



UiT The Arctic University of Norway

Department of Language and Culture

Crosslinguistic Influence in L3A

Acquisition of the English vowel system by Farsi and Gilaki speakers

Siavash Khoshrouz

Master's thesis in English Language Acquisition and Multilingualism, ENG-3991, November 2021

Acknowledgment

Words cannot form how thankful I am to my supervisor Natalia Mitrofanova for sharing with me her expertise and experience and most of all for being there every single time I needed help. I consider myself extremely fortunate to have had you all along the way. Nevertheless, I am merely responsible for any shortcomings in this thesis.

Also, my special thanks goes to Martin Schweinberger who helped me a lot with the data analysis and how to work with tools I used in this thesis.

Table of Contents

1	Introduction	7
2	Literature Review	9
2.1	Crosslinguistic Influence	9
2.1.1	Full Transfer and Full Transfer Potential	9
2.1.2	Phonology and CLI	14
2.1.3	Age and sensitive period in phonology	15
2.2	Previous Studies	16
2.2.1	Direction of the CLI, perception and production	16
2.2.2	Facilitative and non-facilitative CLI	18
2.2.3	What are the criteria for an acceptable production?.....	19
2.3	Farsi, Gilaki and English	23
2.3.1	Farsi and English.....	23
2.3.2	Gilaki.....	25
2.3.3	Code-switching between the L and H variety	31
2.4	Summary.....	32
3	Methodology	35
3.1	The design of the experiment	35
3.2	The production task	35
3.3	Controlling for external factors	37
3.4	The background questionnaire (BQ)	38
3.5	The attitude test (TA)	39
3.6	The participants	41
4	Data Analysis	43
4.1	Introduction	43
4.1.1	Assessing the duration and vowel quality	45
4.2	Coding the data with Praat.....	46

4.3	Data analysis with R	50
4.4	Formant Ceiling	51
5	Results	54
6	Results from the attitude tests	59
6.1	AT1	60
6.2	AT2	61
7	Discussion	63
7.1	Duration	65
7.2	F1 and F2	65
7.3	Effect of Condition	67
7.4	Effect of Proficiency	68
7.5	Attitude, Intentionality and Transfer	69
7.6	Crosslinguistic Influence	72
8	Conclusion	77
8.1	Extra Note	80
	Works cited	83
	Appendix A	89
	Appendix B	89
	Appendix C	90
	Appendix D	92

1 Introduction

This study aims to investigate the production of English vowels by L1 speakers of two closely related Iranian varieties, Standard Farsi and Gilaki. The research idea was initiated based on the phonological differences between Gilaki and Standard Farsi. In contrast to Farsi, Gilaki phonologically distinguishes between several tense and lax vowels, making the Gilaki system closet to the English one and opening for a possibility of facilitative cross-linguistic from Gilaki for Gilaki-Farsi bilinguals. However, there is tangible scarcity of studies on Gilaki, and its comparison to Standard Farsi in particular. Despite the differences in grammar, lexis, phonology, etc. between the two languages, we do not know if Gilaki users behave like monolinguals or bilinguals in certain linguistic contexts. The primary objective in this investigation is to examine phonological differences in the production of L2 English tense and lax vowels between Gilaki-Farsi and monolingual Farsi speakers. The phonological differences between Gilaki and Standard Farsi, according to the only available comprehensive study on Gilaki (Rastorgueva & Lockwood 2012), are mostly in the vowel duration. In addition, we can stipulate that the differences should also include the spectral dimensions. In fact, I hypothesise that Gilaki, as the more-marked variety, has a significant phonological common ground with English, such as differentiation of minimal pairs according to spectral differences. Thus, the two research questions below are going to be explored in this study:

1. Do Gilaki-Farsi speakers demonstrate a different duration of vowels compared to monolingual Standard Farsi speakers, in their English productions?
2. Do Gilaki-Farsi speakers demonstrate a more substantial differentiation in F1 and F2 value of the lax vs tense vowels compared to the Standard Farsi speakers, in their English productions?

My prediction was a 'yes' to both questions and since this can be categorised as a positive transfer I will subsequently argue about the importance of reinforcing positive transfers from learners' L1 and/or L2 and why it is beneficial to investigate and explore such potentials.

To predict the potential answers to these questions requires a review of the theories in crosslinguistic influence (CLI). However, these predictions can vary whether English is considered as the L2 or the L3 depending on whether Gilakai-Farsi users are regarded as monolinguals (having one system with small differentiation between two closely-related varieties) or bilinguals (with two separate linguistic systems). This is one example why studies

in CLI tend to be more complicated in closely related varieties especially if they have been under-investigated before. Therefore, I hope the findings of this study can both contribute to language acquisition and the role of CLI in phonological acquisition of additional languages in general, and fill the knowledge gap related to an understudied minority language/variety, Gilaki.

In the upcoming chapters, I will begin by reviewing two important theories of cross-linguistic influence (CLI) in additional language acquisition (Full Transfer and Full Transfer Potential) and try to explain how they may or may not predict certain outcomes in this study. I will refer back to them in the discussion section and once more associate the research questions with the two theories in the light of the results achieved.

Next, I will narrow the focus down to phonology, as the broad theme of this study, and review some important points such as the sensitive period and the role of age, and discuss how they can affect the acquisition of L2 / L3 phonology. Then, I will go over a few previous studies in L3 phonological acquisition and their results to establish a meaningful link between the CLI theories and the upcoming results of this study. However, phonological acquisition is mostly associated with different paradigms each of which may focus on specific aspects of this process. As a consequence, this appeared to be essential to review related issues such as intelligibility, comprehensibility, their following disputes and clarify what position I take by advocating that facilitative CLI should be explored more. To this end, I will also review how second language acquisition methodologies have benefited from investigations on facilitative and non-facilitative CLI.

The next section will focus on the empirical domain of the thesis: the vowel systems of English, Farsi and Gilaki, and will demonstrate the phonological characteristics, specifically the vowel characteristics of the three varieties of this study. A special section has been allocated to the social relation between Standard Farsi and Gilaki and how low linguistic status has directly affected the linguistic changes and influences within and between the two varieties. I will specifically discuss how Gilaki phonology has been affected by social factors in a way that makes it more complicated to track its effect on L3 production. As a result, I will discuss the importance of considering a combination of linguistic and non-linguistic factors when investigating CLI.

2 Literature Review

2.1 Crosslinguistic Influence

2.1.1 Full Transfer and Full Transfer Potential

It has been about four decades since the term cross-linguistic influence (CLI) came to exist by Sharwood-Smith (Sharwood-Smith and Kellerman 1986) encompassing a wide range of interactions between the languages in language acquisition and processing. This theory-neutral term, as Sharwood Smith puts it, has been an umbrella term to bring together a substantially growing body of research that aim to investigate the interaction between L1, L2 and L3 (and Ln) at various layers of language, its components and subsystems. Perhaps, the ease of mobility, socialisation, education and more recently the increasing rate of immigration coupled with individuals' desire to learn more languages has put more varieties of language in closer contact with one another. And as a result, an ever growing body of research has been dedicated to empirical studies of CLI cases and proposing theories that could explain the phenomenon.

I will review two specific theories that have been put forth to account for CLI in Third Language Acquisition and clarify how and in what direction(s) the influences go across the languages in a multilingual mind. Although there may still be “uncertainty, disagreement and misunderstanding” about these three theories (Westergaard 2021a and Slabakova 2021), it is essential to review each in summary so that, subsequently, a link can be drawn between the results of the present study and any of these theories and hopefully contribute to them as well

1. **The Typological Primacy Model/ Full Transfer**, in summary, argues that the L3 learner reduplicates or makes a full copy of one of the previously acquired languages, especially because it is more economical and preferable rather than having the whole language repertoire active all the time for partial transfers (Schwartz and Sprouse 1996: 41, Rothman et al. 2019: 157, Westergaard 2021a). Full Transfer seems to predicts results in L2 acquisition since L2 is only affected by only one other language involved. However, other studies show that in L3 acquisition, the influence can come from both L1 and L2.

This model puts substantial emphasis on the similarities between languages and argues that the L3 learner looks for those similarities and reduplicates them in the new language. However, TPM has been argued not to explicitly define or clarify what those similarities are and to what extend a language property should be similar to be

transferred (Ionin 2021). An example of this is a study by Archibald and Yousefi (2018) where they show that despite the superficial dissimilarity between English and Persian in having the left-edge onset clusters, such as *cl*, *pl* and *dr* present in English, Persian learners can easily manage to benefit from their right-edge CC clusters to parse the English left-edge CC/onset clusters (the appendices). So they argue that, this could complicate the definition of “similarity” since in the English-Persian example, the absence of the onset clusters indicates a difference, rather than a similarity. However, in practice and beneath the superficial difference, there lies a potential that initially appears like a difference, however, exposed to the new situation, it easily accommodates itself and acts like a similarity. Therefore, one might argue that (superficial) “proximity” may not be a straightforward criterion to predict CLI between two or more languages.

The second argument about the TMP model is that it simply prioritises the similarity between the L3 and previously acquired languages but does not take into account the role that other factors, such as proficiency, metalinguistic knowledge, universal preferences, etc. *do* [my emphasis] play and can even outweigh the similarity of a property between two languages (see Westergaard 2021a). In other words, CLI appears to be an outcome of a multifactorial situation and proximity, though an outstanding factor, is only one of the many factors (Jarvis and Pavlenko 2008, 20) which can predict the source(s) and the extent of CLI.

- 2. The Linguistic Proximity Model/Full Transfer Potential**, on the other hand, lays stress on (all) linguistic properties being potentially transferrable if they receive enough support and/or input from the other languages (Westergaard, Mitrofanova, Mykhaylyk & Rodina 2017). In other words, even typologically dissimilar varieties can be a source of transfer given certain circumstances which are, more often than not, present and affecting the L3 acquisition in one way or another.

