of whether native speakers are involved (Seidlhofer, 2011, p. 7).

As ELF communication typically takes place among speakers with diverse linguistic backgrounds, international intelligibility becomes the key concern (Walker, 2010, p. 7). In such contexts, the ability to understand and be understood takes precedence over accent or fluency. This reorientation of communicative priorities has prompted researchers to examine which aspects of English pronunciation most directly affect intelligibility in international contexts. The subsequent section, therefore, introduces the Lingua Franca Core, a set of phonological features identified as particularly relevant for ensuring successful ELF communication.

4.2 The Lingua Franca Core (LFC)

Given the global spread of ELF, the question of which pronunciation features are essential for ensuring intelligibility among non-native speakers (as discussed in Section 2.3) has received increasing attention. In response to this, Jenkins (2000) introduced the so-called Lingua Franca Core, a set of pronunciation features empirically shown to be crucial for communicative success in ELF interactions. Rather than establishing a new norm or pronunciation model, the LFC provides a framework that refrains from native-speaker norms and aims at achieving intelligibility among diverse English users (Jenkins, 2000, pp. 123-124; Jenkins, 2000, p. 11).

The LFC emerged from Jenkins' exploratory research on naturally occurring ELF interactions, in which she analysed communicative breakdowns between non-native speakers to determine which phonological features most frequently caused misunderstandings. Out of 40 observed breakdowns, 27 were linked directly to pronunciation difficulties, while lexical, grammatical, or other causes played a comparatively minor role (Jenkins, 2000, p. 88). Interestingly, most of the phonological issues stemmed from L1 transfer, whereas suprasegmental errors (e.g., misplaced intonation or stress) rarely affected intelligibility. This, in turn, led Jenkins to question the

long-held focus on suprasegmentals in pronunciation instruction, an emphasis that had dominated the field since the rise of CLT (Kaur, 2017, p. 550; Walker, 2010, pp. 26–27).

At any rate, the key insight of Jenkins' study is that not all features of English pronunciation contribute equally to international intelligibility. Based on her data, Jenkins (2000) proposed a core set of features which should be prioritised in teaching pronunciation to learners who will use English primarily in ELF settings. These features, collectively referred to as the LFC, represent the most frequent causes of intelligibility loss when mispronounced (Jenkins, 2000, p. 132). Drawing on Kaur (2017, p. 549), the LFC does not prescribe correctness but rather identifies non-negotiable phonological features necessary to maintain mutual understanding.

As outlined by Jenkins (2000, p. 132) and summarised by Walker (2010, p. 8), the components of the LFC fall into four broad categories, including:

- most consonant sounds with certain substitutions permissible (e.g., /θ/ and /ð/)
- appropriate consonant cluster simplification
- vowel length distinctions
- nuclear stress placement

These features are further detailed and elaborated in Jenkins' (2000, p. 159) comprehensive table of LFC components, demonstrated in Table 2 below.

Table 2. *The Lingua Franca Core Features proposed by Jenkins (2000, p. 159)*[5]

1) Consonantal inventory with the following provisos	- rhotic [ɻ] rather than other varieties of /r/ - intervocalic /t/ rather than [ɾ] - most substitutions of /θ/ and /ð/, and [ɬ] permissible - close approximations to core consonant sounds generally permissible - certain approximations not permissible (i.e., where there is a risk that they will be heard as a different consonant sound from that intended)

[5] The LFC features listed in this table are not presented in any order of priority or importance.

2) Phonetic requirements	- aspiration following the fortis plosives /p/, /t/, and /k/ - fortis/lenis differential effect on preceding vowel length
3) Consonant clusters	- initial clusters not simplified - medial and final clusters simplified only according to L1 rules of elision
4) Vowel sounds	- maintenance of vowel length contrasts - L2 regional qualities permissible if consistent, but /ɜː/ to be preserved
5) Nuclear stress production and placement and division of speech stream into word groups.	

(Jenkins, 2000, p. 159)

Table 2 highlights that several phonetic decisions are based on practicality and intelligibility. For instance, the realisation of the dental fricatives /θ/ and /ð/ is omitted because, according to Jenkins' findings, their substitution did not lead to reduced intelligibility in the analysed data (Jenkins, 2000, p. 137). Jenkins (2000, p. 138) argues that the acquisition of dental fricatives can be left to learners' individual needs, depending on their specific goals and preferences. Further LFC considerations are evident in the production of /r/. Here, Jenkins prefers the GA rhotic /r/ variant over the non-rhotic RP variant due to its consistency in pronunciation and closer correspondence with English orthography (Jenkins, 2000, p. 139). Similarly, the RP [t] is favoured over the GA flap [ɾ] for intervocalic /t/ because of its phonetic similarity to /d/ (e.g., compare <ma<u>tt</u>er> and <ma<u>dd</u>er>), which increases potential confusion (Jenkins, 2000, p. 140). Overall, these choices reflect Jenkins' effort to select realisations that maximise clarity and minimise phonological ambiguity.

Equally noteworthy are features such as weak forms and schwa, which are traditionally prioritised in pronunciation instruction but are noticeably absent from the LFC. These features are classified as non-core, as they did not lead to communication breakdown in Jenkins' data (Walker, 2010, p. 38). The non-core features, as summarised by Walker (2010, p. 38), include the following:

- /θ/, /ð/, and dark /l/
- exact vowel quality (except for /ɜː/)
- pitch movement (tone)
- word stress[6]
- stress-timing
- vowel reduction, schwa, and weak forms
- certain features of connected speech (linking, assimilation, coalescence)

Within an ELF framework, variation (e.g., L1 transfer) in these non-core areas is not automatically viewed as an error but as acceptable L2 phonological variation (Walker, 2010, p. 38; Jenkins, 2000, p. 159). For instance, if two speakers pronounce the word <sold> using different vowel realisations (e.g., /oʊ/ vs. /əʊ/), this difference is not interpreted as incorrect pronunciation but simply as a reflection of regional or L1-influenced variation, provided intelligibility is maintained (Walker, 2010, p. 38).

Drawing on her empirical findings, Jenkins also challenged the long-standing emphasis on suprasegmental features in pronunciation teaching. While earlier research had considered intonation and rhythm as key to intelligibility (cf. Avery & Ehrlich, 1992; Morley, 1991), Jenkins (2000) found that segmental accuracy, especially of consonants and vowel length, played a significantly greater role in ensuring intelligibility among non-native speakers (Kaur, 2017, p. 550). Moreover, Pennington and Ellis (2000, p. 387) suggest that certain suprasegmental features, such as nuclear stress or prosody, are relatively

[6] Jenkins (2000, pp. 150-151) notes that misplaced word stress rarely causes intelligibility problems on its own, and when it does, it tends to co-occur with other phonological errors. Nevertheless, word stress can indirectly affect intelligibility through its influence on nuclear stress placement and aspiration of word-initial voiceless plosives. As such, she recommends teaching general patterns and dictionary use rather than aiming for full mastery of word stress rules.

accessible to learners, whereas others (e.g., intonation of sentence tags) tend to pose considerably greater difficulty.

Nonetheless, the LFC has also faced criticism as some scholars view it as overly prescriptive, arguing that it risks replacing one norm with another (Berns, 2008, p. 333). Others have questioned the selection of features, suggesting that excluding suprasegmental features, most notably word stress, underestimates their role in intelligibility (Deterding, 2013, p. 8). However, Jenkins (2000, p. 158) emphasises that the LFC is not a fixed model but rather a communicative resource, a tool for learners to rely on depending on the context and their individual communicative goals. It neither prevents learners from developing native-like pronunciation if they wish to, nor forces them to conform to a specific set of features. Rather, it offers a more flexible framework with greater individual freedom to express learners' own identities.

In summary, the LFC clearly represents a shift in pronunciation instruction in a way that it centres intelligibility and responds directly to the realities of global English use. By focusing on empirically established features that impact mutual understanding, the LFC offers a practical and learner-oriented framework, while encouraging context-sensitive language use. To build on this, the subsequent section explores how pronunciation instruction can be informed by such an ELF perspective

4.3 Relevance of the LFC in pronunciation instruction

As emphasised throughout this thesis, intelligibility is nowadays widely recognised as the guiding goal of pronunciation instruction. In line with this focus, an ELF-oriented approach to pronunciation teaching generally seeks to address three main goals: promoting intelligibility over native-like accuracy, focusing on the aspect of teachability, and supporting learner identity (Walker, 2010, p. 20). The previously outlined LFC responds well to these three goals by prioritising intelligible communication while leaving room for learners' individual accent preferences (Walker, 2010, pp. 20-21). This section explores how each of these goals is addressed in the LFC and what this implies for pronunciation instruction in ELF contexts.

A primary strength of the LFC lies in its focus on international intelligibility as a communicative priority. Rather than promoting conformity to native-speaker varieties, the LFC highlights those features, predominantly segmental ones, that are most likely to cause misunderstandings in ELF interactions if mispronounced (e.g., vowel length, initial consonant clusters) (see Section 4.2). It should be noted, however, that the LFC is not intended as a pronunciation model or a prescribed set of norms but as a framework enabling learners to focus on features relevant to the contexts in which they are most likely to use English (Kaur, 2017, p. 550). By targeting these features in pronunciation instruction, the LFC enables learners to develop pronunciation that is functional across diverse interlocutor backgrounds (Jenkins, 2000, p. 131).

Equally important is the aspect of teachability, which further strengthens the pedagogical relevance of the LFC. Jenkins (2000, p. 132) emphasises the importance of distinguishing between what is crucial for intelligibility and what can realistically be taught in the classroom. Drawing on earlier work by Dalton and Seidlhofer (1994, pp. 72–73), some phonological features, such as the distinction between fortis and lenis consonants, are relatively straightforward to teach explicitly. Other features, particularly suprasegmental elements with strong contextual components (e.g., intonation patterns), are far more difficult to isolate and teach systematically. In this regard, Jenkins (2000, p. 133) argues that if a feature is widely unteachable, it becomes irrelevant in the context of ELF. In fact, the LFC includes only features that are both relevant for achieving intelligibility and teachable in the classroom. Along these lines, Walker (2010, pp. 8–9) notes that acquiring as well as teaching LFC features requires significantly less effort and time than developing a native-like accent, making it a realistic goal for learners and teachers with limited instructional time.

In addition to supporting intelligibility and teachability, the LFC also aligns with learners' individual identities and linguistic repertoires. Rather than imposing native-speaker norms as the ideal, the LFC allows for the retention of non-core L1-influenced features that do not hinder intelligibility (see Section 4.2). This flexibility not only validates learners' existing phonological backgrounds but also acknowledges the personal and social dimensions of spoken language. Drawing on Walker (2010, p. 21), the LFC empowers learners to develop a form of English that supports both communicative clarity and self-expression.

Taken together, the three proposed goals (i.e., mutual intelligibility, teachability, and identity) form the foundation of ELF-informed pronunciation teaching for which the LFC offers an evidence-based framework addressing each of these dimensions. By shifting the instructional focus from native-likeness to communicative effectiveness, the LFC encourages a more realistic, inclusive, and learner-centred approach to pronunciation instruction.

4.4 Austrian German and the LFC

Building on the discussion of identity and L1 influence (see Section 4.3), this section explores how specific features of Austrian German may affect learners' production of LFC features in English. In the context of ELF, understanding the phonological characteristics of learners' L1 is essential for identifying potential areas of difficulty that may affect intelligibility. In this thesis, particular attention is given to German, the most widely spoken L1 in the European Union (Berger, 2010, p. 107), with a specific focus on its Austrian variety. Although German encompasses a range of regional accents, certain phonological features are widely shared. Among these, Austrian Standard German diverges from other German varieties in terms of phonetic, lexical, and prosodic features (Ebner, 1998). Notable differences from Standard German include variations in diphthong realisation, vowel closeness, stress placement, and overall prosodic patterns (Richter, 2019, pp. 136–137). While mutually intelligible with Standard German, Austrian German exhibits distinctive traits that may influence EFL learners' English pronunciation. Highlighting these characteristics is particularly relevant to this study, as it allows a more nuanced understanding of the pronunciation challenges Austrian EFL learners may encounter.

To begin with, phonemic overlap between English and German provides Austrian learners with a relatively favourable starting point in certain areas of pronunciation. According to Berger (2010, p. 107), many English consonants, including /tʃ/, /f/, /ʃ/, /h/, /m/, /n/, /l/, and /j/, have equivalent or near-equivalent counterparts in German and are, thus, generally unproblematic for ELF intelligibility. However, research has shown that a number of pronunciation challenges persist among German-speaking learners, many of which are directly related to the previously outlined LFC (Berger, 2010, pp. 108–110; Swan & Smith, 2001, pp. 37–39). These include:

- insufficient aspiration of voiceless plosives /p/, /t/, and /k/ in word-initial position (e.g., town vs. down)
- devoicing of lenis consonants /b/, /d/, /g/, and /z/ in final position due to German Auslautverhärtung (e.g., rib vs. rip)
- difficulty with vowel quantity contrasts, particularly the lengthening before lenis and shortening before fortis consonants
- inaccurate production of the long central vowel /ɜː/, often substituted with the diphthong /øə/ (Kaltenböck & Seidlhofer, 1993, p. 16)
- devoicing of /z/ and confusion between /s/ and /z/, indicating a lack of phonemic distinction in the L1
- replacement of /ʒ/ and /dʒ/ with /ʃ/ and /tʃ/, as the voiced counterparts do not exist in Austrian German
- confusion between /v/ and /w/, with frequent substitution in both directions due to L1 influence or overgeneralisation
- use of the uvular fricative [ʁ] for /r/, which Berger (2010, p. 109) recommends avoiding in ELF contexts due to its potential impact on clarity

Indeed, many of these commonly observed difficulties are identified as essential for mutual intelligibility in the LFC, such as vowel length distinction, aspiration of voiceless plosives, and the exact realisation of the long central vowel /ɜː/ (see Section 4.2). Since these features contribute directly to phonological intelligibility, persistent inaccuracies in their production can hinder mutual understanding in ELF interactions. Hence, while certain deviations rooted in L1 transfer are acceptable as part of ELF variation, others remain critical for successful communication and should be prioritised in EPI for Austrian EFL learners.

Empirical studies reinforce these findings regarding challenging areas for these learners. Interestingly, Wieden and Nemser (1991, p. 54-58) identified in their large-scale longitudinal study of 384 Austrian pupils that vowel production is considerably more problematic compared to consonants. Particularly low accuracy was reported for central and open vowels like /ɜː/ and /ʊə/, with diphthongisation and substitution patterns. The researchers attributed these difficulties primarily to L1 transfer and challenges with vowel quantity in general, especially among high vowels (e.g., /iː/, /ɪ/, /uː/, /ʊ/). Later studies (c.f.

Tatzl, 2011; Mende, 2009; Hrubes, 2008) confirm these widespread issues with consonant devoicing, lenis/fortis distinctions, and imprecise realisations of English vowels and diphthongs.

While some of these errors reflect acceptable phonological variation, particularly with regard to non-core features, others may hinder intelligibility and should, therefore, be prioritised in pronunciation instruction. This is especially true for errors that affect phonemic contrasts relevant to meaning, such as <sad> vs. <sat>. These examples illustrate how difficulties with fortis/lenis consonants may also affect vowel length (Richter, 2015, pp. 114-115).

To sum up, although L1 Austrian German EFL speakers share several cross-linguistic similarities with English, they also encounter challenges with specific LFC features that may affect intelligibility. These difficulties, most notably with vowel length, aspiration, and lenis/fortis distinctions, can largely be traced to L1 transfer and highlight the importance of EPI. However, not all variation in learners' pronunciation can be explained by their linguistic background alone. Even among speakers with the same L1, considerable differences in the retention and development of pronunciation features can be observed. To better understand these individual trajectories, the following chapter explores a range of cognitive, social, and experiential factors that may influence learners' ability to acquire and maintain pronunciation features over time.

5 Factors influencing pronunciation development

Building on the previous discussion of Austrian German and its potential influence on the production of key LFC features, this chapter turns to additional factors that may shape learners' pronunciation development. While phonological transfer from the L1 is undoubtedly important, it does not fully account for the variation observed among EFL learners who share similar linguistic backgrounds. Such variability is frequently attributed to individual difference (ID) variables, a broad set of cognitive, social, and affective influences that may affect how learners perceive, produce, and retain sounds of a second or foreign language.

With regard to factors beyond instruction, individual and contextual variables, such as learners' L1, motivation, attitudes and identity, and exposure to English outside the classroom, may be just as influential as instruction itself in shaping learners' ability to retain acquired pronunciation features over time (Gilakjani, 2012, p. 122). These variables were already identified as influential in early research on pronunciation learning (cf. Kenworthy, 1987) and have since been confirmed in numerous subsequent studies (e.g., Derwing & Munro, 2015; Moyer, 2013), making them highly relevant to the present study. While the analysis in this research study focuses on changes in the production of segmental LFC features over time, it also seeks to understand which learner-related variables may have contributed to their retention or decline. To investigate such influences, participants were asked to complete a questionnaire designed to elicit self-reported insights into their language use, attitudes, and practices following instruction.

Specifically, this chapter explores four key factors that the literature consistently identifies as key influences on L2 pronunciation development, including: learners' L1, their motivation, their attitudes and sense of identity, as well as their exposure to and use of English outside the classroom. Although these ID factors are inherently interconnected (Richter, 2018, p. 352), they are examined individually to allow for a more systematic and focused discussion.

5.1 L1

The first ID factor considered in this chapter is the influence of the learner's L1 on the L2, which plays a central role in shaping L2 pronunciation development. As Schmitt (2016, p. 51) notes, the L1 is "easily the most salient factor for others and one of the most studied" in pronunciation research. This influence can be both facilitative and obstructive, as it affects how learners perceive and produce the sounds of the L2. From a listener's perspective, a foreign accent is often primarily attributed to L1-based phonological transfer, commonly expressed through segmental substitutions, suprasegmental deviations, or other constraints that reflect the learner's L1 phonological system.

In general terms, L1 influence in pronunciation learning refers to the transfer of phonological features from the learner's native language to the target language. In this

regard, Avery and Ehrlich (1992, p. 15) already stressed in early work that "[…] pronunciation errors that L2 learners make are not just random attempts to produce unfamiliar sounds. Rather, they reflect the sound inventory, rules of communication, and the stress and intonation patterns of the native language." This insight is foundational to understanding both the challenges and recurring patterns observable in L2 pronunciation.

Building on this, research has shown that EFL learners often draw on their L1 phonological repertoire when encountering unfamiliar L2 sounds, which can have both facilitating and interfering effects (Celce-Murcia et al., 1996, pp. 19–20). In the literature, this phenomenon, commonly referred to as "interference", "transfer", or "cross-linguistic influence", has been conceptualised in various ways. Historically, the term "interference" was associated with disruptive effects of the L1 on L2 performance, particularly when learners produced forms that deviated from target-like norms (Schmitt, 2016, p. 51; Flege, 1995, p. 241). However, due to its negative connotations, researchers have increasingly favoured more neutral and inclusive terms such as "transfer" or "cross-linguistic influence" (cf. Jarvis & Pavlenko, 2010; Odlin, 1989). These broader terms better capture the fact that L1 influence can be both negative (i.e., interfering) and positive (i.e., facilitating), depending on the degree of similarity between the L1 and L2 (Schmitt, 2016, pp. 51-52).