As one of the recent theories, Westergaard (2021a) introduced Full Transfer Potential (FTP), a concept in line with the core idea of the Linguistic Proximity Model, as an alternative to the traditional SLA notion of Full (Wholesale) Transfer. Westergaard argues that the acquisition/development of language is more of a journey but with “small steps” and also highlights that L2/Ln learners have differences, compared to L1 learners, which makes the transfer more likely to be a “property-by-property” process, rather than

‘full’ or ‘wholesale’ copying of the grammar of one of the previously acquired languages. FTP places more emphasis on L3/Ln acquisition stating that not necessarily “everything” but “anything” from the learners’ repertoire can transfer (See Westergaard 2021a for a review of 1 and 2).

Above I briefly mentioned the work by Jarvis and Pavlenko (2008, 20) in which they propose a classification of “CLI types across ten dimensions” (a part of this classification is illustrated Table 1 below) as an attempt to account for different instantiations of CLI. Though they indicate that the classification might look a little complicated, it also has the potential to help us narrow down our focus on specific aspects of CLI and to be able to relate them to one theory or another.

Area of language/Knowledge use	Mode	Form
<ul style="list-style-type: none"> • Phonological • Orthographic • Lexical • Semantic • Morphological • Syntactic • Discursive • Pragmatic • Sociolinguistic 	<ul style="list-style-type: none"> • Productive • Receptive 	<ul style="list-style-type: none"> • Verbal • Non-verbal
Directionality	Channel	Manifestation
<ul style="list-style-type: none"> • Forward • Reverse • Lateral • Bi- or multi-directional 	<ul style="list-style-type: none"> • Aural • Visual 	<ul style="list-style-type: none"> • Overt • Covert
Intentionality	Cognitive level	Outcome
<ul style="list-style-type: none"> • Intentional • Unintentional 	<ul style="list-style-type: none"> • Linguistic • Conceptual 	<ul style="list-style-type: none"> • Positive • Negative
		Type of Knowledge
		<ul style="list-style-type: none"> • Explicit • Implicit

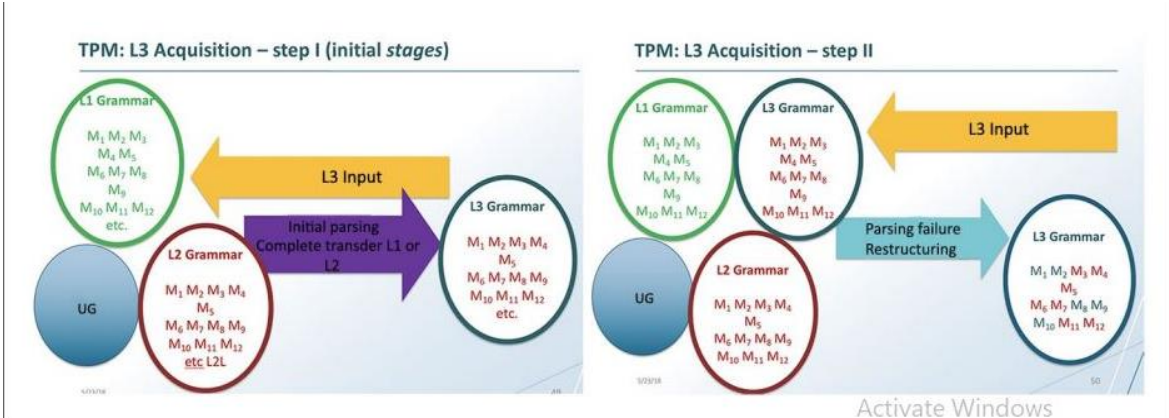
Table 1- Characterisation of CLI types across 10 dimensions (Jarvis and Pavlenko 2008).

Having argued that CLI can be, and is, affected by different factors, it is essential to be aware whether any of these factors ultimately has a more significant weight on the final outcome. As I will discuss in the upcoming sections, we can observe that the attitude (intentionality) of a group of learners can have an impact on the acquisition of a new language. Next, I would like to establish a meaningful relation between one of these ten dimensions, entitled *Intentionality*

and FTP. As proposed by Jarvis and Pavlenko (2008), intentionality (or – conversely - unintentionality) can have an impact on CLI, and adds to the more traditional approach that views transfer as more of an automatic and unintentional phenomena (Fuster and Neuser 2019). Describing transfer as, also, an intentional “communicative strategy” (Jarvis and Pavlenko 2008, 24) opens (or at least highlights) a rather broad spectrum of potential outcomes and empirical research avenues. For example, how is this possible for an L3 learner to intentionally transfer a property from one language to any other language according to TPM/FT which seem to see the CLI as an automatic and one-dimensional process? How is this possible to create a full copy of L1 or L2 while the learner is assumed to have a freedom of intentional choice to transfer a property from any of the languages? Therefore, it is essential to review FTP to have a more detailed view of the L3 acquisition.

Westergaard (2021a) proposes a theory of FTP as an alternative to or a modification of the FT (figure 2 and figure 3). The argument is that rather than an absolute transfer of every property (FT), there is a potential for **any property** [my emphasis] to be transferred. Furthermore, FTP’s contention is that this is rather impractical to gain enough evidence for FT based on L2 acquisition while on the contrary L3/Ln acquisition pretty much supports the property-by-property model of transfer, showing cases where selective groups of features from all the previously learnt languages are present in the new language. However, this view has also been debated for considering both CLI and transfer as one phenomenon, with no distinction (Wrembel 2021). My interpretation of this unification is that CLI and transfer could at least be described as the same phenomenon at certain times yet leaving some room for cases where a distinction has to be drawn. This will become a critical aspect in the following sections since I will be assuming that the potential borrowing of some phonological features from one’s L1 (or L2) in the L2 (or L3) is a case of transfer or CLI with no distinction.

Figure 2 - L3 acquisition according to Full Transfer/TPM (Westergaard 2021b).



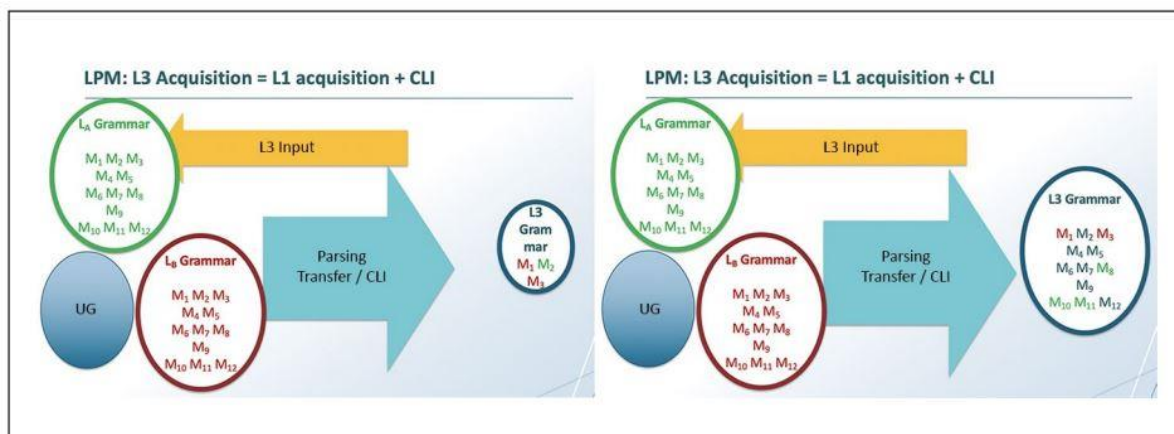


Figure 3 – L3 acquisition according to Full Transfer Potential/LPM (Westergaard 2021b).

More and more, studies are indicating that the cross-linguistic influence is an inevitable outcome in situations when a new language (or additional language) is learned on top of one's L1 (and L2). The influence can be observed in one or more of the language subsystems in several directions, contrary to the traditional view where the influence was believed to happen only from one language to the additional language(s). (Wrembel 2015, 29-47)

One well-known case of CLI, even to laymen, is the use of vocabulary from one language in another, which might be classified as borrowing or code switching. To clarify whether a non-native vocabulary use is a case of borrowing (whether well-established or not) or code switching may not always be so straightforward yet what we can be sure of is that they are all the outcome of an interaction of two or more languages in one mind.

- 1) Jeg liker å spise kabab.
[I like to eat kabab]

Sentence (1) in Norwegian and the English version in the brackets are both using the borrowed Persian word “kabab” meaning, “to fry”. Similarly, more recent loanwords and instances of code-switching can be seen in the language of the youth, in Norway for instance (Roland Kuitert 2013) which indicates that lexical borrowings are common and not limited to a few cases in the history of the language.

Considering CLI in grammar, for an L1 Norwegian speaker learning English as an L2, we can observe transferring the Norwegian V2 structure into English:

- 2) Now must vi go.
[Nå skal vi gå]

- 3) That do you.
[Dette gjør du]

In the examples above, we can see that the sentences in the brackets, though in English, follow the V2 Norwegian structure.

CLI comprises a wide range of phenomena including “transfer”, “interference”, “borrowing”, “language loss” and “avoidance” (Wrembel 2015, 45, Sharwood Smith and Kellerman 1986). Any of the two categories of vocabulary and grammar, mentioned above, are so broad and each falls into series of smaller subcategories of studies. However, the focus of this paper is mainly on the other broad domain of language: phonology, narrowed down to production. It is therefore crucial to review FTP, intentionality and CLI, in regards to phonology to build the foundation for the subsequent methodology section.

2.1.2 Phonology and CLI

The world as we have been experiencing it promises more encounter of languages, at least up in our heads, thanks to the ease of mobility and a demand to learn more languages (Unsworth 2013). Whether one is an early or late bi/multilingual, CLI/transfer is likely to be a typical feature of any additional language acquisition. The historical evidence of such encounter indicates that any of these languages might undergo alterations of different scales, such as the example of the loanwords discussed earlier. However, does it also mean that the way that one sounds, the way that we associate languages with their overall sonic configuration/voice setting can change? The answer to this question does not seem to be very straightforward and it definitely calls for a substantial body of empirical investigations and studies. Nevertheless, what we do know is that though the cross-linguistic influence is a core feature in additional language acquisition, there are multiple linguistic and non-linguistic factors affecting CLI in varying directions and fashions. Overall, albeit this may not lead to a yes/no answer to predictionist questions, specific empirical studies can contribute to our enhanced understanding of the nature of the CLI and subsequent linguistic development of a new language.

The emergence of the Third Language Acquisition (L3A) as an independent field, has led to a new and broader understanding of the ways CLI can occur. L3 phonology in particular, has been an essential part of this new understanding since it has been able to indicate that the source of influence and transfer can stem from all the previously acquired languages at the same time (Wrembel 2015, Wrembel et al.2019a and 2019b). As mentioned earlier, most CLI theories tend to rely on the results achieved from studies on areas other than phonology, such as

grammar and syntax, and this has certainly been one of the main reasons for those theories to take a rather strong position siding with the Whole Transfer models (see Schwartz & Sprouse 2021). However, an increasing body of investigations in L3 phonology has revealed that not all language components, and certainly not language as a whole, could be associated with one model, especially if that model is somehow restrictive in the way it advocates the CLI. Therefore, L3 (Ln) phonology acquisition has been one of the many ways to “various interpretations” of CLI (Wrembel 2021b).

Contracted down to my main focus in here, a large body of research on phonological acquisition of an L2/L3 has focussed on such important lines of research as the effect of age on the ultimate attainment in phonology, and the similarities and differences in production vs perception of L2 phonology. The next sections briefly summarize the core findings of these lines of investigation.

2.1.3 Age and sensitive period in phonology

The traditional point of view regarding the effect of age on language acquisition in general was the *critical period* described as an early phase in a child’s language development process after which the door to most, if not all, changes is closed. On the other hand, the sensitive period, according to growing evidence, describes a phase through which one is more sensitive to input but the possibility to post-sensitive period still remains open (White et al 2013). Nevertheless, it is mostly an undisputed agreement among scholars that age is a strong predictor in many features of language acquisition, especially L2, L3 and Ln. One domain that is believed to be heavily influenced by age, regarding the sensitive period, is phonology. Most scholars consider the end of the critical period from 12 to 18 years of age, however, many of them agree that the period is even shorter for phonology, ending at about the age 9 (Vanhove 2013). The emphasis on this period can be observed in how Kuhl and colleagues (2008, 2017) describe the period: a transition from “the citizens of the world” to “native language specialists”, which occurs due to certain neurological changes in the brain (called ‘pruning’ of neural connections). This is a transition through which infants little by little become ‘deaf’ (*insensitive*) to phonological contrasts that are absent in the language(s) they hear. Though brief, it is clear that the sensitive period slightly differs in its status from the critical period in its flexibility, it still dramatically highlights the critical role of the early phase of language development. Nevertheless, Vanhove (2013) argues that the fact that adults show a faster rate at the initial stage of the L2 acquisition might be a sign of adults’ “sensitivity to language input”. The author also puts emphasis on the fact that critical point is mostly associated with comparing L2 learners with monolinguals which

naturally is a comparison of the L2ers with the native-like criterion. However, as discussed earlier, it is the bilingualism, rather than monolingualism, that is the norm today.

From the neurological point of view, White and colleagues (2013) argue that after the sensitive period, as opposed to the strict and rather dead-end critical period, the “structure and efficiency of pre-existing circuits” (*of brain*) can be refined and changed for the new input to through a top-down process which requires “attention”. Therefore, when approaching CLI in L2 or L3 acquisition, a more untapped potential of influence could be expected, especially if coupled with instructed training.

Overall, it should be noted that the sensitive period describes the general situation, which can manifest itself differently in perception vs production, and could be affected by predictors such as instructed training and learners’ attitude.

2.2 Previous Studies

2.2.1 Direction of the CLI, perception and production

A large number of bi/multilinguals are late learners who have already passed the sensitive period, however, they might still be able to learn, distinguish and/or even produce the sounds that were absent in their previously acquired language(s). Yet, it can be said that studies on L3, though propelling at a lively and rapid pace, are still young compared to the ones of L2 and most models have been proposed to account for L2 acquisition, and may only partly provide explanations for L3/Ln acquisition situations. In addition, even in relation to L3, most studies (and theories, such as FT) have focused on grammar, syntax, and lexicon, while there seem to be somewhat less studies focussing on phonology. How multilinguals (as opposed to bilinguals) acquire sounds of the new language(s) remains underexplained and we do not exactly know the specifics about multiple languages in a multilingual’s repertoire plays out in the CLI (Wrembel et al. 2019, Cabrelli Amaro and Wrembel 2016). In what follows I will review several studies on L3 phonology that I find particularly relevant for the current thesis, and that I believe can help us have a better picture of the phonological CLI. To put the reviewed findings in a broader context of theoretical L3A, the findings from L3 phonological acquisition seem to point to a dynamic interaction of *all* involved linguistic system (rather than copying of one of the systems) and be more in line with the FTP, rather than FT.

Traditionally, CLI was defined as a process through which features from the native language (L1) would manifest themselves in the L2. The steadily growing number of studies in L3 and multilingualism in general, has shed more light on the phenomenon revealing that CLI can be multi-directional for multilinguals and can be affected by a complex variety of factors (Wrembel et al. 2019, Hammarberg and Hammarberg 2005).

In a study conducted with two groups of young multilinguals aged 12-13, (L1 Polish, L2 English and L3 German vs L1 German, L2 English and L3 Polish) it was observed that a five-year-instructed training could result in a considerably accurate perception of L2 English rhotics (Wrembel et al. 2019). Both groups who had been learning their L3 for less than a year showed different perception accuracy for novel sounds in their L3, with the L1 German students outperforming the L1 Polish. In addition, it was also investigated and revealed that there was inconsistency in the acuteness of their L3 phonetic instructions the two groups had received. In fact, the L1 German students benefited from a more accurate instruction. Therefore, the most immediate conclusion to be drawn is that:

- 1) Though there may be a sensitive period for the acquisition of L2 phonology, compensatory strategies such as instructed training can revive possibilities to acquire non-native sounds at a high accuracy level.
- 2) The influence of **various factors** including the instruction itself as a phonological index for the students, the effect of proficiency (English) and the role their L1 / markedness plays a role. The study finds that L1 Germans benefited from having more marked phonetic feature specification in the native language.

The latter is also in line with Markedness Differential Hypothesis (Eckman 1977, Edwards & Zampini 2008) which argues that structures that are different and more marked in the new/additional language should impose more difficulty for the learners, whereas less marked features/structures should be easier to acquire. Thus, put together, this indicates the complexity and a multifactorial nature of the CLI that is even more visible in the L3/Ln acquisition than L2 acquisition.

In relation to production (vs perception), a seminal study on L2 phonological acquisition (Lee et al. 2020) has demonstrated that perception-based instructed training, compared to production based, results in more effective L2 pronunciation and speech learning. Although, this certainly

calls for further studies to examine the effect of different types of instruction on pronunciation, it highlights the significance of two primary results:

- 1) Reconfirming that the sensitive period is rather a general idea which describes human language acquisition abilities on average, without considering the possibilities of compensation in the post sensitive period.
- 2) A tight link between perception and production. If one becomes able to distinguish and perceive a new sound, she/he is more likely to reach a more accurate production of that sound.

It should be noted that although other factors such as attitude, working memory, general auditory aptitude, to name just a few, also affect the final outcome, the findings of the study above are generally still in agreement with the sensitive period idea of a transition from “citizens of the world” to “native language specialists”. In fact, this so-called “economy” can be reshaped and redefined, given new circumstances and provided with new input. In the next section, I will provide an example of how this attitude to CLI has had an essential influence on didactics and teaching methodologies.

2.2.2 Facilitative and non-facilitative CLI

Languages have certain Commonalities (100% overlap), Similarities (partial overlap) and Differences (CSD). These can be found in their lexicon, grammar, sound systems, etc. When it comes to learning a new language, especially as a late learner, any of these CSDs can come in as either facilitative or non-facilitative, having a significant impact on the way CLI can occur. Put into more accurate words, if the transfer of a property from a previously acquired language leads to facilitation in the new language acquisition it is called a positive transfer and on the other hand, it is a negative transfer if the results are the opposite, such as errors or difficulty in the language development (Bardovi-Harlig and Sprouse 2017).

I already reviewed examples where one might overgeneralize features of her/his L1 grammar in producing the L2 sentences. For example, the Norwegian V2 structure in English as an L2 or not following the V2 rule by an English L1 speaker learning Norwegian as an L2. However, in phonology, this is not as concrete as it may seem in grammar (Archibald 2018, 243). To decide whether a sentence is native-like grammatically, one might have an easier task compared to phonology, at least in cases where there is a clear grammatical rule that renders the alternative ungrammatical. As an example, an L1 Farsi speaker may produce sentences with SOV order in a formal/written context may come up with the sentences in (5) and/or (6).

- 4) He wrote letters every Sunday.
- 5) He every Sunday wrote letters.
u hær Yekshanbeh minevesht nemeh
- 6) He every Sunday letters wrote.
u hær Yekshanbeh nameh minevesht

This raises the question whether a potential positive transfer can be intentionally used to facilitate acquisition of a feature. For an example in phonology, a study by Trude and Tokowicz (2011) investigated the positive/negative transfer by studying Spanish-speaking vs non-Spanish-speaking individuals producing Portuguese. The assumption was that having a common background in a Romance language would lead to positive transfer of many features and thus a more accurate production in non-native Portuguese. The results indicated that the Spanish-speaking group had a better performance, as expected; however, they also made more errors on the pronunciation of cognates and non-cognates (*True and false friends*), perhaps on account of overgeneralising what they found facilitative. Although, the study suggests that phonological performance is likely to be affected by a number of various predictors, the core observation that transfer can be positive as well as negative still applies.