Positive transfer occurs when structural similarities between the L1 and L2 support pronunciation acquisition in the L2. For instance, German and English share several consonant phonemes (e.g., /b/, /d/, /f/, /s/) which reduces the need for German-speaking learners to develop entirely new articulatory patterns for these sounds (Schmitt, 2016, p. 52). Due to these overlaps, German speakers tend to find English relatively accessible in the early stages of pronunciation learning (Swan & Smith, 2001, p. 37). Positive transfer can, thus, foster pronunciation development, at least in areas where the phonological systems of both languages overlap.

In contrast, negative transfer, also referred to as interference, arises when features from the L1 conflict with those of the L2, leading to persistent errors or fossilised forms. This influence is of particular relevance in pronunciation as mismatches at the segmental and suprasegmental level can disrupt intelligible production (Mairano et al., 2018, p. 45). For example, differences in vowel length, syllable structure, or stress-timing frequently pose

challenges for German learners and can affect intelligibility if left unaddressed (Mairano et al., 2018, pp. 45–46). Importantly, such L1 transfer is not limited to individual sounds but may extend to broader phonological patterns, including intonation and speech rhythm (Mairano et al., 2018, p. 46-47). Thus, without EPI, these L1-induced patterns may become fossilised, even in advanced learners (Chaira, 2015, p. 481).

In light of these challenges, identifying common difficulties experienced by different L1 learner groups can be helpful in anticipating recurring problem areas and ensuring a more targeted pronunciation instruction (Setter & Jenkins, 2005, p. 3). In the context of the present study, the learner group consists predominantly of L1 Austrian German speakers who are likely to exhibit similar transfer patterns due to shared phonological influences. As such, their L1 background offers a valuable starting point for anticipating potential areas of difficulty in the production of L2 sounds. While L1 transfer plays a central role in shaping pronunciation learning, it is only one of several variables that contribute to the complex process of phonological development. The following sections, therefore, turn to additional factors that also influence how learners acquire and retain pronunciation features over time.

5.2 Motivation

Another influential ID variable which calls for closer examination is motivation. Traditionally, the psychologically driven and rather complex factor motivation, long regarded as central in SLA, has been described as “the primary impetus to initiate language learning” (Dörnyei, 1998, p. 117). Expanding this definition, Moyer (2004, p. 39) frames motivation as a multifaceted construct involving a learner’s conscious effort, intention, and planning toward a specific goal. In line with this view, research on pronunciation learning has shown that motivation contributes not only to learners’ pronunciation development but also to how much effort they are willing to invest in practicing pronunciation (Sardegna & Jarosz, 2023, p. 13).

Research into pronunciation learning defines motivation most commonly as intrinsic, where learners are driven by personal interest or enjoyment, as well as extrinsic, where the drive stems from external rewards (Saito et al., 2017, p. 618). Among university

students in particular, these motivational types are often shaped by instrumental reasons, such as improving job prospects or study opportunities, or integrative reasons, such as building relationships or engaging with English-speaking cultures (Saito et al., 2017, pp. 615-618; Moyer, 2013, p. 68). Hence, these motivational orientations often intersect with learners' personal and professional goals, which can help determine how much emphasis students place on improving their pronunciation. For example, students aiming to work in international contexts or professions that require frequent spoken English use may prioritise pronunciation more strongly than learners without such ambitions. Learners' personal and professional goals are, therefore, closely tied to their motivational intensity (Sardegna & Jarosz, 2023, p. 22; Gilakjani, 2012, p. 123). In support of this, Saito et al. (2017, pp. 627-629) found that Japanese EFL university students with strong instrumental goals improved more in terms of comprehensibility than those without such goals. This aligns with findings by Afshari and Ketabi (2016, p. 76), who observed that even highly motivated learners may not aim for native-like pronunciation but instead focus on intelligibility and successful communication as a personal pronunciation goal.

Further evidence concerning the role of motivation in pronunciation learning stems from longitudinal studies. In the Austrian context, Smit and Dalton (2000, p. 238) conducted a series of studies involving advanced EFL learners, which revealed a high level of motivation to improve pronunciation, with students citing both intrinsic and extrinsic reasons for striving towards near-native pronunciation. Their findings reinforce the idea that motivation, especially in its intrinsic form, can significantly shape learners' pronunciation skills. In a more recent study, Sardegna and Jarosz (2023, p. 23) reported that highly motivated EFL learners who continued to practice independently after EPI maintained significant gains in suprasegmental features, such as linking and sentence stress. In this context, Moyer (2014, p. 431) highlights that it is not only the type of motivation that matters, but also the intensity and consistency of that motivation over time. These findings emphasise the lasting impact motivation can have on learner's pronunciation, especially when practice extends beyond formal instruction. In light of this, the present longitudinal study also considers whether self-reported motivational factors contributed to the participants' ability to retain their pronunciation skills after completing an EPI course.

5.3 Attitude and identity

In addition to motivation, learners' attitudes and sense of identity have also been recognised as influential ID factors in L2 pronunciation development. In pronunciation research, attitude generally refers to learners' evaluative beliefs, feelings and preferences toward different aspects of L2 pronunciation, including accent varieties, pronunciation goals or the value placed on intelligibility (Tsunemoto & McDonough, 2020, p. 26). Identity, by contrast, is understood as the way learners perceive themselves and wish to be seen through their use of the language (Moyer, 2013, p. 65; Setter & Jenkins, 2005, p. 5). These closely related constructs are discussed together in this section, as learners' attitudes toward pronunciation are often shaped by their linguistic identity, both of which are directly relevant to how participants in the present study reflect on their pronunciation development.

A growing body of research highlights the link between positive attitudes toward the target language and greater success in pronunciation development. In fact, learners' views of the language can influence their pronunciation development in that they may attempt to approximate a particular accent (Richter, 2018, p. 343; Afshari & Ketabi, 2016, p. 76; Moyer, 2013, p. 70). Thus, learners who value the target language and its community, and who consider accurate pronunciation an important aspect of communication, are typically more likely to invest effort in practising and refining their L2 pronunciation (Moyer, 2013, pp. 71-72). This increased effort, in turn, may contribute to greater phonological accuracy and long-term improvement (Suzukida, 2021, pp. 52–53). These observations emphasise that pronunciation development is not only shaped by instructional factors but also by learners' underlying attitudes toward the language and its speakers.

In addition to attitudes, learners' identity has emerged as another key influence on pronunciation development. While some learners may aspire to sound like native speakers, others may consciously choose to retain elements of their L1 accent as a marker of identity (Jenkins, 2000, p. 16). For instance, Porter and Garvin (1989, p. 11) found that many L2 users preferred to maintain their L1-accented speech in order to preserve their linguistic background. This tendency is particularly evident in Expanding Circle contexts (see Section 3.3.1), where English is used as a foreign language but is not

used as an L1. In such settings, a local accent may be seen not as a deficiency but as a valid expression of ownership over English as a global resource (Walker, 2010, pp. 19–20). What is more, learners' pronunciation choices are often shaped by their communicative priorities. While some prioritise intelligibility, others strive for greater phonological precision as a way of aligning themselves more closely with the target community (Nunan, 1991, p. 107).

Clearly, pronunciation development is deeply intertwined with learners' attitudes and identities. Whether striving for native-like pronunciation, prioritising intelligible ELF communication, or preserving a distinctive accent that reflects the L1 background, EFL learners make pronunciation-related decisions that extend far beyond phonetics. Against this backdrop, the relevance of attitude and identity becomes especially clear when considering the third research question of this study, which explores self-reported factors contributing to long-term changes in the production of selected segmental LFC features. Differences in learners' attitudes and identity orientations may have influenced not only their engagement during instruction but also the extent to which they retained these features over time.

5.4 Out-of-class exposure and use

The final key factor considered in this chapter is the extent to which learners are exposed to and actively use the target language beyond the classroom. In the present study, these aspects were explored through a questionnaire which elicited participants' self-reported engagement with English outside formal instruction. Items targeting exposure focused primarily on learners' engagement with English-language media such as films, podcasts, or music, while items on use addressed opportunities to speak English in authentic settings, for example during travel, study abroad, or social interactions outside university. Both variables are crucial in supporting advanced learners' L2 pronunciation development beyond formal instruction, as they allow for engagement with varied phonological input and target language practice in real-world contexts (Vong et al., 2024, p. 988).

One particularly influential source of such exposure and use of English is global mobility, most notably through study abroad programmes. In such environments, immersion in an English-speaking context offers rich and authentic input and frequent opportunities for spontaneous spoken interaction. Nagle and Zárate-Sández (2024, p. 542) argue that such real-life communicative experiences foster more dynamic pronunciation learning than classroom instruction alone. Similarly, Martinsen et al. (2014, pp. 73–74) found that learners of Spanish as a foreign language who had spent extended time (e.g., two years) abroad demonstrated greater convergence toward native-like pronunciation than peers without such experience. Although their study focused on Spanish rather than English, the findings underline the potential of sustained, meaningful language use to improve pronunciation. However, not all research on study abroad points to positive outcomes. Some studies have reported limited to no measurable gains in learners' pronunciation development (cf. Avello, 2018, pp. 148-149). This suggests that while staying abroad can provide favourable conditions for pronunciation learning, its effectiveness depends on factors such as length of stay, intensity of interaction, and individual learner differences. What is more, even prolonged exposure may not lead to full native-like speech, especially once learners have achieved intelligibility sufficient for communicative success (Martinsen et al., 2014, p. 74).

In addition to student mobility, digital media has emerged as another key channel for exposing EFL learners to spoken English and enabling active language use. Through global streaming platforms, social media, and user-generated content, learners now encounter English across both formal and informal registers. Such exposure can enhance receptive familiarity with English phonology and encourage spoken output, for instance through online discussions or content creation. As for media-based exposure, Uchihara et al. (2022, pp. 33–34) found that frequent listening via digital media improved learners' comprehensibility, although its impact on productive aspects of pronunciation, such as reducing foreign-accentedness, was less pronounced. Interestingly, learner perceptions align with these findings. In a recent study on Cambodian EFL students, Vong et al. (2024, p. 992) reported that most participants perceived regular English exposure and opportunities to use the language, through conversation or media engagement (e.g., films, music), as beneficial for their pronunciation development.

While neither exposure nor use alone can ensure sustained pronunciation development, both appear to play a decisive role in supporting intelligibility and phonological refinement, particularly when combined with positive attitudes and favourable learner identities. As Afshari and Ketabi (2016, p. 76) point out, learners who are socially engaged and confident in their language use are more likely to take advantage of these opportunities, which links these factors back to the attitudinal and identity-related influences discussed in the previous section. Considering the focus of this study, the addressed ID factors may help explain why learners differ in their long-term retention of pronunciation features. This issue is examined in more detail in Part II, which draws on participants' self-reported experiences and reflections gathered in response to the third research question.

PART II

6 The study – design, methodology and results

Having laid the theoretical framework in the preceding chapters, part two of this thesis shifts its focus to the empirical investigation of the long-term impact of EPI. As highlighted in the literature, relatively few studies have examined the durability of EPI gains over extended periods, particularly among advanced EFL learners with Austrian German as their L1. This study seeks to address this gap by tracking Austrian university students' pronunciation development and retention over time. The chapter begins by outlining the purpose of the present study and the research questions guiding the investigation. This is followed by a detailed description of the study's research design and methodology. The core section of this chapter presents the findings and interprets them in light of previously reviewed research. Finally, the chapter concludes with a discussion of the study's limitations and a brief outlook on directions for future research.

6.1 Purpose of the study and research questions

The overarching aim of this study is to examine the long-term impact of EPI on the production and retention of selected segmental features of the LFC among advanced Austrian EFL learners at the University of Vienna. While considerable research has documented the benefits of EPI for promoting intelligibility-focused pronunciation (see Section 2.2), much of this work has focused on immediate post-instruction gains. Indeed, a number of empirical studies confirm that even relatively brief periods of EPI can lead to measurable improvements in learners' segmental and suprasegmental accuracy (cf. Gordon & Darcy, 2022; Sardegna & Dickerson, 2023).

Despite these encouraging findings, important research gaps remain. First, relatively little is known about the long-term retention of pronunciation gains following EPI. As Nagle (2021, p. 154) notes, longitudinal research into L2 pronunciation remains limited, with only few studies including delayed post-tests or extending beyond a single academic semester. Second, comparatively few studies have addressed the needs of (Austrian) German L1

learners at the tertiary level. Although existing literature acknowledges that German-speaking learners face persistent challenges with segmental and suprasegmental features (see Section 3.4), many of which are known to affect intelligibility, the extent to which EPI can help address these difficulties remains largely underexplored. Third, pronunciation development is shaped by a range of individual and contextual factors, including learners' motivation, practice habits, and exposure to English (see Section 4.1). While these variables are well established in theory, learners' own perspectives on what influences their pronunciation over time have received limited attention. Exploring these self-reported experiences may, therefore, provide valuable insights into the retention or attrition of pronunciation features gained through EPI.

To address these gaps, the present study traces the development and retention of two selected segmental LFC features over three testing points: before EPI (T1), immediately after one semester of EPI (T2), and again three semesters (i.e., one year and four months) after the end of EPI (T3). To guide this investigation, the following research questions have been proposed:

1. How does EFL university students' pronunciation of two selected segmental LFC features develop over one semester of explicit pronunciation instruction, as measured by a pre-test (T1) and post-test (T2)?

2. To what extent do EFL university students retain these two selected segmental LFC features acquired during one semester of explicit pronunciation instruction, based on a comparison between immediate post-test (T2) and delayed post-test (T3) results?

3. Which self-reported factors may have contributed to changes in the production of the selected segmental LFC features over time?

By addressing these questions, this study aims to offer new empirical insights into the long-term effects of EPI in an under-researched learner population, namely advanced L1 Austrian German speakers. By focusing on LFC features, specifically those that are most crucial for international intelligibility, the study explores how key features influencing

intelligibility develop as a result of EPI over time. Additionally, the inclusion of self-reported learner perspectives helps to interpret the observed changes in relation to learners' individual experiences. Collectively, these insights aim to shed light on the lasting impact of EPI and contribute to a more nuanced understanding of pronunciation retention in EFL learning.

6.2 Methodology

Having defined the research aims and questions, this section outlines the methodological considerations of the study. First, the research design is presented, followed by a detailed description of the instructional setting, the participant group and research instruments. Finally, this section also provides an account of how data was collected and subsequently analysed.

6.2.1 Research design

The study at hand can be classified as a longitudinal panel study, tracking participants' development of selected segmental LFC features over a period of approximately one year and seven months (i.e., March 2023 to October 2024) across three testing phases. Given this extended timeframe, the study meets the criteria for longitudinal research. While there is no universally agreed benchmark for defining longitudinal study length, Nagle's (2021, p. 162) review of 39 longitudinal pronunciation studies found that the average duration is approximately 11 months (M = 10.71). The current study, with a duration of 19 months, thus, clearly qualifies as longitudinal. Beyond meeting this temporal criterion, a key advantage of longitudinal research lies in its potential to capture change over time with greater depth and reliability compared to cross-sectional studies (Gorard, 2001, p. 86). Furthermore, since data in this study were collected from the same individuals across all three testing phases, the research design also qualifies as a panel study (i.e., a longitudinal study where data are collected from the same individuals at multiple time points) (Cohen et al., 2007, p. 212). Panel studies are particularly effective for tracking developmental patterns at the individual level (Cohen et al., 2007, p. 215), and provide a robust, non-experimental approach for exploring change without requiring the use of a control group (Dörnyei, 2007, p. 83).

With this in mind, it is worthwhile to clarify that this study is indeed not experimental in nature. No control group was included, both for ethical and practical reasons. While Thomson and Derwing (2015, pp. 337-338) acknowledge that comparison groups can provide useful insights, they also stress that they are not always essential, particularly when dealing with advanced adult learners enrolled in language programmes, where spontaneous improvement without targeted EPI is unlikely. In the case of the present study, all participants were enrolled in a mandatory EPI university course as part of their degree programme. Creating a control group by withholding this instruction would, therefore, have been neither feasible nor ethically justifiable. In line with this rationale, the study does not seek to evaluate the effectiveness of a specific instructional intervention but rather to observe how selected segmental LFC features develop over time within an existing instructional setting. Accordingly, no attempts were made to control the quality or quantity of instructional treatment, and no variables were manipulated proactively. Instead, the instructional setting examined reflects a standard university-level EFL context rather than a tightly controlled experimental environment. In general, the chosen research design is well-suited to the aims of this study, offering an appropriate foundation for observing longitudinal development in a real-world educational setting.

To examine participants' pronunciation development, speech samples were collected at three distinct points (T1, T2, T3). The recordings were elicited through a controlled and standardised reading passage, a method widely used in pronunciation research as it enables comparability across participants and time points. As Schmitt (2016, p. 196) notes, the use of standardised texts allows for consistent tracking of pronunciation features, enabling the observation of individual student progress. Given that longitudinal comparability is a primary aim of this study, the read-aloud task was determined to be the most suitable instrument for eliciting consistent data across all three phases.

In addition to the speech samples, the study also employed a web-based survey comprising both closed- and open-ended questions. Accordingly, a mixed-methods approach was chosen to collect both quantitative and qualitative data (Dörnyei & Dewaele, 2022, pp. 12-13; Creswell & Plano Clark, 2018, p. 3). The use of a web-based format was primarily motivated by reasons of practicality and efficiency. Moreover, a well-designed questionnaire enables rapid and relatively straightforward data processing (Dörnyei &

Dewaele, 2022, p. 7; Dewaele & McCloskey, 2015, p. 229). To ensure the clarity and quality of the instrument, the survey items were reviewed by my thesis supervisor prior to distribution, which is in line with recommendations by Dörnyei and Dewaele (2022, p. 58).

Taken together, the outlined research design enables a comprehensive investigation of participants' pronunciation development and self-reported factors that potentially contribute to these changes, without exceeding the scope of this research project.

6.2.2 Setting

The present study was conducted at the Department of English and American Studies at the University of Vienna and targeted advanced Austrian EFL learners pursuing either a Bachelor of Education (BEd) or Bachelor of Arts (BA) degree in English. More specifically, it focused on students who attended the EPI course 'Practical Phonetics and Oral Communication Skills 1 (PPOCS 1) and its mandatory tutorial, the Language Lab, during the summer semester 2023 (i.e., March – June 2023). PPOCS 1 is usually taken by students in the fourth semester (Richter, 2021, p. 89) and is the only EPI-focused course offered in the bachelor's programmes at the University of Vienna.

PPOCS 1 is a two-hour weekly pronunciation course offering EPI. Prior to enrolment, students choose either BE or GA as their preferred model which they aim to approximate throughout the semester. The course provides explicit instruction on practical phonetics, targeting segmental and suprasegmental features of the selected variety of English (Richter, 2021, p. 90). Particular emphasis is also placed on addressing pronunciation difficulties commonly observed among Austrian EFL learners of English. On the segmental level, these include challenges such as voicing of lenis consonants, vowel length distinctions, and the production of sounds that do not exist in Austrian German. These features are not only problematic for L1 Austrian German learners but are also known to play a key role in ensuring intelligibility in oral communication (see Section 3.2). On the suprasegmental level, instruction focuses on features like word and sentence stress, which often require increased attention among learners from Austrian German-speaking backgrounds. As far as course activities are concerned, students engage in both reading aloud and free speaking tasks, which are designed to practise and consolidate

the features introduced in PPOCS 1 (Richter, 2021, p. 94). In addition to the main course, students attend a mandatory two-hour language lab aligned with the chosen variety, which offers further opportunities for pronunciation practice as well as regular feedback from a tutor (Richter, 2021, p. 91). The language lab is also where I, along with two other tutor colleagues, work as pronunciation tutors, each teaching two groups of approximately 15 students per semester. This direct involvement provided me with convenient access to a large number of advanced EFL learners enrolled in the PPOCS 1 course. Overall, the main course (PPOCS 1) and the language lab amount to four hours of EPI per week.