2.2.3 What are the criteria for an acceptable production?

The criteria to measure or evaluate one's phonological production have experienced an evolution, going from a strict demand to be 'nativelike' to being 'comprehensible/intelligible'. I must also highlight that I am not using the last two terms interchangeably, though it might seem somewhat challenging to differentiate between the two (for a review see Munro & Derwing 1995, 1999 and Derwing and Munro 2008). Nevertheless, today there seems to be a general agreement that comprehensibility should be prioritised over nativelikeness, regardless of what techniques and approaches might be applied (Archibald 2018). However, to whom a speech is comprehensible can differ. If varieties of English such as Jamaican or Chinglish are comprehensible to communities familiar with those varieties, does it also mean that they should be comprehensible to other groups? Or are the criteria for comprehensibility, more often than not, being comprehensible to the native speakers of that language? Is being fully comprehensible enough for certain L2/L3 speakers not to be treated differently and or unfairly?

In fact, the social outcomes of a non-target-like accent may not be easily alleviated by pure linguistic data and facts. Despite all the linguistic findings, a foreign language teacher or instructor who is a native speaker of the taught language is still more desirable than a non-native speaker in many cases (Cook 2000, Jenkins 2000). According to another study (Derwing

2003) most of the participants approved that they would be treated with more status and manner by the native speakers if they did not sound accented. However, this is not only limited to marginalisation and being rejected on the basis of a non-standard accent in the target language society (Sato 1991, Podberesky et al. 1990), but also expands to the L2 speakers in their L1 context. According to a study by Sung (2016) on university students in Hong Kong, the majority of the participants announced that they found their local accent (vs a *native-like* accent) embarrassing and associated it with low proficiency and on the contrary, associated a native-like accent with prestige, power and high proficiency. Put together, although this may seem unfair and unrealistic, many learners would like to overcome foreign accentedness and try to achieve a native-like accent.

English has many varieties with different pronunciations. When it comes to ESL (or even EFL), this might be somewhat of a challenge for both teachers and learners. What type of pronunciation should be taken as the target one by the teacher? What is the desired pronunciation goal? To what extent is any of the assigned goals and strategies realistic and practical? And to what extent can/should the learners reach them? And finally, how do we put to practical use the findings of the studies on phonological acquisition to help the learners improve their pronunciation? Perhaps, the emergence of the term *Englishes* (or world Englishes) and consequently the Lingua Franca Core (LFC) (Jenkins 2000) can be described as efforts to come up with a response to these questions. In summary, LFC proposes a general paradigm that suggests more comprehensibility on a global scale, by proving a list of *dos and don'ts* with “preferable” and “acceptable” in the middle of the spectrum (see Dauer 2005 for a review).

LFC has been debated for its assumptions and practicality. Trudgill (2005, 76) argues, for instance, that LFC is not comprehensive, as the paradigm itself claims. He argues that LFC mostly targets the communication between non-native speakers of English whereas many learners want to communicate with the native English speakers and consequently need to understand them, which most probably requires standard English pronunciation instruction. LFC may not be the optimum or productive solution, however, this, I believe, confirms that the early and general idea of comprehensibility was not enough and not quite objective to work with. In addition, even the LFC itself is comprised of rules some of which begin with an imperative *No* at the beginning. In other words, one needs to attend to certain guidelines to sound just comprehensible. Therefore, whether LFC can “drastically reduce pronunciation teaching load” (Jenkins 2000, 145) or not, coupled with arguments put forward by other

scholars against it, one thing is agreed upon: pronunciation teaching is a delicate and uncertain domain (compared to grammar). However, this uncertainty and notions of comprehensibility and intelligibility may fluctuate when meeting outcomes from either a linguistic or a sociolinguistic point of view. Perhaps, mispronouncing the word “sheet” with an /ɪ/ instead of /i:/ could still be intelligible, given a certain context, however, this might not be easily applied to its social consequences. This uncertainty is exaggerated by the encounter of the sociolinguistic outcomes, and may simply lead to personal preferences and strategies in relation to pronunciation and accent in learners, which may appear to be unrealistic and extreme at times (cf. the example of the university students in Sung 2016 study).

My emphasis on the latter section is that we do not still have a complete list of objective principles to follow when it comes to pronunciation teaching. There are different individual needs and preferences, individual abilities and linguistic backgrounds and a variety of social contexts all of which can have an impact on the decision making and objective in relation to teaching. As Trudgill (2005, 79) points out, we might still be able to rely on the old and traditional pronunciation teaching according to a “native speaker model”, to allow individuals make their own modifications based on their needs and abilities. A native speaker model already includes a set of dos and don’ts that the LFC proposes too, which might minimise the need for LFC or similar approaches.

Furthermore, pronunciation teaching should rely on the phonological features of the taught language, English in the case of our study, and should be aware of the possibility of positive and negative transfer from the previously learnt languages. Consequently, if any of the areas of phonology can be expected to be prone to non-facilitative influence from the previously acquired systems, they can be remedied by putting more emphasis on these problematic aspects, while areas where facilitative CLI can be obtained could be explored to benefit the development of the target language.

A classic example of such an approach (based on similarities and differences between known languages and the new language) can be found in the work of Swan and Smith, the *Learner English* (2001), where the authors provide English teachers with a comprehensive guide for teaching students with a range of first languages. In their book, every language is separately compared to English and the areas where the L2 English learners from that language can face difficulties are highlighted and explained in detail. The book reviews similarities and differences at many levels. For example, it provides teachers with the possible false friends a

Spanish student is likely to use in his or her English production. Or as another example, in phonology particularly, it classifies many instances where negative phonological transfer can occur from a given language (e.g., Spanish) with a set of examples. To sum up, works like Swan and Smith (2001) illustrate an approach to foreign language teaching where instances of potential positive vs negative transfer can be fruitfully explored and inform the curriculum. It also confirms the point of view that having a standard framework in teaching pronunciation and more generally – phonology - is highly in demand by many teaching training centres. Knowing the linguistic background of students, especially in an international setting, is highly important as it can beneficially inform and help the teacher structure the focus of explicit instruction. In addition, it highlights the role of instructed training which itself can be associated with intentionality of the learners.

It is worth noting however, that works that focus on several language learning contexts like *Learner English*, and similar works, cannot encompass all the language varieties. Typically, such books choose to focus on ‘big’ languages like Spanish, Chinese, Arabic etc, and leave smaller languages beyond the scope of the book. As a result, there can be varieties which may never appear in such books.

Moreover, it may often be challenging to draw a clear line between some closely related varieties and decide whether their speakers should be regarded as bilinguals or monolinguals. For instance, a study by Lundquist and Vangsnes (2018) on the grammatical gender system focusses on the differences between two Norwegian dialects, the Oslo and Sogn, shows that there is only one group of Sogn speakers act like bilinguals when predicting the gender of the upcoming word by making use of the gender markers. What is interesting is that in cases where two or more closely related varieties are in an interplay, like Norwegian varieties in Oslo and Sogn, one group of speakers might act like bilinguals and clearly differentiate between the two varieties by having two distinct systems, while the other group may have only a passive knowledge of one of the varieties, and uniformly use only one system in processing. This is critical since it helps us distance ourselves from a black and white definition of bi/multilingualism but rather move towards a smaller scale comparison. This comparison will be crucial for us when reviewing the multilingual situation of the current study: two closely-related varieties (Farsi and Gilaki), which are similar in many respects, but still differ in certain phonological features relevant for the current study.

This brings me to the next section where I am going to review the example of Farsi and Gilaki (both Iranian varieties) in comparison to English.

2.3 Farsi, Gilaki and English

2.3.1 Farsi and English

Coming from the same linguistic background, e.g. Romance languages, can be both facilitative and non-facilitative as discussed in Trude and Tokowicz (2011). The “same family background”, however, might be too broad and not so straightforward, referring to languages that might be close or distant relatives (compare for example Romanian and Spanish vs Spanish and Portuguese). In addition, throughout history, languages from different roots have had encounters as a result of which one or both languages have gained features that used to be absent in their background, for example the encounter of Arabic (Semitic) and Farsi (Indo-European). Despite that, the same or a close family can indicate potentials for certain CLIs or transfer, as I reviewed one example in the study by Trude and Tokowicz (2011). It can lead us to make assumptions and investigate the possible negative and positive transfer that come in the way.

Farsi (or Persian which are often used interchangeably) and English come from the same large family of Indo-European languages. Coming from the same language family most probably means sharing certain commonalities, though with varying degrees, in different linguistic domains. In the case of Farsi and English, one of these similarities are overlaps in the phonological systems of the two languages (Paraskiewicz 2015). Almost every sound in English has a similar (if not identical) counterpart in Farsi. The exceptions are the few unique consonants in English that Farsi doesn't have, such as /θ/ and its voiced version /ð/. Although not every phoneme might be identical to its counterpart, they sound close enough not to cause any major difficulty for a Farsi speaker learning English. Easy and quick phonetic integration may be the reason why English loanwords are easily and at a fast pace adapted by Farsi speakers (for a review see Paraskiewicz 2015). The tables below (table 2 and table 3), demonstrate the similarities in the consonantal systems between the two languages.

	Bilabial	Labiodent.	Dental	Postalv.	Palatal	Velar	Glottal
Plosive	p b		t d			k g	ʔ
Nasal	m		n				
Fricative		f v	s z	ʃ ʒ		x	ɣ h
Affricate				tʃ dʒ			
Trill			r				
Approximant					j		
Lateral Approximant			l				

Table 2 - Consonants in Farsi. IPA

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal	Examples
Plosive	p b			t d			k g		<u>pin</u> <u>tin</u> <u>kin</u> <u>bust</u> <u>dust</u> <u>gust</u>
Affricate					tʃ dʒ				<u>cheap</u> <u>jolly</u>
Nasal	m			n			ŋ		<u>seem</u> <u>scene</u> <u>sing</u>
Fricative		f v	θ ð	s z	ʃ ʒ			h	<u>fin</u> <u>thin</u> <u>sin</u> <u>shin</u> <u>hit</u> <u>van</u> <u>the</u> <u>zoo</u> <u>measure</u>
Approximant				r		j	w		<u>rate</u> <u>yell</u> <u>well</u>
Lateral Approximant				l					<u>late</u>

Table 3 - Consonants in English. IPA

As evident from the charts above, most of the consonants are shared between the two languages¹.