Although PPOCS 1 is oriented toward a native-speaker variety, the course does not promote native-likeness as an explicit goal. What is more, there has been an increasing emphasis on critical engagement with the sociolinguistic realities of global English use. As outlined by Richter (2021, p. 96), students are introduced to concepts such as ELF and the LFC through course materials and in-class discussions, promoting reflection on global English use. While the course remains variety-specific in its instructional approach, its content is nonetheless compatible with the LFC-informed perspective adopted in this study, in that it raises awareness of intelligibility-oriented principles.

Overall, the PPOCS 1 setting, with its focus on EPI, offers a suitable context for investigating the long-term development of selected segmental LFC features. Although the course does not explicitly adopt an intelligibility-based approach, its content and aims aligns well with the objectives of this study.

6.2.3 Participants

As outlined in the previous section, the present study was conducted with BEd and BA students enrolled in the PPOCS 1 course and its accompanying language lab at the University of Vienna. Based on the course description of PPOCS 1 (Richter, 2021, p. 90), learners at this stage are expected to have at least an advanced B2 proficiency level (CEFR, 2020). The participant group, thus, represented a relatively homogeneous sample of tertiary-level EFL learners with comparable instructional backgrounds. All participants had opted for the BE variant of PPOCS 1 and attended the corresponding BE language lab. This ensured that all students received the same quantity and quality of EPI.

Initially, all 80 students, who attended the two courses during the summer semester 2023, were invited to participate in the study. Of these, 48 completed both the pre-test (T1) at the beginning of the semester (March 2023) and the post-test (T2) at its end (June 2023). In October 2024, approximately 16 months after T2, efforts were made to contact these 48 students again and invite them to participate in a delayed post-test (T3). Ultimately, 28 participants agreed to complete all three testing phases. As is typical in longitudinal research, this moderate degree of participant attrition was anticipated (Cohen et al., 2007, p. 214).

The final sample, thus, consisted of 28 participants who completed all three tests (T1, T2, T3) and had taken the BE PPOCS 1 and the language lab. Based on a background questionnaire completed during T3, 60.7% (*n* = 17) were aged between 19 and 24 years, 32.1% (*n* = 9) between 25 and 30, and one participant each belonged to the 31–34 and 35+ age groups. The vast majority of the participants (85.7%, *n* = 24) reported to have German as their L1, while the remaining 14.3% (*n* = 4) were bilingual with German and either Dutch, Hungarian, or Swedish. Thus, all participants spoke German as one of their L1s, and none identified English as their L1. Regarding their language learning backgrounds, participants reported that they began learning English between the ages of 2 and 12. Most began formal instruction at around age 10 (*n* = 10), which typically marks the transition to secondary education in Austria. Another 28.6% (*n* = 8) reported starting at age 6 or 7, corresponding to primary school level. Further background data reveals that only four participants spent time in an English-speaking country after completing PPOCS 1. One reported a five-month stay in Australia (Canberra), and the remaining three spent one week each in the United Kingdom. Table 3 below provides an overview of the participants' socio-demographic information collected via the web-based survey administered during T3.

Table 3. *Socio-demographic information of participants (N = 28)*

age	***n***	**%**	**L1**	***n***	**%**	**age started learning E**	***n***	**%**	**time abroad after PPOCS1**	***n***	**%**
19-24	17	60.7	German	24	85.7	2-5	5	17.9	none	24	85.7
25-30	9	32.1	German & Hungarian	2	7.1	6-7	8	28.6	1 week (UK)	3	10.7
31-34	1	3.6	German & Swedish	1	3.6	9-10	14	50	5 months (Australia)	1	3.6
35+	1	3.6	German & Dutch	1	3.6	12+	1	3.6			

As can be seen in Table 3, the participant group was predominantly composed of L1 Austrian German speakers. This linguistic background aligns well with the aims of this study, which seeks to address the notable lack of research into the pronunciation development of Austrian German-speaking EFL learners.

6.2.4 Instruments

In line with the study's research objectives, two instruments were utilised to collect data: a diagnostic read-aloud task and a web-based survey. The read-aloud task was used to gather speech samples at three time points (T1, T2, and T3), while the survey was administered only during the final phase (T3), after participants had received no further EPI. Prior to each phase, participants were informed about the study's purpose, procedures, and ethical considerations. Participation was voluntary, and participants were reminded that they could withdraw at any time without any consequences (Cohen et al., 2011, p. 80). All gathered data were anonymised and treated confidentially throughout the research process. An overview of the overall data collection timeline is provided in Section 6.2.5.

The primary instrument employed in this study was the controlled, diagnostic read-aloud task, in which participants were asked to read aloud and record a short text passage during each of the three testing phases. Such read-aloud activities are, in fact, the most common form of assessment, as highlighted by Thomson and Derwing (2015, p. 326). While reading aloud differs from spontaneous oral production in that it reduces the

cognitive demands associated with grammatical and lexical retrieval, it allows for analysis of specific phonological features in a controlled context (Derwing & Munro, 2015, p. 88). Although spontaneous speech data might have provided additional insights into students' pronunciation development, its inclusion was beyond the practical scope of this thesis. Hence, a read-aloud format was chosen to ensure consistent data collection and comparability across all three testing points. To this end, the 110-word text passage *The North Wind and the Sun* (Figure 3) was selected due to its established use in phonetic research (Deterding, 2006, p. 187):

> *The North Wind and the Sun were disputing, which was the stronger, when a traveller came along wrapped in a dirty cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew, the more closely did the traveller fold his cloak around him; and at last the North Wind stopped. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.*

Figure 3. *The North Wind and the Sun*

The dataset comprised 84 recorded samples in MP3 format, derived from 28 participants who produced one recording per testing phase (T1, T2, T3). In order to maintain consistency and enable longitudinal comparison, participants read the same text passage for each testing phase (T1, T2, T3). While *The North Wind and the Sun* passage is well-suited for eliciting comparable speech data, it is not without limitations (Deterding, 2006, p. 187). As evidenced by Figure 3, all English monophthongs are represented. However, certain other phonological environments are underrepresented or entirely absent. For instance, the passage does not include any examples of aspirated fortis plosives /p/ in word-initial position. It further lacks instances of word-final consonant clusters ending in /s/ or /z/, which limits the examination of segmental features such as final fricative voicing. This limitation is particularly noteworthy, as the devoicing of word-final voiced fricatives is a commonly observed phenomenon among Austrian learners of English and directly links to key LFC features. However, since the passage does not include /s/ or /z/ in such phonological environments, it does not allow investigation into how learners deal with such

clusters (Deterding, 2006, pp. 188–189). In addition, the passage contains a relatively high degree of lexical repetition. Of the 110 words, only 64 are distinct, meaning that approximately 42% are repeated items. For example, the content words *North*, *Wind*, *traveller*, and *cloak* each appear four times. While such repetition may initially seem limiting, it still provides useful data on pronunciation consistency and variation across different phonetic environments. However, it must be noted that some English phonemes are not represented in the passage at all, which limits the overall range of analysable segmental features (Deterding, 2006, pp. 189–190). Nonetheless, the passage remains a widely used diagnostic tool in phonetic research and was deemed adequately representative for the purposes of this study, as it includes a broad range of English monophthongs and consonants relevant to the targeted LFC features.

Based on a detailed review of the reading passage and insights from previous research, two segmental LFC features were selected for closer analysis, namely vowel length in long monophthongs and word-initial aspiration of the fortis plosives /t/ and /k/. These features were chosen for several reasons. First, these segmentals are widely recognised as particularly challenging for Austrian EFL learners of English (Berger, 2010, pp. 108–110; Swan & Smith, 2001, pp. 37–39), who often fail to maintain appropriate vowel length distinctions and tend to lack aspiration in word-initial plosives. Second, both features play a central role in maintaining intelligibility in ELF communication and are, therefore, prioritised in the LFC framework (see Section 3.2). Given their phonological relevance, examining how these segmental features change over the course of 14 weeks of EPI, and how well they are retained over time, is worth investigating within the context of this longitudinal study. Table 4 provides an overview of the frequency and distribution of these two selected LFC features in the diagnostic text passage employed at T1, T2, and T3.

Table 4. *Distribution of selected segmental LFC features in The North Wind and the Sun*

Selected segmental LFC features	Total number of tokens	Loaded words
vowel length in long monophthongs	18	*North, disputing, dirty, agreed, first, succeeded, blew, hard, more, last, warmly, immediately, two*
word-initial aspiration of /t, k/	17	*traveller, came, cloak, take, considered, could, closely, took, to, confess, two*

As shown in Table 4, the distribution of the selected segmental LFC features is well balanced, comprising 18 tokens of long monophthongs across 13 distinct words and 17 tokens of word-initial fortis plosives across 11 distinct words, as some words occurred repeatedly in the text. This provides sufficient data for conducting longitudinal comparisons across T1, T2 and T3.

In addition to the read-aloud task, an online questionnaire was administered at T3 to help address the third research question. The survey was created using Google Forms and aimed to gather both quantitative and qualitative data on participants' self-reported factors related to their pronunciation development following EPI. To match survey responses with the corresponding audio recordings, participants were asked to enter their matriculation number and create a personal code consisting of the first letter of their first name and last name, and the first two digits of their birthdate (e.g., FM22). This ensured accurate data matching while maintaining a manageable level of anonymisation of the data. Such an approach also aligns with established research practices that utilise anonymous identifiers for matching data in longitudinal studies (Ripper et al., 2017, p. 753). While the personal participant codes were used to match questionnaire responses with recordings, all participants were later assigned anonymised codes (R1–R28) for reporting the results to ensure full anonymity.

The questionnaire consisted of five clearly labelled sections for ease of navigation and included a total of 21 items, most of which correspond to the individual and contextual factors discussed earlier in the thesis (see Chapter 5). The first section comprised primarily factual questions (Dörnyei, 2003, p. 8) that collected basic demographic information such as age and L1 background. This was followed by another set of questions addressing participants' perceptions of PPOCS 1, including their experiences and learning goals. The third section focused on post-instruction pronunciation practice habits, which also provided insights into learners' motivation to continue improving their pronunciation beyond the EPI course. The fourth section explored learners' perceived exposure to and use of English following PPOCS 1 and included attitudinal questions, for which participants expressed their level of agreement using Likert-scale items (Dörnyei, 2003, p. 8). Finally, the last section, labelled 'Other Factors', invited participants to report on additional aspects they perceived as having influenced their pronunciation since completing PPOCS 1. The majority of the questionnaire items were closed-ended, utilising a five-point Likert scale (e.g., from 'very unsatisfied' to 'very satisfied'; 'never' to 'very often'), binary response formats (e.g., yes/no), or multiple-choice questions. To provide participants with a neutral response option, all Likert scales were odd-numbered and included a midpoint (Sudina, 2021, p. 1155; Krug & Schlüter, 2013, pp. 77–78). In order to gain deeper insights, the survey also featured four open-ended questions to elicit more nuanced responses about participants' goals, perceived changes, and contextual influences. Overall, the survey was designed to support the study's mixed-methods approach (Dörnyei & Dewaele, 2022, pp. 12–13; Creswell & Plano Clark, 2018, p. 3) and to help identify self-reported variables that may have contributed to changes in participants' pronunciation. An offline version of the survey is available in the appendix of this thesis.

In essence, the combination of audio recordings and web-based survey responses provide insights into observable changes in pronunciation as well as participants' self-reported views on factors influencing these changes. The subsequent section details when and how these instruments were administered during the study.

6.2.5 Data collection

The data collection for this study was organised into three main phases, spanning a total of 19 months. The longitudinal study began with a pre-testing phase (T1) in March 2023, during which participants completed a controlled read-aloud task. This was followed by a 14-week EPI period in the PPOCS 1 course and its accompanying language lab. Subsequent to this instructional phase, an immediate post-test (T2) was carried out in June 2023, involving the same read-aloud task to allow for direct comparison. The final phase of data collection involved a delayed post-test (T3) in October 2024, during which participants submitted an audio recording and completed a web-based survey. For a clear visual summary, Figure 4 illustrates the three data collection points and the 14-week instructional phase, covering the full 19-month study period.

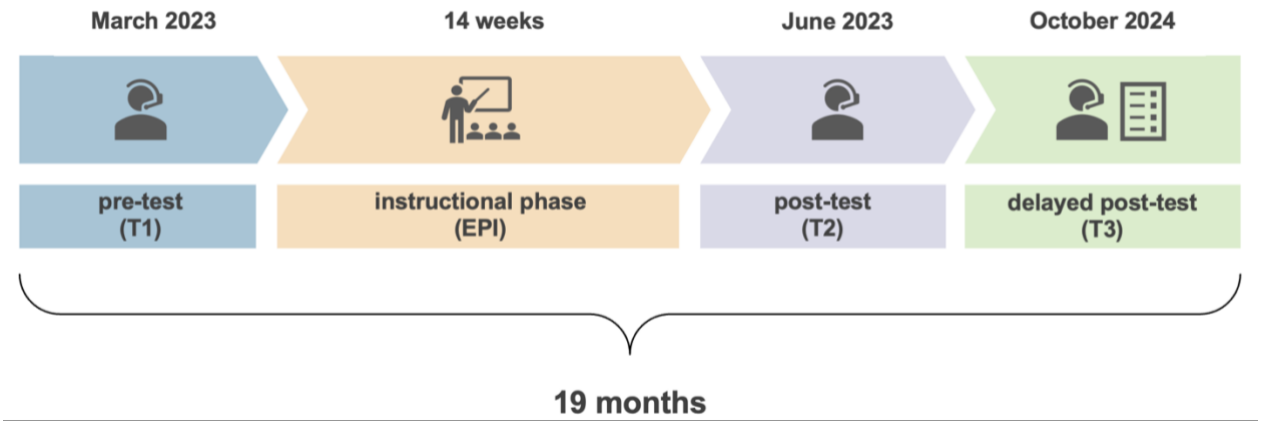


Figure 4. *Overview of data collection points and instructional period*

As shown in Figure 4, the first two testing points (T1 and T2) were aligned with the duration of PPOCS 1 and the language lab. The pre-test (T1) was conducted during the first week of the summer semester (March 2023), before students received any EPI. The post-test (T2) took place 14 weeks later in June 2023, directly following the end of the instructional EPI phase. During both testing sessions, participants recorded themselves reading a diagnostic passage aloud (see Section 5.2.4), using headsets provided in the language lab. All recordings were supervised by a tutor to ensure consistency and data quality. The datasets for T1 and T2 were originally collected for a separate research project on the short-term effects of EPI, which examined participants' pronunciation development over one semester (cf. White-Hautzinger et al., 2025). These recordings subsequently served

as the empirical foundation for the present longitudinal study. In total, 48 students completed both the T1 and T2 phases.

As a last step in the data collection process, delayed post-tests (T3) were conducted in October 2024, approximately 16 months after T2. This final phase was completed by 28 of the 48 participants from the first two testing phases. Unlike the earlier two testing phases, the T3 recordings were completed independently by the participants, who submitted their audio files via a Moodle platform (i.e., a free and online learning management system) created solely for the purpose of this study. Alongside their recordings, participants also completed a web-based survey administered on the same platform. Detailed instructions for submission, including the recording procedure, survey access, and ethical information such as study purpose, confidentiality, and participants' rights, were provided in detail on the Moodle site.

Together, these three data collection phases generated a robust longitudinal dataset consisting of a total of 84 audio recordings (i.e., three per participant) and 28 survey responses. This comprehensive set of data serves as the foundation for the subsequent analysis of participants' pronunciation development and self-reported influencing factors. Given the richness of the collected data, this thesis focuses only on those aspects most directly relevant to the proposed research questions, while additional dimensions of the data will be explored in future research.

6.2.6 Data analysis

To investigate changes in pronunciation over time, this section details the procedures used to analyse the collected data. Participants' audio recordings from three distinct time points (T1, T2, T3) were analysed using a self-developed analytic rating scale, designed specifically for the purposes of this study. This scale was applied to assess the production of two selected segmental features from the LFC, namely vowel length in long monophthongs and the word-initial aspiration of /t, k/. In addition, responses from the web-based survey administered at T3 were analysed to explore participants' self-reported factors related to pronunciation development and retention. To ensure comparability, a consistent data analysis procedure was applied, which is described in more detail below.

As a first step, an analytic rating scale was designed and then applied as the main instrument for evaluating pronunciation changes across the three testing phases. When developing such a rating scale, two key factors strongly influence the reliability of pronunciation assessment, namely the clarity of the scale itself and the rater's expertise (Hirschi & Kang, 2023, p. 214). Designing a robust scale is particularly challenging because descriptors must be interpreted consistently across items and rating levels (Harding, 2017, p. 12). Previous research has shown that vague or overlapping descriptors can compromise both validity and reliability (Zhong, 2019, p. 144; Harding, 2017, pp. 12, 29). To avoid these issues, the present study adopts concise, clearly labelled, and easily interpretable descriptors for each level of performance, following recommendations by Zhong (2019, p. 145) and Harding (2017, p. 30).

In light of these considerations, the rating scale is intended to enable a transparent and reliable auditory analysis of the speech samples collected across all three testing phases. To ensure consistency, the same scale was applied to all recordings from T1, T2, and T3. The rating scale was specifically designed to assess the intelligibility of two selected segmental features of the LFC, namely vowel length in long monophthongs and word-initial aspiration of the fortis plosives /t, k/. It adopts a five-point scale to capture varying degrees of intelligibility, ranging from level 1 (no intelligible production) to level 5 (consistently intelligible production). Each rating level corresponds to a defined range of intelligibly produced tokens (see Table 5), thereby supporting a systematic scoring process. It should be noted that the intervals of the rating scale are not strictly linear, which increases the sensitivity and discriminatory power of the scale, particularly at the higher end (5). Table 5 below presents the full rating scale employed in the assessment of all participant recordings.

Table 5. *Analytic rating scale for investigating selected segmental LFC features*

vowel length in long monophthongs (18 tokens)		
descriptor	level	intelligible tokens
no intelligible production	1	[0;2]
rarely intelligible production	2	[3;6]
occasionally intelligible production	3	[7;11]
mostly intelligible production	4	[12;15]
consistently intelligible production	5	[16;18]
word-initial aspiration of /t, k/ (17 tokens)		
descriptor	level	intelligible tokens
no intelligible production	1	[0;2]
rarely intelligible production	2	[3;5]
occasionally intelligible production	3	[6;10]
mostly intelligible production	4	[11;14]
consistently intelligible production	5	[15;17]

Employing the rating scale presented in Table 5, each participant recording was evaluated for all previously identified tokens in the reading passage: 18 tokens for long monophthongs and 17 tokens for word-initial fortis plosives (see Section 6.2.4). For each token, the rater judged whether the feature was intelligible or not, assigning a binary score of 1 (intelligible) or 0 (not intelligible). The total number of intelligibly produced tokens per feature (i.e., vowel length in long monophthongs and word-initial aspiration) was then summed for every participant. This total was then converted into a score on the five-point scale shown in Table 5. Each level represents a defined range of intelligible tokens: not intelligible production (1), rarely intelligible production (2), occasionally intelligible production (3), mostly intelligible production (4), and consistently intelligible production (5).