¹ It should be noted that the IPA chart for Farsi demonstrates the sounds of the standard variety whereas other local varieties are known to have some other sounds some of which might be common with English, and other languages. For example, the glide /w/ is assimilated into /v/ in standard Farsi:

- a) Windows /vinɪn.dəʊ z/ for /wɪn.dəʊz/

On the other hand, the differences are substantially more distinct in the vowel systems of the two languages (figure 4). There are 5 vowels in English (marked with red circles) that do not exist in Farsi. As a result, the total number of vowels in Farsi equals 6 whereas it is 11 in English.

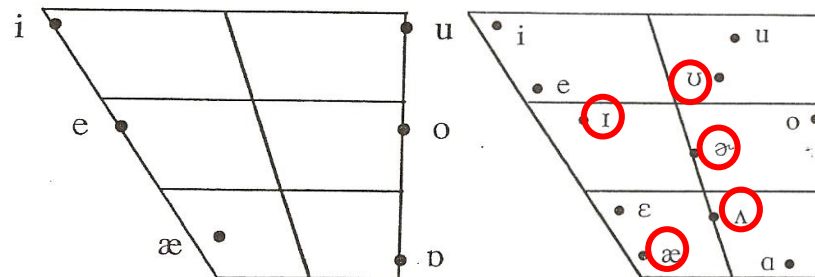


Figure 4 - Vowels in Farsi (left) and in English (right). IPA

This is rather important for ESL or EFL since most of the difficulties Iranian students experience learning English in terms of pronunciation has been reported to be attributed to the differentiation between the minimal pairs such as *ship* and *sheep* assimilating the vowels in both cases to the Farsi /ɪ:/ (Seddighi 2012).

This brings us to the main focus of this study, the acquisition of English vowels that are absent in standard Farsi, but which exist in one of the varieties spoken in northern Iran called Gilaki. This observation was in fact the starting point for the current thesis. Before I proceed with the argument why this difference may be an important resource for Gilaki speakers acquiring English, I will give a brief overview of Gilaki from a linguistic and a sociolinguistic perspective.

2.3.2 Gilaki

2.3.2.1 The general background and phonological features

Gilaki, one of the north-west Iranian languages, is the language spoken/associated with the province of Guilan, located in northern Iran, the southern part of the Caspian Sea. The estimated

However, in some of the varieties in southern Iran, or in Kurdish (Kurdi in Farsi) the /w/ exists in its glide form:

- b) /wolek/ informal exclamation to call someone
- c) /wafr/ or /wafer/ for /bærf/ meaning snow.

population of the province is 2,531 million with its capital city Rasht around 679, 995 residents². Gilaki vocabulary has been under strong influence of Standard Farsi. According to Rastorgueva & Lockwood (2012), the phonology of the Gilaki has been only “partially” affected by Farsi. To my interpretation however, Gilaki phonology in its current state may have become more affected and changed compared to the description in Rastorgueva & Lockwood (2012). The Gilaki language has mostly existed as an unwritten variety. Thus, most of the few written texts resembled the standard Farsi and consequently the reader had to rely on their own knowledge of “how-to-read properly” in Gilaki. In fact, this has partly been due to the differences that lie between the vowels of the two languages (figure 5).

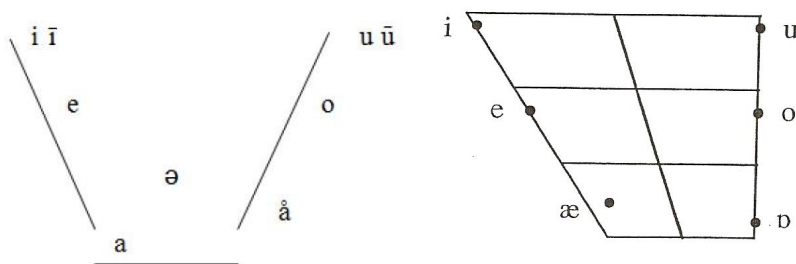


Figure 5 - Gilaki vowels (left) vs standard Farsi vowels (right)

The Gilaki vowels chart demonstrated above, is almost the only one available and it is not according to the IPA format and goes back to the only comprehensive study of the variety (Rastorgueva & Lockwood 2012)³ on Gilaki. Although there are differences in the representations, it can immediately be seen that the number of vowels in Gilaki is more than the ones of Farsi with 9 vs 6. Rastorgueva and Lockwood report that there are only duration differences within /u/ and /i/ pairs (figure 5 – left and right top corners). Yet there might be slight differences in the way other studies represent this sound system. It should also be noted that the representation of this chart is mainly according to the variety of the dialect in Rasht, the capital of Guilan, whereas Gilaki has 3 main varieties, Bie-Pas, Bie-Pish and Galeshi which indicates a possibility of a variation in phonetic repertoires. Moreover, it should also be noted that the original study was conducted about 50 years ago and in the meantime, there has been a

² Most surveys refer to the last data from 2006. Today, new estimations are about an average of 1 million inhabitants living in Rasht. The day population is also estimated to be roughly higher.

³ This is a translation of the original work in Russian entitled: *Giljanskij jazyk* by V. S. Rastorgueva, A. A. Kerimova, A. K. Mamedzade, L. A. Pireiko, and D. I. Edel'man. Editor in Chief: V. S. Rastorgueva, Doctor of Philological Sciences, Moskva: Izdatel'stvo "Nauka". 1971.

new wave of migration from other parts of the Guilan province into Rasht, its capital, which could have resulted in contact between several closely-related dialects of Gilaki. My personal interpretation, as a Gilaki native speaker, is that even according to the variety from Rasht, the two above-mentioned pairs have spectral differences and the chart can still include at least 2 more vowels (/ʌ/ and /ɐ/), however this needs thorough investigation and analysis to be scientifically reliable. Nevertheless, the first conclusion is that the Gilaki vowel system, which phonologically distinguishes between several lax and tense vowels, is closer to the English system than the Farsi is (figure 6).

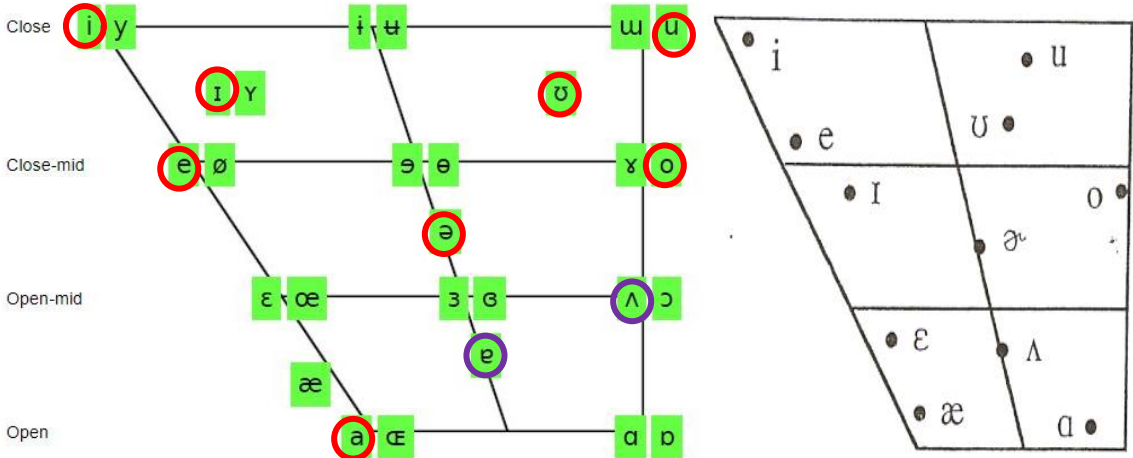


Figure 6 - Gilaki vowels (left) according to the IPA chart vs English vowels (right).

Earlier, I mentioned how the standard Farsi orthography did not suffice for Gilaki to become and exist as a written language on account of the vowel differences between the two varieties. However, a rather comprehensive orthography was finally introduced as a result of a collective effort of about 30 scholars in late 2020. The newly introduced orthography claims to have the essential characteristics and characters needed to convey the pronunciation of the Gilaki vowels. However, among the three main Gilaki varieties, the new orthography seems to be closer to only one of them and it might become a reason for the other groups not to fully identify with it. But in general, this seems to be a natural outcome of a movement that was founded about a decade ago to revive the Gilaki language. Although many of the efforts were individual and experimental, such as translations into and performing famous plays in Gilaki, or the creation of original Gilaki lyrics and poetry, finally led to the first ever written standard for this variety. This, though, is only a beginning and requires certain amount of time and observation to see if the new system succeeds and is recognized by the Gilaki speakers as the written standard.

2.3.2.2 The sociolinguistic situation of Gilaki

Being a spoken language in general, might be one reason for many local varieties not to be able to sustain or have less to offer compared to a standardised or codified variety. According to Weth and Juffermans (2018, 6) “writing can influence processes of language change” along with being a criterion for setting standards as “correct” and “incorrect”, and norms such as “literate” and illiterate”. Therefore, I assume that the absence of a standard writing system can have certain influences on the language change. In the example of Gilaki, I will be shortly reviewing how the absence of a codified standard writing system, coupled with other factors, has led to a certain language change. The overall situation of Gilaki is not an exception from most minority languages.

Most Gilaks, (people of Guilan), see their variety as one inferior and non-prestigious variety of the (more superior) standard Farsi. Also, they mostly do not consider the “standard language” (or even the variety spoken in Tehran) a dialect. This is not unique to the Gilaki-Farsi and other Standardised varieties might have the same relation with the non-standardised varieties spoken in a country. (Wardhaugh 2010, 28). The lower-class self-image of local varieties can be observed in many language-dialect situations. The very notion of standardisation itself has certainly resulted in the imbalance of the social outlook of this relation. Wardhaugh (2010) points out that the standardisation of a language can be driven by various social, cultural, educational and political dimensions, rather than linguistic ones. As a consequence, the other varieties are demoted in a sociolinguistic hierarchy, which might lead to their subsequent lower status (social, cultural etc.).

On the other hand, the users of the codified and standardised variety begin to see themselves as having a higher status, prestige and/or power. The amalgamation of the dimensions mentioned above with the standardisation process naturally creates a gateway, or a bridge, to the class that is now associated with all the power and prestige: education. During the early phase of standardisation in Iran, not so many people were aware of this process taking shape. For example, during the modernisation period in Iran by Reza Shah Pahlavi (1921-1941), the first language standardisation procedures were implemented which were also fuelled by a puristic attitude and a sense of nationalism (Paul 2010, Perry 1985). This was when the first comprehensive and systematic education schemes were initiated. The training of teachers and then educating the new generation of students, indeed, required a standard language. That is where diglossia started looking like a problem (Wardhaugh 2010, 89-92). In other words, those educated in the 1970’s had to become bilingual, speaking standard Farsi in addition to their

local variety. Therefore, the language of the educated vs non-educated had a more tangible manifestation, which resulted to a public conclusion:

- the (more) educated the more Farsi-like and the more local-like, the less educated (figure 7).

This can also explain the social burden shouldered by the public who felt less powerful, lower in class, non-prestigious. This generalisation may not be applicable to all local varieties to the same extent; however, Gilaki has been a proper example of such situation.

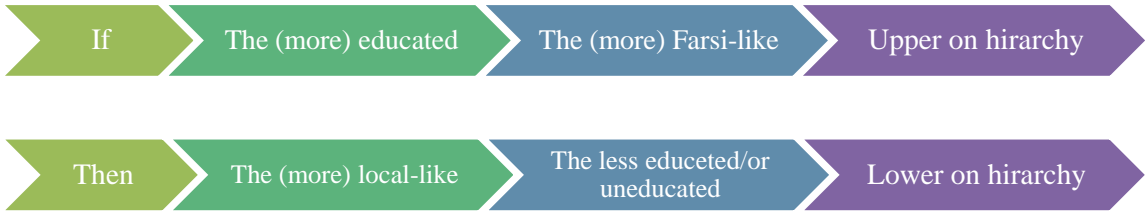


Figure 7 - Inverse conclusion making based on standard-local relation.

What Gilaki has experienced is not only a common lower status that many local varieties are associated with, but in fact, has been there another driving force that both amplified its low-class associations and altered its function in society. As mentioned earlier, Gilaki has not had a written form in general and the standardisation of Farsi coupled with the first systematic nationwide education system marginalised the use and the status of Gilaki. As a part of a familiar scenario, families stopped talking to their children in Gilaki. This led to a generation with a passive knowledge who heard and understood the conversations of their parents’ generation in Gilaki but never needed to use the language either at home or at school.

The one-way conversations of the older generation were the only input until 1998 when the local television was established. It then became the second major language input after real human conversations; however, it gradually started to becoming more specialized (and compartmentalized) in terms of the issues and the content it covered. The most important feature of the TV channel was the use of the local language in different programmes some of which people found difficult to connect with, such as the news in Gilaki since it sounded quite unauthentic and unnatural to people. I believe the example of the Gilaki news was (has been?) a case of a diglossia with the low (L) variety in a context associated with the high (H) variety. This is not certainly an exception to Gilaki and L varieties either fail to take on such functions or require persistence and longevity so that new association with the L variety are formed (Wardhaugh 2010, 89-92).

Above all else, one of the most (if not the most) popular shows on that TV channel were the soap operas whose casts talked Gilaki. However, most of these TV series had one theme in common: low comedy. Therefore, the generation who only had passive yet authentic knowledge began to receive an exaggerated load of language input almost only through joking and jesting of unusual TV characters. Soon, the catchy phrases were used by many young people, and it became their first and only Gilaki utterances.

Since the standardisation, Gilaki has had a trend which can be characterized as both a decline and an increase. The decline has been in the number of its natural users after the standardisation and recently a moderate increase in its users plus an evolution in its function, in a way that the younger generations who know the language (to any extent) use it for a restricted set of purposes (figure 8) rather than communication on a daily basis.



Figure 8 - Post standardisation changes of Gilaki

It is worth mentioning that the new or exaggerated function is still an L variety which has also become more limited in terms of function and its users (figure 9).



Figure 9 - Gilaki, still an L variety but with an altered/ing function.

2.3.3 Code-switching between the L and H variety

Code-switching is a tactic/strategy bi/multilinguals use in their communication to bridge the gap, to create solidarity, express identity, etc. The kind of code switching that happens between a Gilak (a person from Guilan) and a speaker of the standard Farsi mostly manifests itself in the phonological aspect. Like many L-H varieties, the switching is upwards on the hierarchy with the Gilaks using the Farsi phonological features. It is also true for two Gilaks if they intend to express a prestigious and powerful position to each other (or at least to avoid being classified as the opposite). How can this affect a young Gilak learning an L2 or L3? Put in more precise words, what shares of transfer does Gilaki and Farsi will have, knowing that one is regarded as an L and the other as an H variety? My prediction is that the formal nature of an instructed training (English or any other language), exposure to the H variety of the new language, avoidance of low-class associations by the learners may lead to a rather subconscious superiority of the phonological transfer of the native H variety (figure 10).

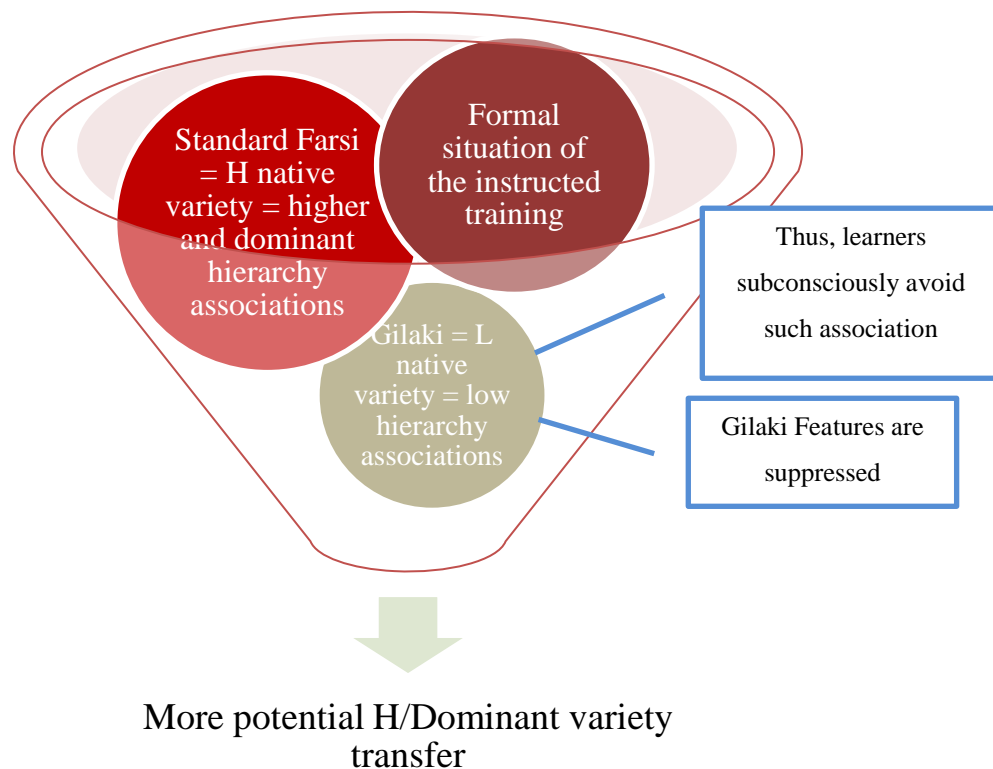


Figure 10 – The effect of school (education) context on the activation of H and L varieties.

Figure 11 also schematically represents how Farsi can function as a phonological filter which results in the suppression of the Gilaki vowels, and a dominance of Farsi-based pronunciation.

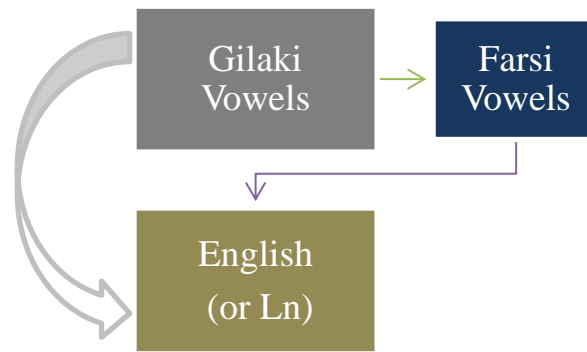


Figure 11 - Phonological dominance of Farsi as the H variety.

I have tried to put emphasis on the fact that the Farsi vowel repertoire is smaller compared to the Gilaki in the figure above (11) by manipulating the size of the box, nevertheless the position it has on the hierarchy provides it with more power so that it can be more active than a larger vowel system. On the other hand, Gilaki vowels can be facilitative in the acquisition of the English system yet the suppression form the H variety may be able to inhibit the positive transfer from Gilaki to varying degrees. By the end of this study, we might be able to shed light on the question if Gilaki can prove to be facilitative in the acquisition of English.

2.4 Summary

Many studies in the rapidly developing field of L3 acquisition have proposed accounts to the effect how CLI should be defined and understood. L3 acquisition and CLI in L3 are now regarded as more complex phenomena compared to the view on CLI in the more traditional field of generative SLA of the 1990-ies-early 2000-s. Recent findings from the L3/Ln acquisition are in my view more in agreement with the LPM/FTP theory which argues that that CLI is a complex and dynamic process driven by co-activation of *all* available linguistic systems in a multilingual mind, rather than with the FT ‘wholesale copying’ proposal. The fine-grained property-by-property nature of the transfer may explain why some features may be more prominent at one time but disappear or minimise at another time. In addition, not being “whole” and “at once” is indeed in agreement with the changing direction of the CLI/transfer, since traces of L1, L2, L3 and Ln features can be found in any of the languages of the learners’ repertoire. This would have been hard to account for within the FT or Wholesale Transfer.