To illustrate, if one participant produced 14 of 18 long monophthongs intelligibly, the production was rated as ‘mostly intelligible production (4)’. If the same participant produced 15 of 17 fortis plosives intelligibly, the rating for that feature was ‘consistently intelligible production (5)’. A similar rating procedure has been employed in other research

(cf. White-Hautzinger et al., 2025, p. 122) and was repeated in the present study for all 84 recordings. This approach enabled a quantifiable and consistent assessment of pronunciation accuracy for both segmental features across all three testing points.

To examine participants' pronunciation development and retention over time, the resulting five-point scores for each participant and feature were subsequently analysed using descriptive statistics, conducted in the statistical software JASP. Mean scores and standard deviations were calculated for each testing phase (T1, T2, T3), with particular attention to changes between T1 and T2 (i.e., to assess the effects of EPI) and between T2 and T3 (i.e., to evaluate the extent of long-term retention). Given the ordinal nature of the data and the relatively small sample size (N = 28), the initially considered Wilcoxon signed-rank test was not applied, as the distribution of scores in some comparisons resulted in a zero-test statistic. Instead, changes in performance were analysed descriptively and illustrated using bar charts generated in Excel. These visualisations help illustrate participants' average performance across test phases and offer insight into the instructional impact and durability of their pronunciation gains (see Chapter 6.3).

All speech samples were evaluated by the same rater (i.e., the author of this thesis) who has approximately three years of experience as a pronunciation tutor at the University of Vienna and can, therefore, be regarded as a trained listener. In this role, the author has worked closely with advanced EFL learners, regularly providing targeted feedback on their pronunciation. As a result, the rater is highly familiar with the typical difficulties faced by this learner group. Given this background, a high level of rating accuracy and consistency can be expected (Hirschi & Kang, 2023, p. 214; Cucchiarini et al., 2000, pp. 110–112; Thompson, 1991, p. 184).

To further strengthen the reliability of the auditory analysis, a high degree of intra-rater reliability was established. First, all 84 speech samples were anonymised by assigning codes. Second, the recordings were randomised prior to rating, ensuring that the rater did not know whether a given sample belonged to T1, T2, or T3. Together, these steps helped to minimise the risk of potential rater bias. In addition, a re-rating procedure was conducted. This means that approximately 25% of the total speech samples (i.e., 21 out of 84 recordings) were selected at random and rated a second time by the same rater two

weeks later. This time gap was intended to minimise the influence of memory effects on the ratings (cf. Isaacs & Thomson, 2013, pp. 29-30). To assess the degree of agreement between the two rounds of ratings, Spearman's Correlation was calculated using JASP. This measure is commonly used to evaluate the internal consistency of repeated ratings by one rater (Field, 2018, pp. 141-142). Intra-rater reliability was examined separately for vowel length and aspiration ratings. The resulting Spearman's correlation coefficients were $\rho = .92$ ($p < .001$) for vowel length and $\rho = .84$ ($p < .001$) for word-initial aspiration, both indicating very strong consistency between the two rating rounds. This step was included to verify the stability of the rating scale and to ensure that the assessment of the two segmental LFC features was applied consistently over time.

The second source of data in this study is a web-based survey, which aimed to capture participants' self-reported views on pronunciation retention. As outlined in Section 6.2.4, the survey included both closed- and open-ended items, which required a mixed-methods approach to data analysis (Creswell & Plano Clark, 2018, p. 3). Responses to closed-ended questions were analysed quantitatively using Excel, with descriptive statistics such as mean values, frequencies, and percentages used to summarise participants' ratings. In contrast, the open-ended responses were analysed qualitatively by identifying recurring themes across participants' responses. Specifically, these open-ended questions allowed participants to elaborate on individual goals, pronunciation practices, and other factors not captured by the closed-ended items. As Creswell and Plano Clark (2018, p. 73) observe, qualitative data provide researchers with "emergent themes and interesting quotes that can be used to validate and embellish the quantitative survey findings", which was exactly the purpose of including open-ended questions in the questionnaire.

Taken together, the data analysis procedures outlined above allow for clear comparisons of participants' pronunciation performance across the three testing phases of this study. By analysing both speech samples through an analytic rating scale and self-reported data from the questionnaire, the study is well positioned to generate new insights into the long-term impact of EPI. The findings will be presented in the subsequent section.

6.3 Results and discussion

Building on the preceding chapter, which introduced the study's design, research questions, and methodological considerations, this chapter presents a comprehensive analysis and interpretation of the data collected at three data collection points (T1, T2, T3). The analysis draws on two main data sources, including ratings of participants' recorded audio samples at T1, T2, and T3 and responses to the web-based survey collected at T3. To ensure anonymity, each participant was assigned an anonymised code (R1-R28), which is used throughout this chapter when referring to individual data points.

To maintain a clear and coherent structure, this chapter is divided into three subsections, each addressing one of the three research questions. The first section looks at how participants' production of two selected segmental LFC features developed between T1 and T2, shedding light on the impact of EPI. The second section investigates whether these features were retained in the long term comparing T2 and T3 data. The final subsection considers learners' self-reported factors that may have influenced their pronunciation development and retention of the two LFC features under investigation.

Throughout this chapter, the findings of the study are presented, interpreted, and discussed in relation to the theoretical frameworks and empirical studies reviewed in earlier chapters. This integrated approach allows for a nuanced analysis of the data and meaningful connections between the results and existing research on EFL learners' pronunciation development and retention. In doing so, the chapter provides insights into both the immediate and long-term effects of EPI, and individual and contextual factors that may have influenced advanced Austrian EFL learners' production of two key LFC features.

6.3.1 Development of segmental LFC features (T1 – T2)

The analysis begins by addressing the first research question, which investigates how Austrian EFL learners' pronunciation of two selected segmental features from the LFC, namely vowel length in long monophthongs and the word-initial aspiration of fortis plosives, developed over the course of one semester of EPI. The aim was to explore whether the instruction provided in the EPI course PPOCS 1 led to observable gains in learners' intelligibility in accuracy in these two specific areas.

To examine this development, speech samples were collected from 28 participants at two testing points: the pre-test (T1) and the post-test (T2), which were conducted at the beginning and end of PPOCS 1, respectively. At each stage, participants were asked to read the text passage *The North Wind and the Sun* to ensure comparability across test points (see Section 6.2.4). As described in Section 6.2.6, each feature was rated on a five-point rating scale, with (1) indicating 'not intelligible production' and (5) representing 'consistently intelligible production'. In order to provide an initial overview of the overall development, total scores were calculated for each feature at T1 and T2. These scores offer a helpful first insight into patterns of change.

Figure 5 below illustrates the summed total scores for both targeted segmental LFC features at T1 and T2. This visual comparison highlights the initial effects of EPI and serves as a starting point for the subsequent interpretation and discussion of the results.

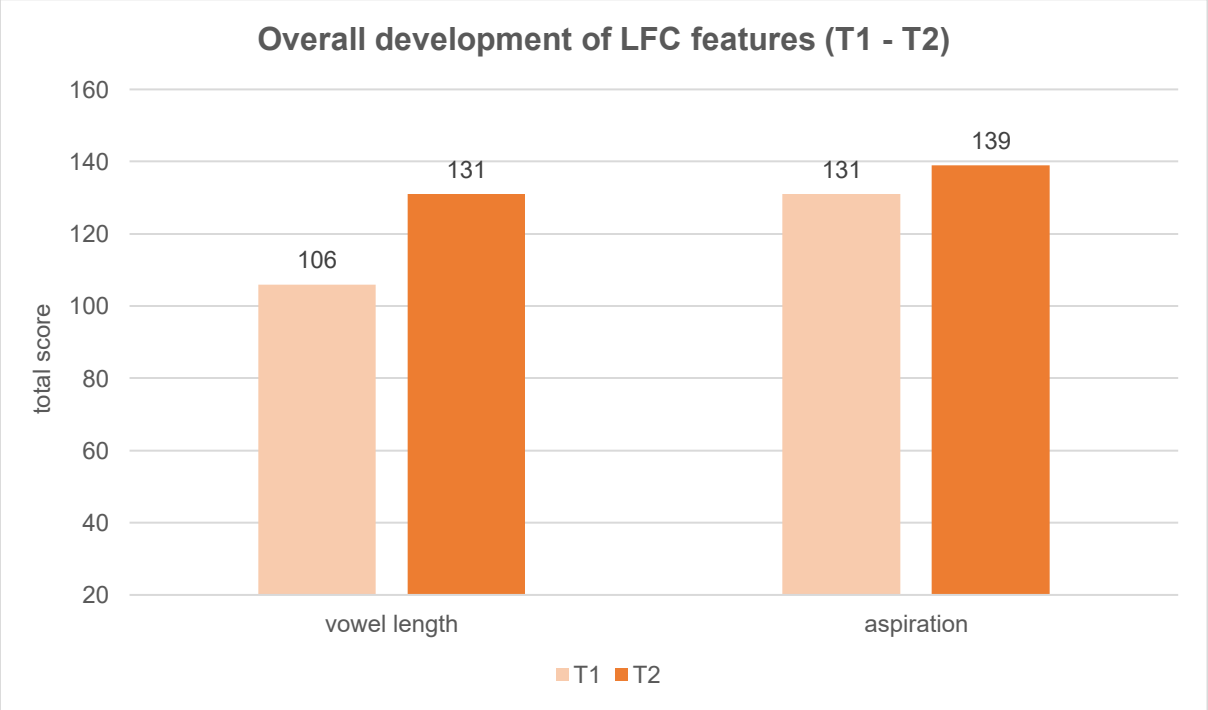


Figure 5. *Overall development of segmental LFC features between T1 and T2*

As shown in Figure 5, both segmental LFC features under investigation demonstrate an overall upward trend between T1 and T2, indicating measurable improvement following one semester of EPI. The total score for vowel length, assessed in the production of long

monophthongs, rose from 106 at T1 to 131 at T2, indicating notable gains in the production of vowel quantity. Similarly, the aspiration of word-initial fortis plosive sounds /t, k/ improved from an already high score of 131 to 139. Although both features show improvement after the completion of PPOCS 1, the increase in intelligible production is more pronounced for long monophthongs.

These overall findings in both long monophthongs and word-initial aspiration provide initial support for the effectiveness of EPI in enhancing segmental accuracy among advanced EFL learners. In fact, the observed improvement between T1 and T2 aligns with a substantial body of research suggesting that EPI can lead to measurable gains in learners' production of segmental features (e.g., Thomson & Derwing, 2015, p. 327; Saito, 2013, p. 47; Saito, 2011, p. 222; Derwing & Munro, 2005, p. 387). Taken together, these findings support the view that raising learners' awareness of specific phonological targets, such as vowel quantity and aspiration of fortis plosives, can lead to noticeable improvement in intelligibility-related pronunciation features.

In order to provide a more detailed picture of the gains observed in both targeted segmental features, Table 6 summarises the descriptive statistics for vowel length in long monophthongs and word-initial aspiration of fortis plosives at T1 and T2. These findings offer additional insight into patterns of development that may not be evident from the overall scores alone.

Table 6. *Descriptive statistics for vowel length and word-initial aspiration at T1 and T2*

	T1 (vowel length)	T2 (vowel length)	T1 (aspiration)	T2 (aspiration)
Valid	28	28	28	28
Mean	3.786	4.679	4.679	4.964
Std. Deviation	0.787	0.476	0.548	0.189
Minimum	2.000	4.000	3.000	4.000
Maximum	5.000	5.000	5.000	5.000
Increase in %		+23.6%		+6.1%

As illustrated in Table 6, both targeted LFC features improved between T1 and T2, with vowel length in long monophthongs showing the larger relative gain. To be more precise,

for this feature, the mean score increased from 3.79 (SD = 0.79) at T1 to 4.68 (SD = 0.48) at T2, corresponding to a 23.6% gain. The initial mean at T1 ($\bar{x}$ = 3.79, SD = 0.79) suggests that some learners experienced difficulties in producing the length of long monophthongs (e.g., /u:/ in <blew>; /ɑ:/ in <hard>) prior to receiving EPI. The increase of nearly one full scale point following one semester of EPI indicates that learners generally progressed from 'rarely' or 'occasionally' intelligible production towards 'consistently intelligible' production of long monophthongs. The reduced standard deviation at T2 (SD = 0.48) reflects greater consistency in learners' performance, suggesting that the gains were not restricted to a few high-achieving participants but were evident across the group. This consistency is further supported by the increase in the minimum rating score from 2 to 4, showing that even the lowest-scoring learners at T1 achieved mostly or consistently intelligible production of the investigated vowel sounds by T2. Notably, 27 out of 28 participants either improved or maintained their score across the two testing points. The single case of deterioration (R14) (i.e., from score 5 to 4) appears to be linked to the reading pace. In fact, the auditory analysis revealed that the participant read the passage at T2 at an unusually fast rate, which likely resulted in shortening some of the long monophthongs.

As for word-initial aspiration of fortis plosive sounds (/t, k/), the mean score also increased, from 4.68 (SD = 0.55) at T1 to 4.96 (SD = 0.19) at T2, representing a 6.1% gain (see Table 6). In contrast to vowel length, which showed a lower T1 mean, the relatively strong pre-test performance in word-initial aspiration reveals that the vast majority of the participants (*n* = 27) were already producing these sounds in a mostly or consistently intelligible manner prior to the EPI course. Nonetheless, the reduction in standard deviation from 0.55 to 0.19 points to greater consistency at T2, with fewer instances of reduced or inconsistent word-initial aspiration affecting intelligibility. This trend is further supported by the increase in the minimum score from 3 to 4, indicating that lower-scoring participants at T1 also achieved a high level of accuracy in aspirating fortis plosive sounds by T2. By the end of the EPI course, 23 participants maintained their score, four improved by one point, one improved by two points, and none deteriorated. All participants scored either 4 or 5 at T2, reflecting uniformly high intelligibility in the production of word-initial aspiration following EPI.

The improvements observed across the two selected segmental LFC features can be better understood when considered in relation to previous findings on the effectiveness of EPI. In fact, vowel length and word-initial aspiration have been identified as particularly challenging for German-speaking EFL learners and as potential sources of unintelligible language production if left unaddressed (Mairano et al., 2018, pp. 45–46; Berger, 2010, pp. 108–110; Swan & Smith, 2001, pp. 37–39). While both features showed measurable gains following one semester of EPI, the more substantial improvement of 23.6% was observed in producing long monophthongs. This finding is consistent with Wieden and Nemser's (1991, p. 54-58) large-scale study, which found vowel production to be me more problematic for Austrian learners than consonant production, an outcome attributed to L1 transfer effects and general difficulties with maintaining vowel quantity. Participants' initial difficulties with producing long monophthongs are also in line with Berger's (2010, p. 110) notion that German-speaking learners often require dedicated practice in distinguishing and producing different vowel quantities.

In contrast, the minimal improvement in word-initial aspiration by only 6.1% likely reflects a ceiling effect, as most participants already demonstrated high accuracy at T1. Although aspiration of /t, k/ in stressed initial position has been reported as a recurring difficulty for German-speaking EFL learners (Berger, 2010, pp. 107–108), the strong pre-test results in this study may indicate participants' prior exposure to, or familiarity with, English aspiration patterns. Nevertheless, the increased consistency at T2 (SD = 0.19) suggests that EPI can still develop and refine the production of aspirated fortis plosives, even when accuracy is initially high. Similar findings were obtained in previous studies which found that targeted segmental instruction consistently yielded significant improvements in consonants and consonant clusters, leading to greater accuracy at the segmental level (Naeini & Adni, 2017, pp. 116-117; Saito & Leyster, 2012, p. 23).

Overall, the present findings corroborate previous evidence that targeted EPI at the tertiary level can facilitate the development of advanced learners' production of two segmental LFC features (e.g., Halahan et al., 2023, p. 139). Given that both vowel length and aspiration are integral to the LFC and directly affect intelligibility (Jenkins, 2000, p. 159), the results provide further empirical support for prioritising these features in pronunciation instruction for advanced EFL learners.

6.3.2 Retention of LFC features (T2 – T3)

To answer the second research question, which examines the extent to which advanced EFL learners retain the two targeted segmental LFC features acquired during one semester of EPI, participants' post-test (T2) and delayed post-test (T3) performances were compared. As shown in Section 6.3.1, participants improved in both vowel length and word-initial aspiration following the one semester long EPI course. The present analysis now investigates to what extent these developed LFC features were maintained over time. For this, T3 data were collected approximately 16 months after the completion of the EPI course (see Section 6.2.5). To ensure comparability, the T3 ratings were based on the same diagnostic passage (i.e., *The North Wind and the Sun*) and focused on the identical segmental LFC features, namely vowel length in long monophthongs and word-initial aspiration of fortis plosives.

In order to provide a general overview of participants' retention of the two segmental LFC features, Figure 6 below visualises the total scores at T2 and T3. These totals represent the sum of individual scores for each feature across all participants, offering a first impression of any changes in learners' performance.

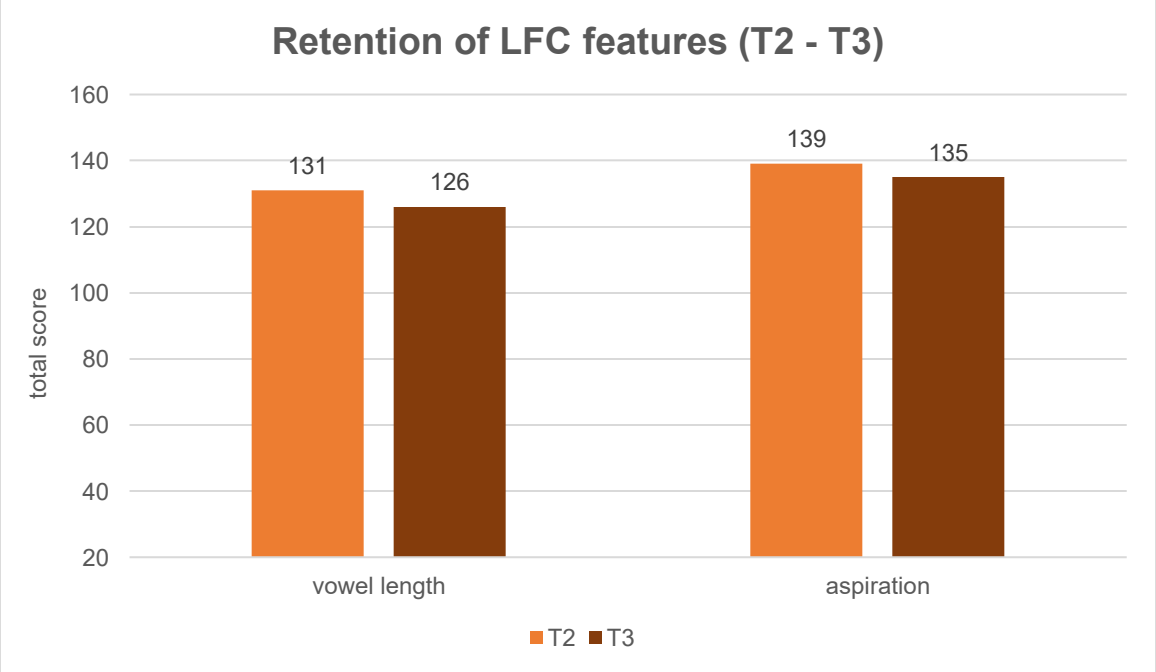


Figure 6. *Retention of LFC features between T2 and T3*

As can be seen in Figure 6, both targeted features display a slight decline in total scores between immediate the post-test (T2) and the delayed post-test (T3). For vowel length, total scores decreased from 131 to 126, representing a modest reduction in rating scores. A similar pattern can be observed for word-initial aspiration, with total scores dropping from 139 to 135. While this demonstrates a slight decrease in intelligible production of the two targeted LFC features, scores at T3 remain clearly above the pre-test level (see Figure 5). This suggests that pronunciation gains achieved through EPI were largely sustained over time, despite some reduction between T2 and T3.

These overall findings take on greater significance when viewed against the limited research on the long-term impact of EPI in L2 pronunciation. In fact, findings on the long-term effects of EPI remain relatively scarce, with most research focusing on short-term gains measured immediately after instruction (see Section 3.2.2) (Nagle, 2021, p. 154; Algethami, 2016, p. 297; Derwing & Munro, 2013, p. 180). The slight decline observed for both vowel length and word-initial aspiration between T2 and T3 is, therefore, of particular interest as it adds to the limited body of evidence on pronunciation retention. Some studies have shown that pronunciation gains can be maintained over extended periods, particularly for features that are relatively easily acquired (Sardegna & Jarosz, 2023, p. 23; Brekelmans, 2017, pp. 45-52). Others, however, report partial loss or inconsistency, most notably for features that require sustained practice to stabilise, such as vowel quality (Saito & Hanzawa, 2018, pp. 21-22; Brekelmans, 2017, p. 52; Algethami, 2016, pp. 300-302). In the present study, both investigated segmental features showed only a moderate decrease from T2 to T3. Given that T3 scores remained well above pre-test levels, much of the progress achieved through EPI appears to have been retained over the 16-month period, indicating its potential for lasting benefits.

A more detailed examination of the post-test (T2) and delayed post-test (T3) results is presented in Table 7, which summarises the descriptive statistics for vowel length and word-initial aspiration. This allows a closer analysis of the subtle changes in participants' performance measured at T3.

Table 7. *Descriptive statistics for vowel length and word-initial aspiration at T2 and T3*

	T2 (vowel length)	T3 (vowel length)	T2 (aspiration)	T3 (aspiration)
Valid	28	28	28	28
Mean	4.679	4.500	4.964	4.821
Std. Deviation	0.476	0.577	0.189	0.476
Minimum	4.000	3.000	4.000	3.000
Maximum	5.000	5.000	5.000	5.000
Decrease in %		- 3.8%		- 2.9%

Table 7 presents a more detailed picture of participants' development of acquired LFC features over time. Evidently, both targeted features exhibited a slight decline between T2 and T3, with vowel length showing the larger relative decrease. Specifically, for vowel length, the mean score dropped from 4.68 (SD = 0.48) at T2 to 4.50 (SD = 0.58) at T3, representing a 3.8% decrease. The high mean at T2 reflects the substantial improvement achieved through EPI, yet the lower T3 mean suggests some decline in performance after the completion of PPOCS 1. The increased standard deviation from 0.48 to 0.58 points at greater variability in learners' performance at T3, implying that the decline was not uniform across participants. Hence, some learners showed a greater decrease than others and some maintained their previous level of performance. Interestingly, two participants showed improvement between T2 and T3. One of them was the single case of deterioration (R14) between T1 and T2, which had previously been attributed to an unusually fast reading pace. At T3, however, this participant reached the highest score on the scale, supporting the assumption that the lower T2 rating was indeed due to the increased pace rather than a genuine difficulty in producing long monophthongs intelligibly. The other notable case was R5, who demonstrated a steady progress across all three testing phases, despite the absence of further EPI after T2. This suggests that individual factors, such as continued personal practice or increased exposure to English outside the course, may have contributed to sustained development beyond the instructional period. Such factors will be examined more closely in Section 6.3.3.

For word-initial aspiration of /t, k/, the mean decreased from 4.96 (SD = 0.19) at T2 to 4.82 (SD = 0.48) at T3, corresponding to a 2.9% reduction. Given the already high mean at T2, this smaller decrease suggests a degree of stability in the production of aspiration.

However, the standard deviation increased from 0.19 to 0.48, indicating reduced consistency across the group. The minimum score also fell from 4 to 3, pointing to isolated cases of lower performance at T3. One such case was R5, who progressed from a score of 4 at T1 to 5 at T2, but declined to 3 in the delayed post-test (T3). This pattern contrasts with R5's vowel length development, which improved even beyond the instructional period, and may suggest that aspiration was a less stable feature for this learner. Despite such individual declines, the maximum score remained at 5, and the majority of learners continued to produce aspirated fortis plosives in a mostly or consistently intelligible manner.

The reported findings contribute to the growing body of literature on the long-term effects of EPI by focusing on segmental LFC features among tertiary-level EFL learners. While earlier research in this field has often emphasised suprasegmental features such as linking, word stress, and prosody (e.g., Sardegna & Jarosz, 2023, p. 23; Saito & Hanzawa, 2018, pp. 21-22) or examined only individual phonemes (e.g., Algethami, 2016, pp. 300-302), the present study investigated two features that have received less longitudinal attention. The results demonstrate a slight decline between T2 and T3. However, mean scores for both features remained above pre-test levels. This pattern suggests that the gains achieved through EPI were largely maintained over the one-and-a-half-year period, supporting the view that targeted instruction can yield benefits that persist beyond the instructional phase (cf. Brekelmans, 2017, pp. 45–54).

The findings also highlight a feature-specific dimension of long-term retention in pronunciation. Earlier studies have reported that suprasegmental gains tend to be more robust over time, whereas segmental improvements can be less consistent (e.g., Sardegna & Jarosz, 2023; Saito & Hanzawa, 2018). This trend is echoed in Algethami's (2016, pp. 300-302) study, which tracked Arab EFL learners' production of the /p/ sound and found limited long-term success despite short-term improvement. Similarly, while the majority of participants in the present study retained much of their progress, individual variation was evident. To exemplify, R5 improved in vowel length even beyond T2, yet their word-initial aspiration scores, after reaching consistent intelligibility at T2, dropped notably at T3. This supports Brekelmans' (2017, p. 47) observation that pronunciation features do not necessarily develop with equal stability. Features that are more difficult to

acquire may require ongoing post-instruction practice, a point also emphasised by Derwing (2010, p. 27) and Thompson and Derwing (2015, p. 339).

Taken together, the findings provide further support for the potential of EPI to promote durable gains in advanced EFL learners' production of certain segmental LFC features, while also underscoring the need to account for both learner-specific and feature-specific factors when evaluating long-term outcomes. To gain a deeper understanding of potential reasons for the observed changes between T2 and T3, the following section examines participants' self-reported factors that may have influenced their changes in pronunciation over time.

6.3.3 Learner-reported factors influencing pronunciation changes

Beyond examining participants' development and long-term retention of the two segmental LFC features under investigation, this study also addresses a third research question, which seeks to identify the factors that may have influenced learners' ability to largely maintain these features over time. To this end, a web-based survey was administered at T3 to capture learner-reported factors shaping their pronunciation, such as their goals regarding their pronunciation development, practice habits, exposure to and use of English beyond the EPI course (PPOCS 1), as well as other influences that may have influenced their pronunciation since completing PPOCS 1. Importantly, the data reflects participants' self-reported perceptions rather than objectively measured behaviours, providing insight into learners' own perceptions of the factors that shaped their pronunciation over time. Since many of these factors are closely interconnected and often overlap, they are not presented as separate categories but discussed in relation to one another. Finally, this section does not report on all data gathered in the survey but focuses only on the survey items most relevant to addressing the third research question.

In order to gain a deeper understanding of what motivated participants at the beginning of the EPI course, one of the survey items asked participants to retrospectively state their main goals for PPOCS 1 which revealed a range of motivational orientations. Respondents could select multiple options, which are summarised in Figure 7 below.

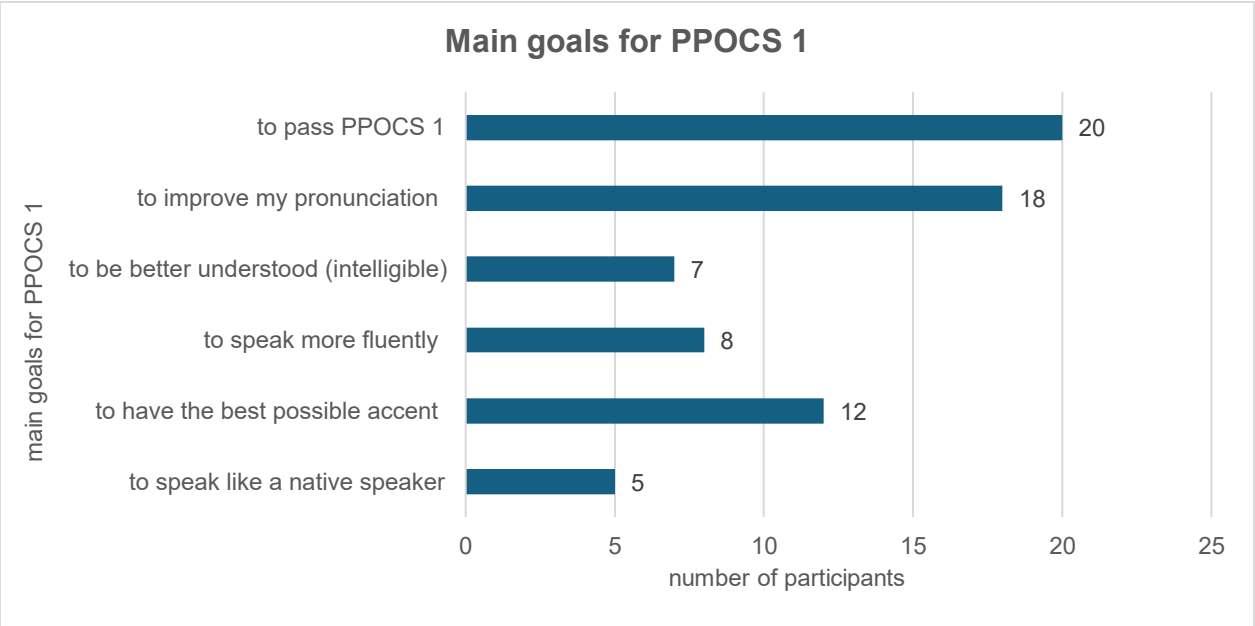


Figure 7. *Participants' main goals for the EPI course PPOCS 1*

As shown in Figure 7, the majority of participants (*n* = 20, 71%) reported extrinsically motivated goals as their primary aim, namely passing PPOCS 1. At the same time, intrinsically motivated goals were also strongly represented, with 64% (*n* = 18) seeking to improve their pronunciation and 43% (*n* = 12) aiming to achieve the 'best possible accent'. Fewer participants reported instrumental, intelligibility-oriented goals, such as being better understood (*n* = 7, 25%), while an even smaller group (*n* = 5, 18%) pursued native-like attainment as their personal goal for the EPI course. Hence, the data reveal that participants' goals encompassed predominantly externally driven but also personal aspirations to develop their pronunciation skills.

In fact, these findings align with the common distinctions made in previous research between intrinsic and extrinsic motivation in pronunciation learning (Saito et al., 2017, p. 618). The prominence of 'passing PPOCS 1' as a main goal reflects the instrumental, extrinsic motivation often observed among tertiary level students (Saito et al., 2017, pp. 615–618). In addition, the considerable number of participants who reported intrinsic aspirations, such as improving their pronunciation, suggests that their motivation extended beyond course requirements and reflected personal investment in developing their pronunciation during PPOCS 1. This combination of externally and internally driven goals suggests that students entered the course with both practical goals and genuine

interest in pronunciation development, which may help explain differences in how much effort they later devoted to maintaining their acquired skills.

Further motivational patterns emerged when participants were asked whether their initial pronunciation goals had remained the same after completing the EPI course. The questionnaire data show that most participants (*n* = 23, 82%) report their goals as unchanged, while a smaller group (*n* = 5, 18%) indicate shifts over the one-and-a-half-year period. Open responses reveal that some participants moved away from native-like targets toward intelligibility. To exemplify, R27 noted that although their initial aim was to "speak like a native speaker," their current priority is "better and more clearer [sic] communication." Similarly, R22 explained that they are "now focused on speaking intelligibly […]". Beyond the shift toward intelligibility, R11 also mentioned instrumental reasons for their change of goals stating that "as a future teacher, [they] would like to be as comprehensible as possible for [their] students."

The fact that the majority of participants reported no change in their pronunciation goals suggests that their initial motivational orientations were relatively robust. However, the changes observed in the remaining 18% are particularly noteworthy. Some participants shifted from aspiring native-like pronunciation toward prioritising intelligibility, which points to a change not only in motivation but also in attitude. As Tsunemoto and McDonough (2020, p. 26) argue, attitudes reflect the value learners place on certain aspects of pronunciation, and in this case, intelligibility emerged as a more meaningful goal than native-likeness. This finding also resonates with Afshari and Ketabi (2016, p. 76), who similarly observed that highly motivated learners may prioritise communicative success over accent approximation. The explicit mention of future teaching responsibilities of R11 further highlights how professional and instrumental reasons can shape both motivation and attitudes (Sardegna & Jarosz, 2023, p. 22). In this sense, the reported shifts suggest that a small group of learners gradually aligned their goals with communicative and professional needs rather than the idea of sounding like a native speaker.

Closely linked to participants' motivation is their self-reported engagement with pronunciation practice beyond the EPI course. To explore this aspect, the survey asked participants to indicate the extent to which they had actively worked on their pronunciation

since completing PPOCS 1. Responses were collected on a five-point Likert scale ranging from ‘not at all’ (1) to ‘a lot’ (5), which are depicted in Figure 8 below.

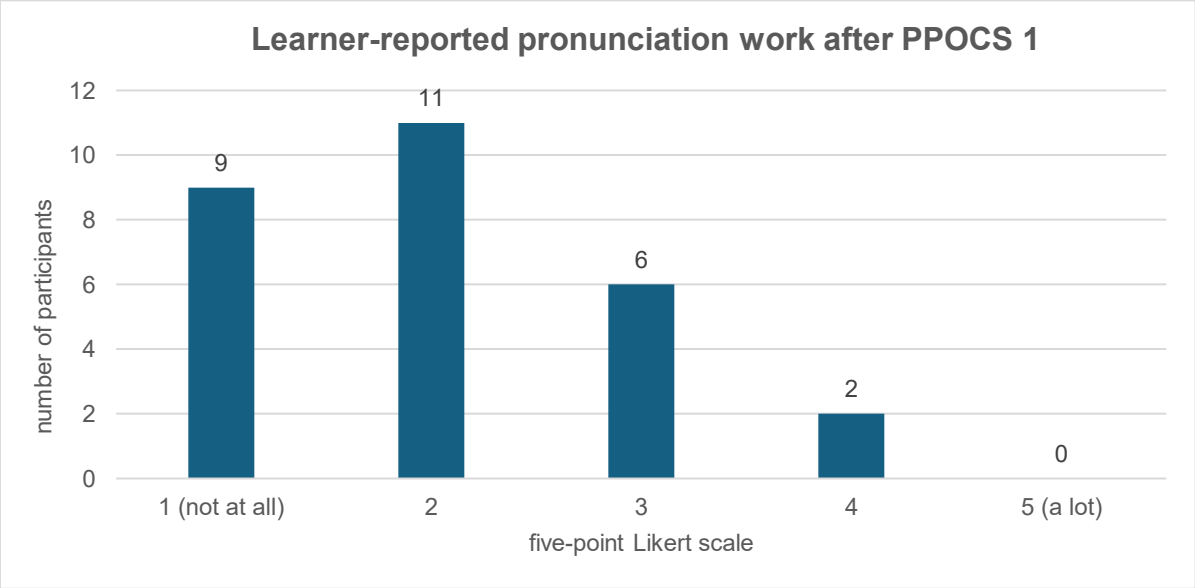


Figure 8. *Learner-reported pronunciation work after PPOCS 1*

A closer look at Figure 8 reveals that most participants reported only minimal engagement with pronunciation practice after the EPI course. The majority selected ratings at the lower end of the scale, with 9 participants (32%) indicating that they have not worked on their pronunciation at all and 11 participants (39%) a little bit. A smaller group reported moderate practice (n = 6, 21%), while only two participants (7%) indicated more consistent practice. Within this overall pattern, it is worth highlighting the two participants R11 and R25 who reported more sustained engagement with pronunciation practice beyond the EPI course. Their cases will be revisited in later sections to shed light on how these learners engaged in such practices. In any case, the overall mean score of 2.04 suggests that active work on pronunciation was, on average, rather minimal across the group.

In addition to reporting the overall extent of pronunciation work after PPOCS 1, participants were also asked to specify the ways in which they had practiced pronunciation since completing the EPI course. More precisely, the questionnaire invited them to indicate how frequently they engaged in a range of practices, using a five-point Likert scale from ‘not at all’ (1) to ‘a lot’ (5). Their responses were summarised by calculating mean scores for each practice, as illustrated in Figure 9.

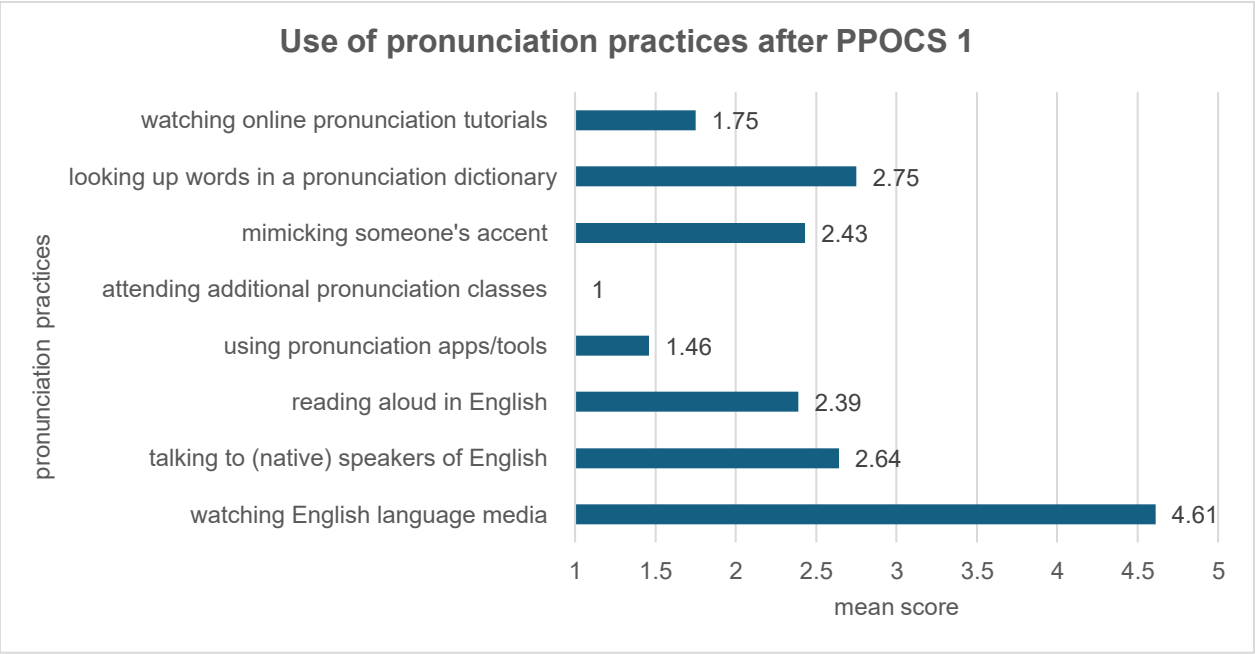


Figure 9. *Learner-reported use of pronunciation practices after PPOCS 1*

As depicted in Figure 9, the most frequently reported pronunciation practice was watching English-language media, including films and YouTube videos, which reached a mean score of 4.61. This indicates that nearly all participants engaged with this form of input regularly, suggesting that digital media serves as the most accessible and widely adopted means of exposure to spoken English. The second most commonly reported practice was looking up words in a pronunciation dictionary ($\bar{x}$ = 2.75). Interestingly, one participant (R24) noted in the open responses that this strategy was not only used for checking English pronunciation but also for exploring other accents they found interesting, such as Swedish, Russian, or Australian. Talking to (native) speakers of English was also relatively common ($\bar{x}$ = 2.64), followed closely by mimicking someone's accent ($\bar{x}$ = 2.43) and reading aloud in English ($\bar{x}$ = 2.39). By contrast, more formal and deliberate pronunciation-focused practices were used less frequently. Watching online pronunciation tutorials ($\bar{x}$ = 1.75) as well as using pronunciation apps or tools ($\bar{x}$ = 1.46) both received low mean scores, while attending additional pronunciation classes was reported as virtually absent ($\bar{x}$ = 1.00). Among the two learners who reported higher overall engagement (see Figure 8), R11 most frequently engaged in watching English-language media, looking up words in a pronunciation dictionary, reading aloud in English, and mimicking accents. Similarly, R25 reported using media input most often together with interactive and productive

practices, such as talking to other speakers of English, reading aloud, and mimicking accents. Overall, the data suggest that participants tended to rely more on informal and accessible opportunities for exposure and practice, whereas more explicit forms of pronunciation work were rarely pursued beyond the EPI course, a pattern also observed in previous research (cf. Sardegna & Jarosz, 2023, p. 23).