In the sections above I reviewed a few recent theories that looked at the questions of teaching, focussing on nativelikeness comprehensibility/intelligibility, Lingua Franca Core, etc. I highlighted that it is quite essential to consider these arguments since they have a direct impact on the way phonology and pronunciation should be taught. For instance, they could clarify to

what extent we need to detect and overcome negative transfer or encourage facilitative transfer. All of these could be affected by the criteria we assign for the “acceptable” pronunciation. Subsequently, I tried to indicate that the agreement on “comprehensibility” might not be as objective and straightforward as it appears to be. In addition, it is also important to know the strategies that have been offered regarding identifying facilitative and non-facilitative CLI in a given learner population and the target language and using this knowledge in classroom to the learners’ benefit. I concluded that regarding the usability of CLI, it sounds justifiable of the teaching methodologies/ists to attend to potential facilitators and non-facilitators in a given language combination and provide learners with a broader set of options and possibilities to choose from.

Furthermore, FTP also provides us with the potential to include extra-linguistic features such as learners’ attitudes towards certain features of a given language. One outstanding characteristic of L3/Ln, explained by FTP and confirmed by a growing number of studies, is that CLI, or transfer in general, is multifactorial. With all of that said, a pure linguistic investigation may not always be able to explain why and how any of these features can or cannot be transferred. Languages have different social and socio-political associations attached to them which do not manifest themselves in the way a learner might intentionally or unintentionally suppress or encourage CLI. For instance, an L3 Russian learner might not be willing to sound quite native-like (Archibald, 2018, 251) if she or he is a politician in the US. In most L3 studies, such extra-linguistic factors are either not clearly defined or taken into account (for instance whether learners have different attitudes towards certain sounds from certain languages etc.). There is also a scarcity of studies around L3 acquisition that involve speakers of L varieties in contexts associated with the formal status (Wrembel, Marecka & Kopečková 2019). Moreover, it is not clear if learners associate the same position and power for their additional language(s), especially if those languages are close enough in such a way that one might be considered a “dialect” of the other. Just like the example of Gilaki which offers a richer phonological repertoire to the learner but fails to have the position and power of the Standard Farsi that is less marked or in other words, has a smaller set of vowels.

Finally, as of today, there has been no study on whether Gilaki users (passive and active) act like bilinguals or not. Therefore, the results of his study may be able to make a small contribution to this area.

All this brings me to the end of this section which was intended to provide a rationale and a motivation for the current study. With this in mind, I will move on to the introduction of the methodology in the next chapter. In what follows, I will show how every core point discussed so far has motivated a particular feature in the experimental design and subsequently in the data analysis and the discussion. Generally, my core assumption was that since Gilaki has a more elaborated vowel repertoire compared to Farsi, distinguishing between long and short (lax and tense) vowels, Gilaki speakers might find it easier to acquire and produce English vowels (figure12). And if that is true, we would be able to observe facilitative CLI in L3 phonology, namely facilitation in pronunciation.

Figure 13, as a summary, demonstrates how FTP/LPM, and consequently intentionality, enables us to picture a multidirectional path for CLI in which it is possible to selectively reinforce a property from either L1 or L2 to reach a desirable outcome in L3. In addition, it is under the FTP that we can substitute intentionality and instructed training with other factors such as social status, universal preference, age and even socio-political status of a language and expect a different result. However, there is also a possibility for L3 learners to “intentionally”, inhibit the transfer of a certain features from either L1 or L2 the result of which would differ from case to case, and even person to person, in the case the languages involved.

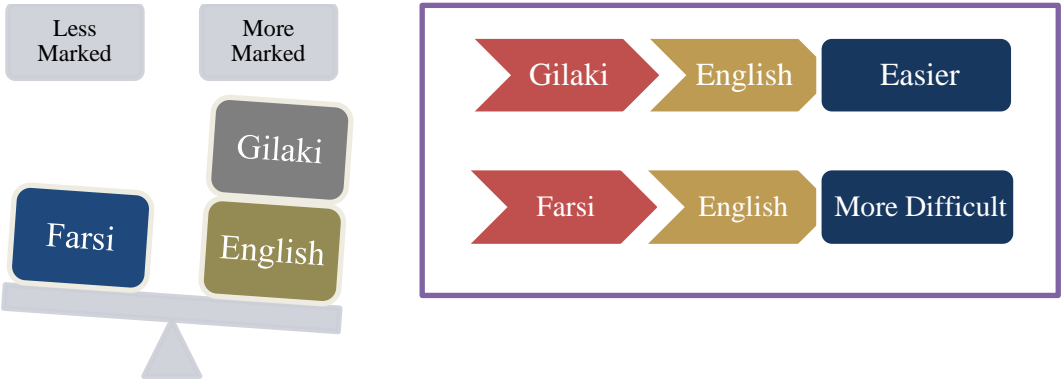


Figure 12 - More markedness of Gilaki should facilitate the English pronunciation for its speakers.

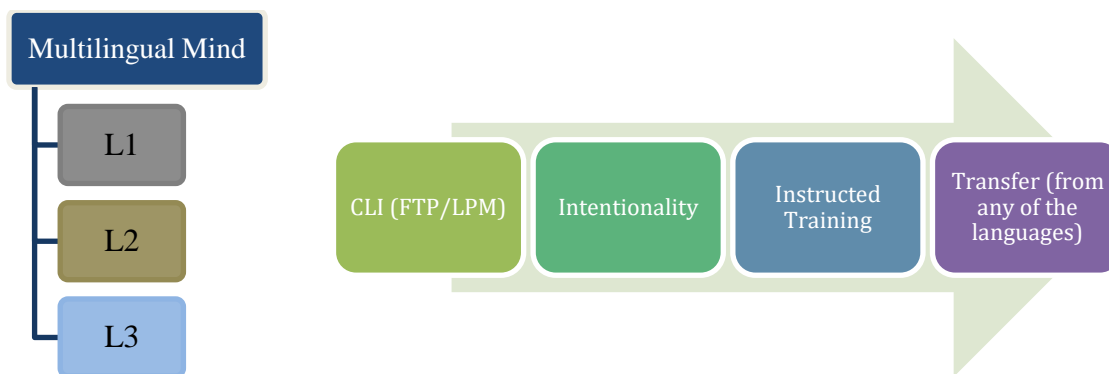


Figure 13 - How positive transfer could be intentionally reinforced according to LPM/FTP.

3 Methodology

3.1 The design of the experiment

The experiment was comprised of two tests, 1) a production task and 2) a background questionnaire plus an attitude test. In this section, I begin with the production task (figure 14).

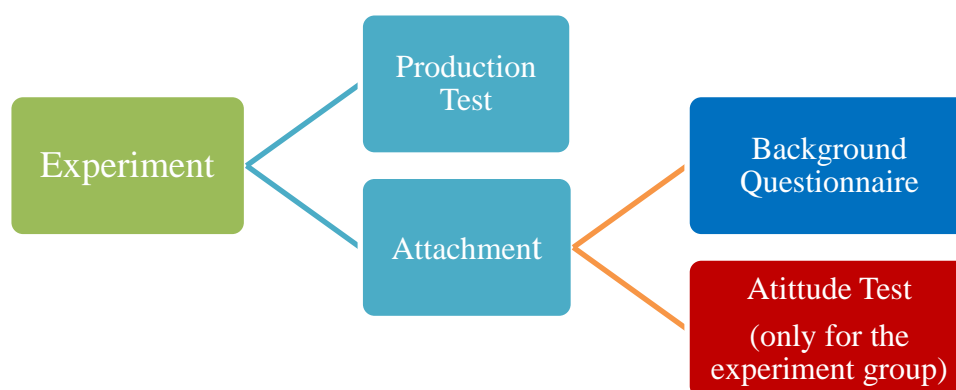


Figure 14 - The general layout of the experiment

3.2 The production task

The design of the experiment was originally initiated from the observation that was made on the phonological differences between the three languages of this study. Earlier, it was discussed that Gilaki has more phonological commonalities with English than with Farsi when it comes to the vowel system. In fact, Gilaki and English share a similar vowel repertoire which is also larger than the one of Farsi. In addition, both Gilaki and English are more marked compared to Farsi. Therefore, the assumption was made that a Gilak, whether a passive or an active user, might benefit from the positive transfer from Gilaki in producing English words/sentences. In this study, I assumed that the similarities between Gilaki and English vowels are the /ə/, /ʌ/, /ʊ/

and /ɪ/ which are assimilated to sounds close to them in Farsi. Therefore, both as a filler and a criterion for comparison the vowels /æ/, /ɪ/, /ɑ/, /ɑ:/ and /ö/ were included. In addition, one extra group including diphthongs such as /o-ɔ-əʊ-ou/ was added to the rest of the sentences (table 4).

1. /ə/	2. /æ/
3. /ɪ/	4. /ɪ:/
5. /ʌ/	6. /ɑ - ɑ:/
7. /ʊ/	8. /ö/
9. o-ɔ-əʊ-ou	

Table 4 - Target vowel sounds (left and highlighted) and fillers

- 1) For every vowel, six sentences were made with only one target vowel embedded in the target word (same for fillers).
Example:
 - /ə/
 - a) I hope she will **accept** the invitation.
 - b) Please double check your bank **account** number.
- 2) The sentences were designed to be structurally simple. To achieve this, most sentences were short (min 3 and max 12 words), no subordinate clauses or complex structures were included.
- 3) The sentences were designed not to contain hard to read or advanced vocabulary, especially the target words.
Example of target words: *such, sugar, live, stupid, teens, push and bank*.
- 4) Sentences were organised based on the random order to prevent any possibility of guessing or prediction by the participants.
- 5) The model for sentences were double-checked with examples from the Cambridge Advanced Learner's Dictionary – 3rd Edition.
- 6) Since the position of a word in a sentence and its pre/proceeding words can affect the stress patterns, the embedded target words were placed at different positions of the sentences. For example:

1. He **did** it
2. **Some** people find it hard to study in the morning
3. It's so **clean** here
4. The people at the seminar were all wearing the same **badge**
5. I guess we need a stronger **glue**

3.3 Controlling for external factors

Prior to anything, participants were given a guideline when files were sent to them asking them to open the second file (attachment) only after they have completed the first task (the experiment). This was especially critical for the experimental group since they had an additional attitude test in the attachment which could potentially influence the outcome of their production task. To conduct the test and obtain a more ecologically valid result required the participants to read the sentences spontaneously with the least possible preparation time, least distraction by the number of sentences, and possibly most minimal stress effect. Therefore:

- 1) The experiment was designed to be taken offline via a pdf file and a mobile phone recorder for the following reasons:
 - a) Firstly, avoiding any probable internet filtering in Iran, making sure that the files are accessible to every participant.
 - b) Secondly, an offline test provided the participants with their own preferred method of filling in the form. For instance, if one did not feel comfortable with filling in the PDF file they could simply choose the pen-and-paper option and send the photo instead of the filled in file.
- 2) It started with an instruction section on the first page both in English and Farsi asking the participants to:
 - a) Turn their recorders on.
 - b) Read the sentences as they turn the pages (in a PDF document).
 - c) Read each sentence **once only**.
 - d) Turn the page after reading each sentence. (Do not go back)
 - e) Continue until the last sentence in the same way.