The survey responses reveal that participants engaged only minimally in deliberate pronunciation work after completing PPOCS 1, with most reporting little or no active practice. This suggests that despite short-term gains achieved through EPI (see Section 6.3.1), the long-term retention of the LFC features investigated was barely supported by consistent post-instruction pronunciation practice. This tendency aligns with Moyer's (2014, p. 431) observation that not only the type but also the intensity and sustainability of motivation determine whether learners continue investing in their pronunciation. In the present study, the low engagement ($\bar{x}$ = 2.04) points to limited motivational intensity beyond PPOCS 1. When examining the types of practices pursued, however, participants cited a strong reliance on informal and easily accessible exposure, particularly through English-language media ($\bar{x}$ = 4.61), a trend similarly observed in Vong et al.'s (2024, p. 992) recent study on Cambodian learners. Finally, the relatively frequent mention of interactive practices, such as talking to native speakers ($\bar{x}$ = 2.64), points to the social dimension of pronunciation learning. As Afshari and Ketabi (2016, p. 76) argue, socially engaged learners are more likely to take advantage of opportunities for practice. Overall, these findings indicate that participants' limited reliance on deliberate practice may also have influenced how they came to perceive changes in their pronunciation since completing PPOCS 1. Yet, individual cases such as R11 and R25, who reported engaging in more sustained pronunciation practice beyond the EPI course, suggest that continued practice (e.g., media exposure, reading aloud) may have helped them maintain their acquired segmental LFC features over time. In fact, Sardegna and Jarosz (2023, p. 23) report that highly motivated learners who maintained pronunciation practice after EPI retained considerable improvements in suprasegmental features. Notably, R11 retained both vowel length and word-initial aspiration from T2 to T3. R25, by contrast, showed a stable performance in word-initial aspiration but a slight decline in the production of long monophthongs (from a score of 5 at T2 to 4 at T3). These individual cases suggest that

while continued engagement with pronunciation may support long-term retention of LFC features, its effects are not necessarily uniform across all segmental categories. Thus, Sardegna and Jarosz's (2023, p. 23) observation regarding suprasegmentals appears to extend, at least partly, to segmental LFC features, although individual variation remains evident.

Besides reporting on the specific practices they engaged in, participants were also invited to evaluate how strongly they believe various factors and practices influenced their pronunciation development since PPOCS 1. To indicate their responses, they rated each item on a five-point Likert scale ranging from 'not at all' (1) to 'very much' (5). Figure 10 below displays the mean ratings for each item.

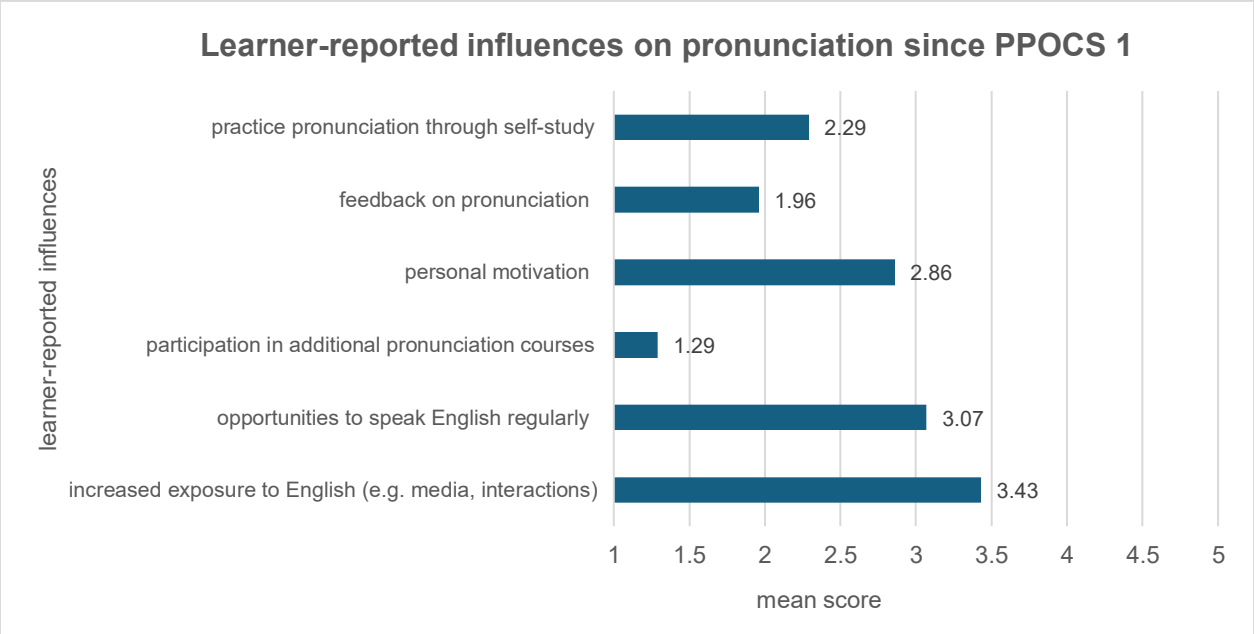


Figure 10. *Learner-reported influences on pronunciation since PPOCS 1*

As illustrated in Figure 10, participants attributed the strongest influence on their pronunciation development since completing PPOCS 1 to increased exposure to English through media or social interactions with other speakers of English, which reached a mean score of 3.43. The distribution of responses shows that nearly half of the participants (*n* = 12) felt this factor influenced their pronunciation 'quite a bit,' while an additional four reported it had influenced them 'very much.' Further, opportunities to speak English regularly emerged as the second most influential factor, with an average rating of 3.07. Interestingly, personal motivation was rated only moderately ($\bar{x}$ = 2.86), indicating that

learners did not perceive their own motivational drive as a particularly decisive factor in shaping their pronunciation over time. Finally, receiving feedback on pronunciation ($\bar{x}$ = 1.96) and participating in additional pronunciation courses ($\bar{x}$ = 1.29) were rated as the least influential factors. The overall pattern evident in Figure 10 indicates that learners attributed their long-term pronunciation development primarily to everyday exposure and use of English rather than to explicit practice or formal instruction beyond PPOCS 1.

These findings resonate closely with the previously reported trends in Figure 9, where participants indicated frequent reliance on media input as well as spoken interaction with other speakers of English, while more intentional and explicit forms of practice were rather rare. What learners perceived as most influential (i.e., exposure to English and opportunities for use), therefore, reflects the types of activities they engaged in most frequently. At the same time, the relatively low rating of personal motivation ($\bar{x}$ = 2.86) helps explain why sustained, active work on pronunciation after receiving EPI remained minimal overall (see Figure 8). Taken together, these findings suggest that while participants viewed general exposure to and use of English as the main drivers of their pronunciation development, they did not perceive motivation or deliberate practice as equally decisive factors. This limited emphasis on deliberate practice may also help explain why vowel length in long monophthongs proved less stable between T2 and T3. Compared to word-initial aspiration, this feature was less strongly established at T1 and T2, which suggests that it requires more sustained practice to be retained over time. In line with Brekelmans' observation (2017, p. 47), this indicates that pronunciation features do not necessarily develop or stabilise with equal consistency over time.

In addition to the closed-ended items, the survey also invited participants to share any further reflections on their pronunciation development since completing PPOCS 1. These open responses provide more nuanced insights into individual experiences among the participant group. A thematic analysis of the responses revealed three recurring themes: perceived decline in pronunciation skills over time, the resulting effects on learners' confidence and self-perception as speakers, as well as identity-related and attitudinal influences on pronunciation use. Each of these themes is discussed in the following paragraphs, illustrated with selected participant excerpts.

The first theme that emerged concerns participants' realisation that their pronunciation skills had declined since completing PPOCS 1. Specifically, respondents described a sense of regression, noting that the progress achieved during PPOCS 1 had not been fully maintained over time. Their open responses are demonstrated in the examples below:

> "I feel like my pronunciation was quite okay immediately after PPOCS 1 but deteriorated since then due to the lack of practice." (R24)
>
> "Overall, I feel like my pronunciation has gone back to the way it was before PPOCS […], especially in spontaneous speech." (R27)

Evidently, both participants report a perceived decline in pronunciation skills. Notably, R24 links their deterioration to a lack of practice which reinforces the survey findings that revealed generally low levels of active work on pronunciation beyond the EPI course (see Figure 8). R27, on the other hand, emphasises a reversion to their pre-instruction pronunciation skills, particularly in spontaneous speech, suggesting difficulties in transferring gains from PPOCS 1 to less controlled communicative contexts. This is especially relevant given that PPOCS 1 offered ample opportunities to engage in both reading aloud and free speaking tasks designed to practise and consolidate the features introduced in the course (Richter, 2021, p. 94). Such consistent practice for the duration of 14 weeks (see Section 6.2.2) resulted in measurable gains in the selected segmental LFC features. However, what appears to be missing is the consistent and ongoing practice beyond the course that Sardegna and Jarosz (2023, p. 23) emphasise as crucial for sustaining gains. For learners like R24 and R27, continued practice would likely have been necessary, particularly for features that are more difficult to acquire. This view is supported by previous research showing that certain pronunciation features (e.g. vowel quality) require sustained post-instruction practice to stabilise over time (cf. Saito & Hanzawa, 2018, pp. 21–22; Brekelmans, 2017, p. 52; Algethami, 2016, pp. 300–302; Thomson & Derwing, 2015, p. 339; Derwing, 2010, p. 27).

Another consequence of perceived decline in pronunciation skills was its effect on learners' confidence and their willingness to speak. While some participants focused on the partial loss of pronunciation features over time, others also described affective

aspects, such as anxiety, self-doubt, and fear of being judged. These responses suggest that the deterioration of skills was also accompanied by a negative shift in how participants evaluated themselves as speakers of English. The following two examples highlight this theme:

> "I'm afraid to speak to native speakers because I'm afraid of making mistakes. I know it's counterproductive and I really want to change it, but I feel like everyone is judging me - especially people from the English department." (R17)

> "I feel like my accent has gotten worse, which I think is really unfortunate. I used to be very comfortable speaking British English but now I sometimes feel bad because I know my accent used to be so much better. I have worked so hard in PPOCS 1 to improve my accent and now, a bit over one year later, I just do not have the same skills anymore." (R7)

These examples show how participants' perception of decline in their pronunciation skills is intertwined with affective consequences, such as reduced confidence and increased speaking anxiety. In fact, R17 explicitly connected their difficulties with a fear of speaking to native speakers, describing anxiety about making mistakes and feeling judged by peers. This reflects how their perceived pronunciation skills can become an affective barrier, leading to withdrawal from communicative opportunities despite a stated desire to improve. R7 reflected on a perceived decline in their pronunciation skills and emphasised the loss of comfort they once felt when speaking English. Their response highlights how a perceived deterioration in pronunciation can undermine learners' confidence, leading to feelings of dissatisfaction and self-consciousness about their current abilities.

Such accounts resonate with previous research emphasising that pronunciation development is closely linked to confidence and willingness to engage in spoken interaction. As Sharma (2021, p. 65) observes, targeted EPI can boost speaking confidence, yet without sustained support, learners may experience difficulties that reduce their readiness to communicate. This also aligns with Baran-Łucarz's (2014, p. 445) argument that persistent pronunciation difficulties may contribute to increased anxiety and reluctance to speak. R7's frustration at no longer being able to maintain previously achieved skills exemplifies this tendency, as the lack of practice translated not only into a perceived decline in pronunciation skills but also into negative self-evaluation. Similarly, R17's fear of being judged highlights how social pressures can intensify pronunciation-

related anxiety. Interestingly, however, the T2–T3 data show that both participants in fact maintained stable performances in the two investigated segmental LFC features, both receiving the highest rating scores. Their perceptions, therefore, appear to reflect a broader sense of decline in their overall pronunciation ability, which may extend beyond the two segmental features analysed in this study. In light of these findings, Clément et al. (2003, p. 191) and Baran-Łucarz (2014, p. 467) emphasise that consistent pronunciation training can help mitigate such affective barriers by gradually strengthening learners' confidence.

Beyond issues of confidence and self-evaluation, participants also highlight how their pronunciation is shaped by identity-related and attitudinal factors. Their responses reveal that learners' relationship with English, the contexts in which they use it, and the way they wish to be perceived can influence how much importance they attach to pronunciation accuracy. Two illustrative examples are provided below:

> "[...] I used to work in service and payed attention to my pronunciation more with English speaking customers, but now I mostly speak "denglisch[7]" in private with the people in my life so I don't care about perfect pronunciation at all." (R18)

> "I have been on an Erasmus exchange and most of my friends had an accent leaning more to the American one. I felt a little akward at times, as they kept pointing it out, so I gradually changed to a more "neutral" / american accent as well." (R4)

These responses highlight how identity and social context can shape learners' attitude toward pronunciation. For instance, R18 noted that while they used to pay close attention to their pronunciation when talking to English-speaking customers, they now mainly use "Denglisch" in private contexts and therefore "don't care about perfect pronunciation at all." In contrast, R4 highlighted how their Erasmus experience led them to adjust their pronunciation to a more "neutral" or American accent, influenced by the predominance of American accents among peers.

As can be seen in these examples, for the two respondents, pronunciation is not merely a matter of linguistic ability but also of how they position themselves socially and culturally. This is largely supported by Moyer (2013, pp. 65) as well as Setter and Jenkins (2005, p.

[7] The term "Denglisch" is a portmanteau of "Deutsch" (German) and English. The participant appears to use it to describe speaking English with a strong influence of German in terms of accent and pronunciation.

5) who note that identity is tied to how learners wish to be perceived through their language use. Indeed, R18 stated that they are less concerned about “perfect” pronunciation which illustrates how context and personal preferences can lower the value attached to accuracy. This finding also resonates with Jenkins’ (2000, p. 16) observation that some learners deliberately wish to retain L1 features as part of their identity. R4’s experience, on the other hand, emphasises how social dynamics and peer influence can encourage alignment with a particular variety. This corroborates Nunan’s (1991, p. 107) early notions asserting that pronunciation choices are often shaped by the desire to fit in a certain target community. What is more, the shift toward a more “neutral” or American accent likely reflects the kind of integrative motivation discussed by Saito et al. (2017, pp. 615-618), where learners adapt their way of speaking to build relationships and feel part of the community. Together, these examples demonstrate that learners’ pronunciation development is closely intertwined with issues of identity, attitude, and the cultural or social contexts in which they use English.

Overall, the survey findings largely corroborate previous research on the influence of individual and contextual factors in L2 pronunciation development. The analysis showed that participants initially entered PPOCS 1 with both extrinsic and intrinsic goals, though only a few reported changes over time, shifting their focus from nativeness toward intelligibility. After completing the course, however, most engaged only minimally in deliberate practice, instead relying on easily accessible opportunities for practice (e.g., media input, spoken interactions). This reliance suggests that long-term engagement was shaped more by convenience than sustained effort. Open responses further revealed that some learners linked their perceived decline in pronunciation skills to lack of practice, while others, for example, highlighted issues of confidence. Taken together, these self-reported perspectives suggest that the somewhat lower long-term retention of vowel length, compared to aspiration, may partly be explained by insufficient ongoing practice, combined with other factors such as learners’ confidence, motivation, and attitudes toward pronunciation. It should be noted, however, that these findings concern only two selected segmental LFC features and reflect learners’ subjective perceptions, which may not always correspond to their objectively measured pronunciation performance.

7 Conclusion

This Master's thesis investigated the long-term effects of EPI on advanced Austrian EFL learners, focusing on the production of two segmental LFC features: vowel length in long monophthongs and word-initial aspiration of fortis plosives. Employing a longitudinal mixed-methods design, data were collected through a controlled read-aloud task at three time points (T1, T2, T3) to examine participants' development and retention of these features. Additionally, a survey administered at T3 elicited participants' reflections on factors they believed had influenced their pronunciation over the one-and-a-half-year period. This research design provided insights into both the measurable development and retention of the targeted LFC features as well as learners' subjective evaluations of their pronunciation trajectories.

Findings from the first testing phase (T1–T2) reveal clear gains for both targeted features after one semester of EPI. Vowel length, which had posed greater challenges initially, improved more substantially, with mean scores increasing by nearly one full scale point and variability among learners decreasing. This suggests that EPI was particularly effective in addressing vowel quantity, a feature known to affect intelligibility in EFL learners. Improvement was also evident for word-initial aspiration, though to a smaller degree. Since most participants already produced aspirated fortis plosives accurately before EPI, the gains were more modest and mainly reflected greater consistency. Overall, the results confirm that one semester of EPI can significantly enhance advanced learners' production of two segmental LFC features.

Between T2 and T3, much of the improvement was retained, though not without some decline. Both vowel length and aspiration scores decreased marginally, with vowel length showing the greater reduction. Nevertheless, average T3 scores remained well above pre-test levels (T1), indicating that gains from EPI persisted over time. The relative stability of aspiration may reflect its stronger establishment at T1 and minor need for reinforcement. Overall, these findings align with previous research showing that pronunciation retention varies across features and learners which highlights the importance of both factors when evaluating the impact of EPI. The study, thus, shows that targeted EPI can yield lasting segmental gains, contributing valuable evidence on long-term pronunciation retention.

Additionally, the survey data offered insights into participants' views on what influenced their pronunciation after completing the EPI course PPOCS 1. Most participants reported engaging only minimally in deliberate pronunciation practice following the EPI course, instead relying heavily on more accessible forms of exposure, such as media input and occasional spoken interaction in English. These everyday forms of engagement were perceived as most influential, whereas explicit feedback or further pronunciation courses were rated less decisive. Some learners linked a decline in performance to a lack of practice, while others cited affective or identity-related factors that influenced how much attention they paid to intelligible pronunciation. Consistent with previous research, these findings suggest that maintaining accuracy in segmental production requires ongoing attention and practice outside the instructional context.

Beyond the immediate findings, this study holds important implications for pronunciation pedagogy and future research. The rich longitudinal dataset collected provides a solid foundation for further analyses extending beyond the scope of this thesis. A key insight emerging from this study is that pronunciation learning is an ongoing, dynamic process shaped by continued practice, exposure, motivation, and identity, and that it needs to continue even after EPI. The results confirm the effectiveness of EPI for two segmental LFC features, yet the differing retention patterns observed suggest that some features, such as vowel length, may require more sustained reinforcement to ensure long-term consistency. These findings call for a reconsideration of how pronunciation is embedded within tertiary curricula, advocating for a more continuous integration of intelligibility-oriented pronunciation training rather than a single EPI course. Further, learners should also be supported in developing autonomous strategies for maintaining their pronunciation beyond the instructional setting (e.g., using pronunciation dictionaries, apps). Ultimately, pronunciation development emerges as an ongoing process that relies on both explicit teaching and learners' independent pronunciation practice over time.

In essence, this thesis provides new longitudinal evidence on the long-term effects of EPI among advanced Austrian EFL learners. Focusing on two segmental LFC features, the study offers context-specific insights into pronunciation development and highlights the need to consider linguistic and individual factors for long-term retention.

8 Limitations and directions for future research

The present study makes a valuable contribution to research on EPI through its longitudinal design and its focus on an underrepresented learner group, namely advanced Austrian German-speaking EFL students at the tertiary level. Nevertheless, several limitations need to be acknowledged, as they constrain the generalisability of the findings, while also pointing to directions for future research.

To begin with, one limitation that is worth addressing is the narrow setting of the study. Although the sample size (N = 28) is relatively modest and typical in longitudinal research due to participant attrition (Cohen et al., 2007, p. 214), it nevertheless limits the statistical generalisability of the findings. In addition, while the specific focus on advanced Austrian EFL learners from the University of Vienna provides valuable insights into a learner group that has received little attention in previous research, it also restricts the extent to which the results can be transferred to other learner populations or educational contexts. Future studies could, therefore, broaden the scope by including larger and more diverse groups of participants from different institutional and linguistic backgrounds.

A second limitation relates to methodological considerations, particularly the research design. The study employed a carefully constructed longitudinal panel design, yet not without shortcomings. First, no control group was included for practical and ethical reasons (see Section 6.2.1), which is why the observed developments cannot be attributed solely to EPI. Hence, other external factors, such as learners' practice habits, increased exposure to English, as well as other unmeasured influences, may also have contributed to the observed effects. Second, T1, T2, and T3 data were elicited exclusively through a controlled read-aloud task. While this ensured comparability across the three testing points, it did not capture spontaneous or interactive speech. Subsequent research could, therefore, address these issues by incorporating experimental designs with control groups as well as diversifying elicitation methods to include both controlled and spontaneous speech tasks to further explore the impact of EPI.

Another very important limitation concerns the scope of features examined. Only two segmental aspects of the LFC (i.e., vowel length in long monophthongs and word-initial aspiration of fortis plosives) were investigated. This restriction was mainly due to the diagnostic text passage '*The North Wind and the Sun*', which did not represent other LFC features as prominently, as well as the scope of this thesis. As a result, the auditory analysis excluded many other segmental as well as suprasegmental features. Given this narrow focus, caution is required when attempting to generalise the findings to learners' pronunciation development as a whole. A more comprehensive picture could be obtained by expanding future research to a wider range of features, combining segmental and suprasegmental analyses.

Finally, the auditory analysis of participant recordings was conducted by a single rater, which constitutes an additional limitation. Although the rater can be considered a trained listener in assessing pronunciation and high inter-rater reliability was established prior to the rating process, reliance on one rater still entails the risk of potential rater bias. To strengthen the reliability, future studies could, therefore, involve multiple trained raters and ideally combine auditory judgements with acoustic analyses.

Taken together, these limitations point to several promising directions for future research in the field of pronunciation learning. Further studies could investigate more diverse learner groups (e.g. participants differing in L1 background) and adopt experimental designs that incorporate both controlled and spontaneous language production tasks. Expanding the analysis to include additional segmental and suprasegmental LFC features would further enhance our understanding of learners' pronunciation development over time. Addressing these directions would allow future research to build on the present study and provide more nuanced and generalisable insights into the long-term effects of EPI.

9 References

Afshari, S. & Ketabi, S. (2016). Changing paradigms in teaching English pronunciation: A historical overview. *International Journal of Research Studies in Language Learning*, *5*(5). https://doi.org/10.5861/ijrsll.2016.1423

Algethami, G. (2016). Long term effect of phonetic instruction on the production of /p/ by EFL Arab learners: An exploratory study. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2898628

Almusharraf, A. (2024). Pronunciation instruction in the context of world English: exploring university EFL instructors' perceptions and practices. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-03365-y

Atar, C. (2018). Should we teach pronunciation explicitly in L2/EFL classrooms? *International Journal of Contemporary Educational Research, 5*(2), 95–102. https://doi.org/10.33200/ijcer.486044

Archer, A. L. & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching.* Guilford Press.

Avello, P. (2018). Assessing learners' changes in foreign accent during study abroad. In C. Pérez-Vidal, S. López-Serrano, J. Ament, & D. J. Thomas-Wilhelm (Eds.), *Learning context effects: Study abroad, formal instruction and international immersion classrooms* (pp. 131–154). Language Science Press. https://doi.org/10.5281/zenodo.1300622

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. Oxford University Press.

Bai, B., & Yuan, R. (2018). EFL teachers' beliefs and practices about pronunciation teaching. *ELT Journal*, *73*(2), 134–143. https://doi.org/10.1093/elt/ccy040

Baran-Łucarz, M. (2014). The link between pronunciation anxiety and willingness to communicate in the foreign-language classroom: The Polish EFL context. *Canadian Modern Language Review/ La Revue Canadienne Des Langues Vivantes*, *70*(4), 445– 473. https://doi.org/10.3138/cmlr.2666

Berger, A. (2010). German. In R. Walker (Ed.), *Teaching the pronunciation of English as a lingua franca* (pp. 107-110). Oxford University Press.

Berns, M. (2008). World Englishes, English as a lingua franca, and intelligibility. *World Englishes*, *27*(3), 327–334.

Brekelmans, G. (2017). The value of phonetics and pronunciation teaching for advanced learners of English. *Linguistica*, *57*(1), 45–58. https://doi.org/10.4312/linguistica.57.1.45-58

Brown, A. (1991). *Pronunciation models*. Singapore University Press.

Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language. *Journal of Language and Social Psychology*, *22*(2), 190– 209. https://doi.org/10.1177/0261927x03022002003

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A course book and reference guide*. Cambridge University Press.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.

Chaira, S. (2015). Interference of first language in pronunciation of English segmental sounds. *English Education Journal*, *6*(4), 469–483. http://jurnal.unsyiah.ac.id/EEJ/article/download/2856/2718

Charpentier-Jiménez, W. (2019). University students´ perception of exposure to various English accents and their production. *Actualidades Investigativas En Educación*, *19*(2), 1–27. https://doi.org/10.15517/aie.v19i2.36908

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). Routledge.

Council of Europe (2020). Common European framework of reference for languages: Learning, teaching, assessment (Companion Volume). Council of Europe Publishing. https://www.coe.int/en/web/common-european-framework-reference-languages

Creswell, J. W., Plano Clark, V. L. (2018). *Designing and conducting mixed methods research.* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Cucchiarini, C., Strik, H., & Boves, L. (2000). Different aspects of expert pronunciation quality ratings and their relation to scores produced by speech recognition algorithms. *Speech Communication*, *30*, 109–119.

Dalton, C. & Seidlhofer, B. (1994). *Pronunciation*. Oxford University Press.

Dalton-Puffer, C., Kaltenböck, G., & Smit, U. (1997). Learner attitudes and L2 pronunciation in Austria. *World Englishes, 16*(1), 115–128. https://doi.org/10.1111/1467-971X.00052

Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal, 30*(1), 13–45. https://doi.org/10.5070/B5.35963

DeKeyser, R. (2003). Implicit and explicit learning. In M. Long & C. Doughty (Eds.), *Handbook of second language acquisition* (pp. 313–348). Blackwell.

Derwing, T.M. (2010). Utopian goals for pronunciation teaching. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 24–37). Ames, IA: Iowa State University.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379–397. https://doi.org/10.2307/3588486

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, *29*, 359–380. https://doi.org/10.1093/applin/amm041

Derwing, T. M., & Munro, M. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476–90. https://doi.org/10.1017/s026144480800551x

Derwing, T. M. & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, *63*(2), 163–185. https://doi.org/10.1111/lang.12000

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (1st ed.). John Benjamins Publishing Company.

Deterding, D. (2006). The pronunciation of English by speakers from China. *English World-Wide, 27*(2), 175–198. https://doi.org/10.1075/eww.27.2.04det

Dewaele, J. & McCloskey, J. (2015). Attitudes towards foreign accents among adult multilingual language users. *Journal of Multilingual and Multicultural Development 36*(3), 221-238.

Dörnyei, Z. (1998). *Motivation in second and foreign language learning*. Cambridge University Press.

Dörnyei, Z. (2003). *Questionnaires in second language research: construction, administration, and processing*. Lawrence Erlbaum Associates.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Dörnyei, Z., & Dewaele, J.-M. (2022). *Questionnaires in second language research: Construction, administration and processing* (3rd ed.). Routledge.

Ebner, J. (1998). *Wie sagt man in Österreich? Wörterbuch des österreichischen Deutsch*. Bibliographisches Institut. http://ci.nii.ac.jp/ncid/BB03315643

Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, J. Philp, R. Erlam, & H. Reinders (Eds.), *Implicit and explicit*

knowledge in second language learning, testing and teaching (pp. 3–25). Multilingual Matters.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399–423. https://doi.org/10.2307/3588487

Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics.* (5th ed.). Sage, Newbury Park.

Firth, A. (1996). The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics*, *26*(2), 237–259. https://doi.org/10.1016/0378-2166(96)00014-8

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.). *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press.

Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a second language*. Department of Education, Training and Youth Affairs (DETYA).

Ghorbani, M. R., Neissari, M., & Kargozari, H. R. (2016). The effect of explicit pronunciation instruction on undergraduate EFL learners' vowel perception. *Language and Literacy*, *18*(1), 57-70. https://doi.org/10.20360/g2xw2k

Gilakjani, A. P. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science*. http://ijhssnet.com/journals/Vol_2_No_3_February_2012/17.pdf

Gorard, S. (2001). Quantitative Methods in Educational Research: *The Role of Numbers Made Easy*. Continuum.

Gordon, J., & Darcy, I. (2022). Teaching segmentals and suprasegmentals: Effects of explicit pronunciation instruction on comprehensibility, fluency, and accentedness. *Journal of Second Language Pronunciation, 8*(2), 168–195. https://doi.org/10.1075/jslp.21042.gor

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201–223. https://doi.org/10.2307/3588378

Halahan, Y. V., Bezuhla, I. V. & Leshchinska, A. V. (2023). Study of phonological fossilization in the processof the English language acquisition. *Scientific Notes of V I Vernadsky Taurida National University Series Philology Journalism*, *1*(1), 134–140. https://doi.org/10.32782/2710-4656/2023.1.1/23

Harding, L. (2017). What do raters need in a pronunciation scale? The user's view. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment:*

Interdisciplinary perspectives (Vol. 107, pp. 12–34). Multilingual Matters. http://www.jstor.org/stable/10.21832/j.ctt1xp3wcc.6

Harmer, J. (2011). *The practice of English language teaching*. (4th ed.). Pearson Longman.

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek- Klimczak, E, Levey, D, Cunningham, U, & Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, *10*(1)*,* 5-26.

Hodgetts, J. (2020). *Pronunciation instruction in English for academic purposes: An investigation of attitudes, beliefs and practices*. Springer.

House, J. (1999). Misunderstanding in intercultural communication: interactions in English as a lingua franca and the myth of mutual intelligibility. In Gnutzmann, Claus (Ed.). *Teaching and learning English as a global language*. Tübingen: Stauffenburg, (pp. 73- 89).

Hrubes, P. (2008). Some problems with the pronunciation of English typical of native speakers of German (a tentative case study). BA thesis. Masaryk University, Brno.

Hirschi, K. & Kang, O. (2023). How many raters can be enough: G theory applied to assessment and measurement of L2 speech perception. *Language Teaching Research Quarterly*, *37*, 213–230. https://doi.org/10.32038/ltrq.2023.37.12

Jarvis, S. & Pavlenko, A. (2010). *Crosslinguistic influence in language and cognition*. Routledge

Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford University Press.

Jenkins, J. (2002). A sociolinguistically-based, empirically-researched pronunciation syllabus for English as an international language. *Applied Linguistics 23*(1), 83-103.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford University Press.

Kachru, B. B. (1985). *The Alchemy of English: The spread, functions, and models of non-native Englishes.* University of Illinois Press.

Kaltenböck, G. & Seidlhofer, B. (1993). English pronunciation and the German-speaking learner. *Speak Out! 12*, 15-19.

Kaur, J. (2017). Intelligibility in global contexts. In *Routledge eBooks* (pp. 542–555). https://doi.org/10.4324/9781315145006-34

Kelly, G. (2000). *How to teach pronunciation*. Pearson Education.

Kenworthy, J. (1987). *Teaching English pronunciation*. Longman.

Kim, D. & Clayards, M. (2019). Individual differences in the link between perception and production and the mechanisms of phonetic imitation. *Language Cognition and Neuroscience*, *34*(6), 769–786. https://doi.org/10.1080/23273798.2019.1582787

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners?. *The Modern Language Journal 97*(3), 720-744.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon.

Krug, M., & Schlüter, J. (Eds.). (2013). *Research methods in language variation and change*. Cambridge University Press.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369–377. https://doi.org/10.2307/3588485

Levis, J. M. (2015). Learners' views of social issues in pronunciation learning. *Journal of Academic Language and Learning, 9*(1), 64–78.

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation* (1st ed.). Cambridge University Press.

Lightbown, P., & Spada, N. E. (2013). *How languages are learned* (4. ed.). Oxford University Press.

Mairano, P., Bassetti, B., Sokolović-Perović, M. & Cerni, T. (2018). Effects of L1 orthography and L1 phonology on L2 English pronunciation. *Revue Française de Linguistique Appliquée*, *Vol. XXIII*(1), 45–57. https://doi.org/10.3917/rfla.231.0045

Martinsen, R. A., Alvord, S. M. & Tanner, J. (2014). Perceived foreign accent: Extended stays abroad, level of instruction, and motivation. *Foreign Language Annals*, *47*(1), 66–78. https://doi.org/10.1111/flan.12076

Mende C. (2009). The pronunciation of Austrian students of English at university level: A descriptive analysis. [Unpublished master's thesis]. Universität Graz.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481–520. https://doi.org/10.2307/3586981

Moyer, A. (2004). *Age, accent and experience in second language acquisition. An integrated approach to critical period inquiry*. Clevedon, UK: Multilingual Matters.

Moyer, A. (2013). *Foreign accent: the phenomenon of non-native speech*. Cambridge University Press. https://doi.org/10.1017/cbo9780511794407

Moyer, A. (2014). Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics*, *35*(4), 418–440.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*, 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., Derwing, T. M. & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, *28*(01). https://doi.org/10.1017/s0272263106060049

Munro, M. J., & Derwing, T. M. (2011). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*, *44*(3), 316– 327. https://doi.org/10.1017/s0261444811000103

Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, *42*, 258–269. https://doi.org/10.1016/j.system.2013.12.007

Murphy, J. M. (2018). Teacher training in the teaching of pronunciation. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 298–319). Routledge. https://doi.org/10.4324/9781315145006-19

Naeini, M. B., & Adni, Z. (2017). The impact of explicit vs. implicit instruction on pronunciation intelligibility among Iranian EFL learners. *Journal of Studies in Learning and Teaching English, 6*(2), 103–123.

Nagle, C. L. (2018). Examining the temporal structure of the perception–production link in second language acquisition: A longitudinal study. *Language Learning*, *68*(1), 234–270. https://doi.org/10.1111/lang.12275

Nagle, C. & Zárate-Sández, G. (2024). The phonetics and phonology of adult L2 learners after study abroad. In *Cambridge University Press eBooks* (pp. 542–559). https://doi.org/10.1017/9781009105767.025

Nakuma, C. (1998). A new theoretical account of 'fossilization': Implications for L2 attrition research. *IRAL, 36*(3), 247–256. https://eric.ed.gov/?id=EJ573684

Newbrook, M. (1991). Received pronunciation in Singapore: A sacred cow? In A. Brown (Ed.), *Teaching English pronunciation: A book of readings* (pp. 72–84). Routledge.

Nguyen, L. T., Hung, B. P., Duong, U. T. T., & Le, T. T. (2021). Teachers' and learners' beliefs about pronunciation instruction in tertiary English as a foreign language education. *Frontiers in Psychology, 12*, Article 739842. https://doi.org/10.3389/fpsyg.2021.739842

Nguyen, L. T., & Newton, J. (2020). Pronunciation teaching in tertiary EFL classes: Vietnamese teachers' beliefs and practices. *TESL-EJ, 24*(1), Article 1.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528. https://doi.org/10.1111/0023-8333.00136

Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Prentice Hall International Ltd.

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning.* Cambridge University Press.

Pennington, M. C. & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *Modern Language Journal*, *84*(3), 372–389. https://doi.org/10.1111/0026-7902.00075

Pennington, M. C., & Rogerson-Revell, P. (2019). *English pronunciation teaching and research: Contemporary perspectives.* Springer Nature Limited. https://doi.org/10.1057/978-1-137-47677-7

Plag, I.; Arndt-Lappe, S.; Braun, M. & Schramm, M. (2015). *Introduction to English Linguistics.* (3rd edition). De Gruyter Mouton.

Porter, D. & Garvin. S. (1989). Attitudes to pronunciation in EFL. *Speak Out!* 5, 8-15.

Rajadurai, J. (2007). Intelligibility studies: a consideration of empirical and ideological issues. *World Englishes*, *26*(1), 87–98. https://doi.org/10.1111/j.1467-971x.2007.00490.x

Richter, K. (2015). English-medium instruction and pronunciation. A longitudinal case study of Austrian business students at the University of Applied Sciences Vienna. Unpublished PhD Thesis. University of Vienna, Austria.

Richter, K. (2018). Factors affecting the pronunciation abilities of adult learners of English. A longitudinal group study. In *English language education* (pp. 339–361). https://doi.org/10.1007/978-3-319-91917-1_18

Richter, K. (2019). *English-medium instruction and pronunciation: Exposure and skills development* (1st ed.). Multilingual Matters.

Richter, K. (2021). Foreign accent and the role of identity in the adult English as a foreign language classroom. In Berger, A., Heaney, H., Resnik, P., Rieder-Bünemann, A., & Savukova, G. (Eds.). *Developing advanced English language competence: A research-informed approach at tertiary level.* Springer.

Richter, K. & Schölzky, L. (2023). Ready to teach English pronunciation? Student teachers' views and experiences. *CELT Matters 9*, 1-8.

Richter, K., & Weissenbäck, A. (2021). British English or American English? Investigating Austrian English language students' choice of a model accent. In L. McCallum (Ed.),

English language teaching: Theory, Research and Pedagogy. (pp. 89–106). Springer. https://doi.org/10.1007/978-981-19-2152-0_6

Ripper, L., Ciaravino, S., Jones, K., Jaime, M. C. D. & Miller, E. (2017). Use of a respondent-generated personal code for matching anonymous adolescent surveys in longitudinal studies. *Journal of Adolescent Health*, *60*(6), 751–753. https://doi.org/10.1016/j.jadohealth.2017.01.003

Roach, P. (2009). *English phonetics and phonology: A practical course.* Cambridge University Press.

Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. Zed Books.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language Awareness, 20*(1), 45–59.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly, 46*(4), 842–854. https://doi.org/10.1002/tesq.67

Saito, K. & Hanzawa, K. (2018). The role of input in second language oral ability development in foreign language classrooms: A longitudinal study. *Language Teaching Research*, *22*(4), 398–417. https://doi.org/10.1177/1362168816679030

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning, 62*(2), 595–633. https://doi.org/10.1111/j.1467-9922.2011.00640.x

Saito, K., Dewaele, J. & Hanzawa, K. (2017). A longitudinal investigation of the relationship between motivation and late second languages speech learning in classroom settings. *Language and Speech*, *60*(4), 614–632. https://doi.org/10.1177/0023830916687793

Saito, K. (2021). What characterizes comprehensible and native-like pronunciation among English-as-a-second-language speakers? Meta-analyses of phonological, rater, and instructional factors. *TESOL Quarterly*, *55*(3), 866–900. https://doi.org/10.1002/tesq.3027

Sardegna, V. G., & Dickerson, W. B. (2023). Improving the pronunciation of English polysyllabic words through orthographic word-stress rules. In V. G. Sardegna & A. Jarosz (Eds.), *The Routledge handbook of second language pronunciation* (pp. 81–97). Routledge.

Sardegna, V. G., & Jarosz, A. (2023). *English pronunciation teaching: Theory, practice and research findings*. Multilingual Matters.

Scales, J., Wennerstrom, A., & Richard, D. (2006). Language learners' perceptions of accent. *TESOL Quarterly, 40*, 715–738.

Schertz, J. & Clare, E. J. (2019). Phonetic cue weighting in perception and production. *WIREs Cogn Sci., 11*(2), 1-24. https://doi.org/10.1002/wcs.1521

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

Seidlhofer, B. (2011). *Understanding English as a lingua franca.* Oxford University Press.

Sesma, M. G. F., Arias, K. R., & Montaño, S. A. (2022). Exploring EFL teachers' attitudes toward the use of technology to teach pronunciation. *Mextesol Journal*,*46*(2), 1– 14. https://doi.org/10.61871/mj.v46n2-8

Setter, J. & Jenkins, J. (2005). State-of-the-art review article. *Language Teaching, 38*, 1–17. https:// doi: 10.1017/S026144480500251X

Sharma, L. (2021). Significance of teaching the pronunciation of segmental and suprasegmental features of English. *Interdisciplinary Research in Education, 6*(2), 63–78.

Smit, U., & Dalton, C. (2000). Motivational patterns in advanced EFL pronunciation. *IRAL*, *38*, 229–246.

Sudina, E. (2021). Study and scale quality in second language survey research, 2009–2019: The case of anxiety and motivation. *Language Learning*, *71*(4), 1149–1193. https://doi.org/10.1111/lang.12468

Suzukida, Y. (2021). The contribution of individual differences to L2 pronunciation learning: Insights from research and pedagogical Implications. *RELC Journal*, *52*(1), 48–61. https://doi.org/10.1177/0033688220987655

Swan, M., & Smith, B. (Eds.). (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.

Szpyra- Kozłowska, J. (2015). *Pronunciation in EFL instruction: a research-based approach*. Multilingual Matters.

Tatzl, D. (2011). English-medium masters' programmes at an Austrian university of applied sciences: Attitudes, experiences and challenges. *Journal of English for Academic Purposes, 10*(4), 252–270. http://dx.doi.org/10.1016/j.jeap.2011.08.003

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning*, *41*(2), 177–204. https://doi.org/10.1111/j.1467-1770.1991.tb00683.x

Thomson, R. I. & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics 36*(3), 326-344. https://doi.org/10.1093/applin/amu076

Tsunemoto, A. & McDonough, K. (2020). Exploring Japanese EFL learners' attitudes toward English pronunciation and its relationship to perceived accentedness. *Language and Speech*, *64*(1), 24–34. https://doi.org/10.1177/0023830919900372

Uchihara, T., Webb, S., Saito, K. & Trofimovich, P. (2022). Frequency of exposure influences accentedness and comprehensibility in learners' pronunciation of second language Wwrds. *Language Learning*, *73*(1), 84–125. https://doi.org/10.1111/lang.12517

Uzun, T. (2022). Preservice English teachers' attitudes and beliefs about learning and teaching pronunciation. *Dil Dergisi*, *173*(2), 22–42. https://doi.org/10.33690/dilder.1040120

Vong, C., Sam, R. & So, S. (2024). Cambodian EFL students' perception of factors influencing pronunciation. *LLT Journal a Journal on Language and Language Teaching*, *27*(2), 985–996. https://doi.org/10.24071/llt.v27i2.8031

Wagner, E., & Toth, P. D. (2017). The role of pronunciation in the assessment of second language listening ability. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment* (pp. 72–92). Multilingual Matters.

Walker, R. (2010). *Teaching the pronunciation of English as a lingua franca*. Oxford University Press.

Wang, Y. & Wen, X. (2023). Nativeness versus intelligibility as goal of English pronunciation teaching in China: Changing attitudes in national syllabi and curriculum standards. *Asian-Pacific Journal of Second and Foreign Language Education*, *8*(1). https://doi.org/10.1186/s40862-023-00189-2

White-Hautzinger, V., Djenadic, M., Rumpler, C. & Fiala, M. (2025). The effects of explicit pronunciation instruction: A study of advanced Austrian EFL students. In A. Kirkova-Naskova, P. Humanez-Berral & A. Henderson, (Eds.). *Proceedings of the 8th International Conference on English Pronunciation: Issues and Practices* (pp.118-128) HAL open science. https://doi.org/10.5281/zenodo.16696495

Wieden, W., & Nemser, W. (1991). *The pronunciation of English in Austria: A developmental and regional study*. Narr.

Yenkimaleki, M. & Van Heuven, V. J. (2019). Prosody instruction for interpreter trainees: Does methodology make a difference? An experimental study. *Across Languages and Cultures*, *20*(1), 117–133. https://doi.org/10.1556/084.2019.20.1.6

Yule, G. (2010). *The study of language* (4th edition). Cambridge University Press.

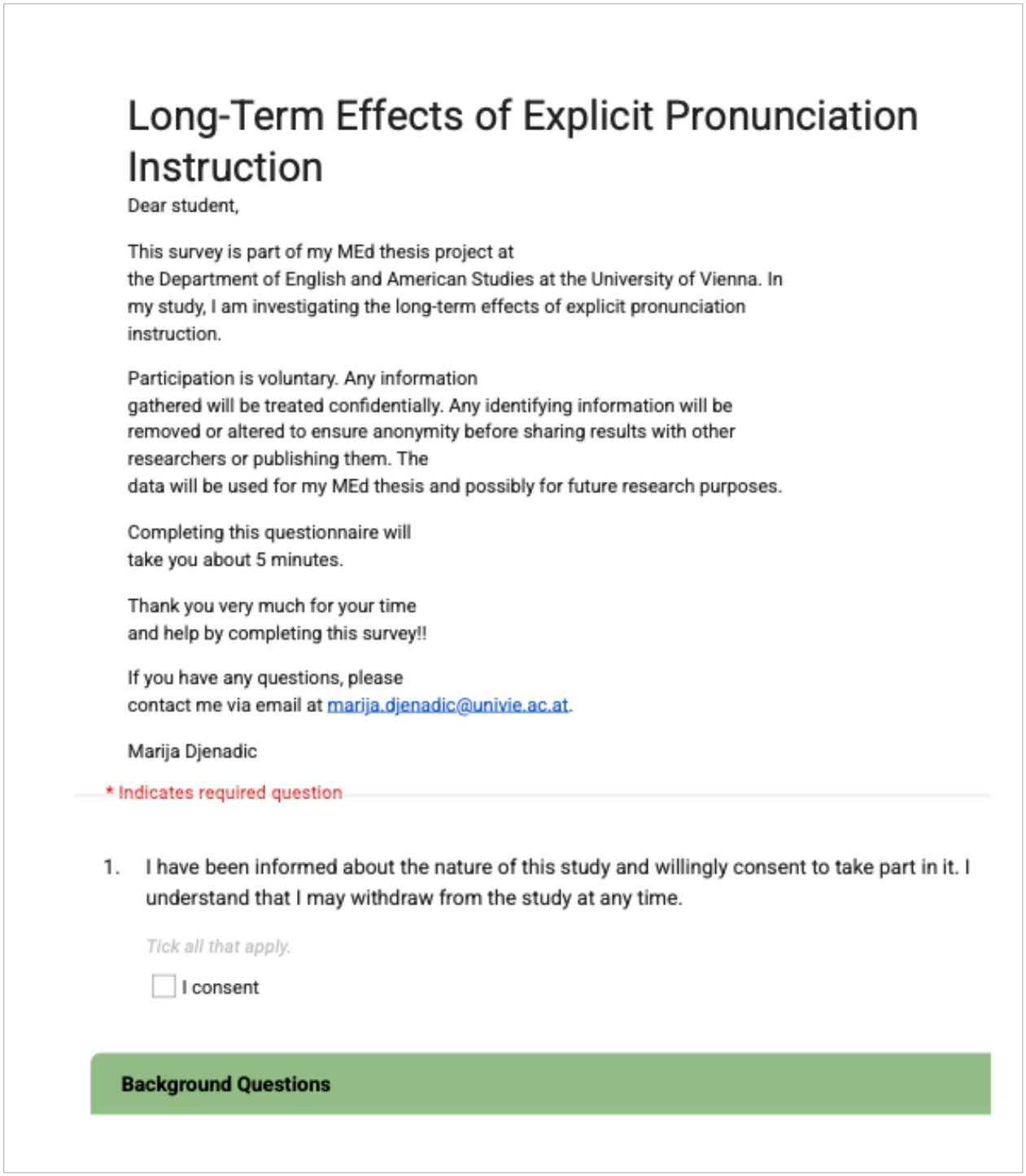
Zhang, R., & Yuan, Z. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition, 42*(4), 905–918. https://doi.org/10.1017/S0272263120000121

Zhong, W. (2019). Pronunciation rating scale in second language pronunciation assessment: A review. *Journal of Language Teaching and Research, 10*(1), 141–149. https://doi.org/10.17507/jltr.1001.16

Zielinski, B. (2015). The segmental/suprasegmental debate. In M. Reed, & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 397-412). (Blackwell handbooks in linguistics). Wiley-Blackwell, Wiley. https://doi.org/10.1002/9781118346952.ch22

10 Appendix

10.1 Web-based survey



Long-Term Effects of Explicit Pronunciation Instruction

Dear student,

This survey is part of my MEd thesis project at the Department of English and American Studies at the University of Vienna. In my study, I am investigating the long-term effects of explicit pronunciation instruction.

Participation is voluntary. Any information gathered will be treated confidentially. Any identifying information will be removed or altered to ensure anonymity before sharing results with other researchers or publishing them. The data will be used for my MEd thesis and possibly for future research purposes.

Completing this questionnaire will take you about 5 minutes.

Thank you very much for your time and help by completing this survey!!

If you have any questions, please contact me via email at marija.djenadic@univie.ac.at.

Marija Djenadic

* Indicates required question

1. I have been informed about the nature of this study and willingly consent to take part in it. I understand that I may withdraw from the study at any time.

Tick all that apply.

☐ I consent

Background Questions

2. **Please create a code using the following format:** *
   - first letter of your first name
   - first letter of your last name
   - first two digits of your birthdate

   e.g. Mia Smith, born on 08 October --> "MS08"

3. Your Matriculation Number *

4. Age *

   Mark only one oval.

   - 19-24
   - 25-30
   - 31-34
   - 35+

5. What is/are your first language(s)? *

6. At what **age** did you **start learning English**? *

7. How often do you **speak** English in the following contexts? *

Mark only one oval per row.

	never	rarely	occasionally	sometimes	often	usually	alway
with your friends	◯	◯	◯	◯	◯	◯	◯
with your family	◯	◯	◯	◯	◯	◯	◯
at home	◯	◯	◯	◯	◯	◯	◯
at university	◯	◯	◯	◯	◯	◯	◯
at work	◯	◯	◯	◯	◯	◯	◯
in online communication	◯	◯	◯	◯	◯	◯	◯

8. Have you spent time in an **English-speaking country** after completing PPOCS1? *

Mark only one oval.

◯ Yes
◯ No

9. If yes, **where** did you stay and for **how long**?

PPOCS 1

10. What were your **main goals** for PPOCS 1? *

Tick all that apply.

- [ ] to speak like a native speaker
- [ ] to have the best possible accent
- [ ] to speak more fluently
- [ ] to be better understood (intelligible)
- [ ] to improve my pronunciation
- [ ] to pass PPOCS 1
- [ ] Other: _____

11. Since PPOCS 1, have your **goals** related to pronunciation **changed**? *

Mark only one oval.

- ◯ Yes
- ◯ No

12. If yes, please describe **how** your goals have changed.

_____

_____

_____

_____

_____

13. How satisfied were you with your pronunciation **immediately after** completing PPOCS 1? *

Mark only one oval.

	1	2	3	4	5	
very	◯	◯	◯	◯	◯	very satisfied

14. How satisfied are you with your pronunciation **right now**? *

Mark only one oval.

	1	2	3	4	5	
very	◯	◯	◯	◯	◯	very satisfied

15. Since completing PPOCS 1, do you feel that your acquired accent has ... *

Mark only one oval.

- ◯ deteriorated?
- ◯ stayed the same?
- ◯ improved?
- ◯ Other: _____

Practice Habits after PPOCS 1

16. Since completing PPOCS 1, to what extent have you **actively worked** on your pronunciation?

Mark only one oval.

	1	2	3	4	5	
not	◯	◯	◯	◯	◯	a lot

17. Since completing PPOCS 1, to what extent have you **used** the **following pronunciation practices**?

Mark only one oval per row.

	not at all	rarely	sometimes	often	a lot
watching English-language media (e.g. YouTube, films, podcasts)	◯	◯	◯	◯	◯
talking to (native) speakers of English	◯	◯	◯	◯	◯
reading aloud in English	◯	◯	◯	◯	◯
using pronunciation apps/tools	◯	◯	◯	◯	◯
attending additional pronunciation workshops/courses	◯	◯	◯	◯	◯
mimicking someone's accent	◯	◯	◯	◯	◯
looking up words in a pronunciation dictionary	◯	◯	◯	◯	◯
watching online pronunciation tutorials (e.g. YouTube)	◯	◯	◯	◯	◯

18. If there are **any other practices** you have used since completing PPOCS 1, please **describe** them and **explain why** you used them.

Exposure after PPOCS 1

19. Which English accent(s) have you been **most exposed** to since completing PPOCS 1? *

Tick all that apply.

- [ ] British English
- [ ] American English
- [ ] Australian English
- [ ] Irish English
- [ ] Non-native English (e.g. Austrian)
- [ ] Other: _____

20. How strongly do you **agree** or **disagree** with the following statements? *

Mark only one oval per row.

	strongly disagree	disagree	neutral	agree	strongl agree
I frequently speak English outside of academic settings (e.g. uni).	◯	◯	◯	◯	◯
I often listen to or watch English-language media with the intention of improving my pronunciation.	◯	◯	◯	◯	◯
I regularly pay attention to how others pronounce words in English when I listen to them.	◯	◯	◯	◯	◯

Other Factors

21. To what extent do you think the following **factors** have **influenced** your **pronunciation** since completing PPOCS 1?

Mark only one oval per row.

	not at all	a little	somewhat	quite a bit	very much
increased exposure to English (e.g. media, social interactions)	◯	◯	◯	◯	◯
opportunities to speak English regularly (e.g. studies, travel, exchange programme)	◯	◯	◯	◯	◯
participation in additional pronunciation or language courses	◯	◯	◯	◯	◯
personal motivation	◯	◯	◯	◯	◯
feedback on pronunciation from native or proficient speakers of English	◯	◯	◯	◯	◯
practice pronunciation through self-study	◯	◯	◯	◯	◯

22. Is there anything else you would like to share about your pronunciation development since completing PPOCS 1?

Zusammenfassung

Für Lernende von Englisch als Fremdsprache (EFL) ist Aussprache ein zentraler Bestandteil kommunikativer Kompetenz, insbesondere im tertiären Bildungsbereich, wo Englisch häufig in akademischen und beruflichen Kontexten verwendet wird. Forschung hat gezeigt, dass expliziter Ausspracheunterricht (EPI; Explicit Pronunciation Instruction) hierbei eine entscheidende Rolle spielt und sowohl phonetische Genauigkeit als auch kommunikative Effektivität verbessern kann. Bisher haben sich jedoch die meisten Studien auf kurzfristige Effekte von EPI konzentriert, sodass nur begrenzt belegt ist, ob diese Fortschritte langfristig bestehen bleiben. Weiters ist wenig darüber bekannt, inwieweit fortgeschrittene EFL-Lernende mit Deutsch als Erstsprache phonetische Eigenschaften beibehalten oder verlieren, die für die Sprachverständlichkeit entscheidend sind, und welche individuellen und kontextuellen Faktoren ihre Ausspracheentwicklung beeinflussen. Folglich ist die langfristige Wirkung von EPI für diese Gruppe von Lernenden bislang weitgehend unerforscht. Die vorliegende Masterarbeit untersucht daher die langfristigen Effekte von EPI auf die Produktion zweier segmentaler Merkmale des Linga Franca Core (i.e., Vokallänge und Aspiration) bei fortgeschrittenen österreichischen EFL-Studierenden. Im Rahmen eines longitudinalen Mixed-Methods-Designs wurden Daten von 28 Bachelorstudierenden der Universität Wien erhoben, die einen EPI Kurs besuchten. Ihre Ausspracheentwicklung wurde zu drei Zeitpunkten erfasst: vor Beginn des Kurses (T1), nach Abschluss des Semesters (T2) und rund sechzehn Monate später (T3). Ergänzend wurde beim T3 ein Online-Fragebogen durchgeführt, der weitere Einflussfaktoren auf die Ausspracheentwicklung der Proband*innen erhob. Die Ergebnisse zeigen deutliche kurzfristige Fortschritte in der Vokallänge und bei der Aspiration, wobei die Verbesserungen bei der Vokallänge besonders ausgeprägt waren. Zum Zeitpunkt T3 blieb ein Großteil dieser Fortschritte erhalten. Zwar kam es zu einem leichten Rückgang in der Genauigkeit, die Mittelwerte lagen jedoch weiterhin über dem Ausgangsniveau (T1). Die Befragungsergebnisse verdeutlichten zudem, dass die langfristige Entwicklung weniger auf gezieltem Üben beruhte, sondern stärker durch alltägliche Formen der Sprachkontakts geprägt war, etwa durch Medienkonsum oder gelegentliche Interaktion. Insgesamt zeigt die Studie, dass EPI nachhaltige Fortschritte in der Sprachverständlichkeit von EFL-Lernenden ermöglichen kann.