The instruction did not provide any information on the number of sentences since the number “54” might have caused reluctance to do the test or anxiety.

- 3) For minimising distraction, every single sentence was located in the middle of a blank page. Therefore, the participants would have no idea of how long or short etc. the next sentence would be.

3.4 The background questionnaire (BQ)

The production task was accompanied by a background questionnaire (Appendix A) that included the following questions about the participants

1. age
2. birthplace
3. place of residence
4. Native language (s)
5. Foreign languages you can understand / speak
6. Age they started learning Farsi
7. Age they started learning English
8. English self-evaluation scale with 1, 2 and 3 standing for the beginner, intermediate and advanced respectively (or the level they study at):
 - a. -Speaking (beginner, intermediate, advanced)
 - b. -Reading (beginner, intermediate, advanced)
 - c. -Understanding (beginner, intermediate, advanced)
9. Educational background (high school, Bachelor, MA)
10. Current occupation

What language they use:

11. In family
12. At work
13. With friends
14. In studies
15. Social media
16. Leisure time (books, movies, podcasts ...).

The BQ was identical for both groups, however, for the experiment group it could indirectly reveal if any of them either mentioned Gilaki as one of their languages or used it in one of the contexts mentioned in the BQ. This could then be used and compared with the data from the attitude test which I will explain in the next section. Knowing from the literature that CLI in L3/Ln may be multifactorial, it was necessary to obtain some general information about

participants to see if there are any outliers and to potentially see the distribution of the participants based on the language background factors.

3.5 The attitude test (TA)

The AT was designed and included for two major reasons.

1. For the scarcity of data around Gilaki, I needed to obtain an authentic and up-to-date observation about the status of Gilaki that I mentioned in the literature review such as its current functioning.
2. As discussed in the state of the art section, I argued that the speakers' attitudes may have the potential to influence the activation level of the language and affect the source of transfer.

The AT in this study was a slightly revised version of a survey I had conducted a few months prior to this study. Regarding the first attempt as a pilot study which worked so well, I included the revised version in this study. Therefore, it can also be interesting to look at the two tasks and compare the results. In addition, the number of people (116) who participated in the first survey was substantially larger both in general and in comparison to the one of this study (30) (figure 15).

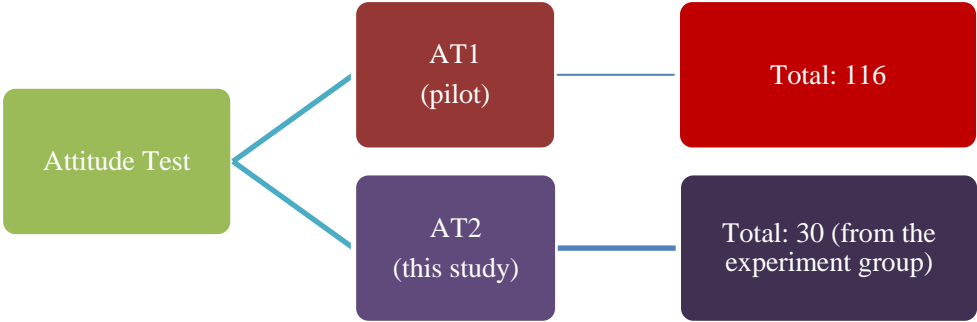


Figure 15 - The two attitude tests and their total number of participants.

AT1 survey was also an interesting and practical experience which can certainly become a valuable tool for similar studies in the future. AT1 was conducted on Instagram, through the feature provided by the Instagram *story* (IS)⁴. One of the options available on IS is the poll

⁴ In general, IS has the potential to substantially increase the access to participants if that account (Instagram page) already has a considerable number of people. For instance, I ran this

where you can raise a question and a yes/no (or any binary response) option for the audience. What makes this feature unique can be summarised as follows:

1. Instagram appears among the 10 top apps on most surveys making it one of the most familiar platforms available, especially to youth.
2. Not only familiar, it is one of the most user-friendly platforms (both for being familiar and the simple design of the IS “poll”)
3. The immediate contact of the one’s audience (followers) with the poll makes it more authentic in many respects, especially for responses that require to be given in the most spontaneous and immediate way, avoiding any preparation or thinking. In other words, they only know about the “poll” or survey, once they encounter with it.

AT2, a slightly revised version of the AT1 was sent directly sent to participants. Having had the experience from the AT1, I posted an announcement on Instagram story asking people to send me their emails if they were willing to participate in such a study. There was a substantial decline in the number of people who expressed their willingness to take part in the study from the original pool of 116 participants of AT1. Subsequently, there was a second fall in the number of people who did send their emails but never participated regardless of the follow-up afterwards (figure 16). I must also indicate that no detail of the study was revealed in the

experiment on my personal Instagram page with 520 followers at the time and an average number of 116 people participated in that. Although it is roughly a 25% percent participation, 116 is almost 4 times bigger than 30, the number of people I have managed to contact by email and networking. There is also another interesting fact about IS which I will clarify in the next paragraph. Nevertheless, I must indicate that, this is only a rather new experience which is young and needs to be studied and investigated for the possibilities it can or cannot offer. For instance, one might have less control over the population taking part in the poll. In the present experiment, some participants refused to answer some of the questions and that is why there is a min and max number of participants in figure 15. There can be other examples such as age, educational background ,etc. which one cannot filter or have control on. However, the large number of participants may compensate for these downsides, depending on the nature of the study and the poll.

announcement. Therefore, it may again demonstrate the potential Instagram poll has to be used in studies as such.

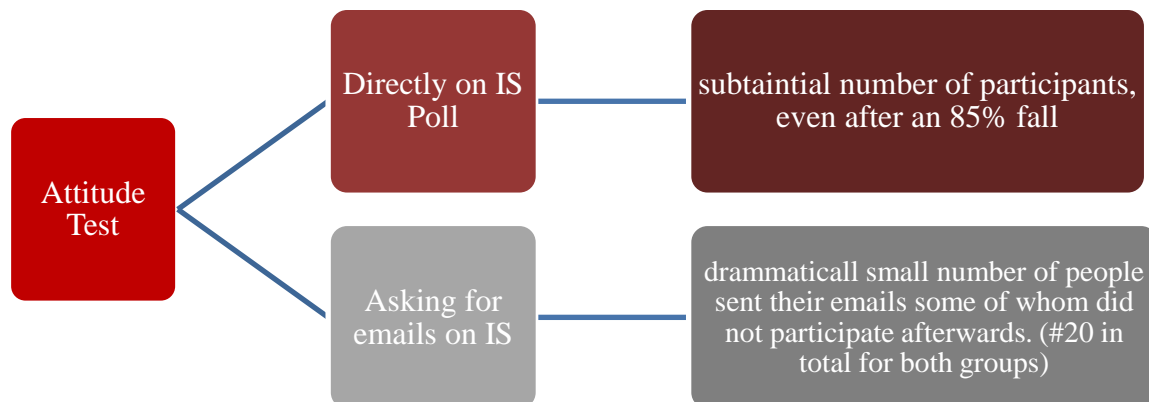


Figure 16 - the number of participants compared between AT1 and AT2.

The AT2 was the second part of the questionnaire file that was attached to the production task (figure 14). It was comprised of 12 yes/no questions which were designed to assess the socio-political and socio-linguistic status of Gilaki and the general information regarding L varieties discussed in the previous chapter. This survey was only completed by the experimental group (and not the control Farsi group).

3.6 The participants

In order to conduct this study, I needed to access Farsi-English speakers from Iran, with the experiment group being from Guilan only and the control group from Tehran. Tehran was chosen as the city of the control group as a base for Standard Farsi for two reasons:

1. The variety spoken in Tehran is usually regarded as the closest to the standard Farsi, pronunciation-wise at least.
2. Other local varieties spoken in Iran have the potential to include some of the vowels Standard Farsi misses. This was a confounding factor I needed to avoid as much as possible.

In terms of English proficiency, the ideal participants were those with no more than intermediate command of English. As mentioned in the literature review, proficiency and a considerable time of training can have a large effect on the target-likeness pronunciation of the new language. Therefore, I needed a group of bilinguals roughly between A2 and B1 proficiency levels of the Common European Framework so that there are more chances of observing transfer from their L1. However, the critically low number of the recruited

participants forced me to remove the proficiency filter so that more people could take part in the study.

In addition, for the experimental group, the ideal population would be two separate groups of active vs passive users of Gilaki, however, this seemed to contract the study population down much further, given the limited time. Another contributing factor that could have been attended to more strictly is *age*, when Gilaki-English participants over a certain age would be classified as an additional group. However, there is not much inconsistency in terms of age among the participants, the study fails to include people above 45 which would include the population who are more proficient in Gilaki. Overall, the recruited group of participants turned out to be quite homogeneous in terms of age.

In Summary, the study was designed aiming to recruit more participants - an ideal population (figure 17) and was subsequently modified according to the available population (figure 18).

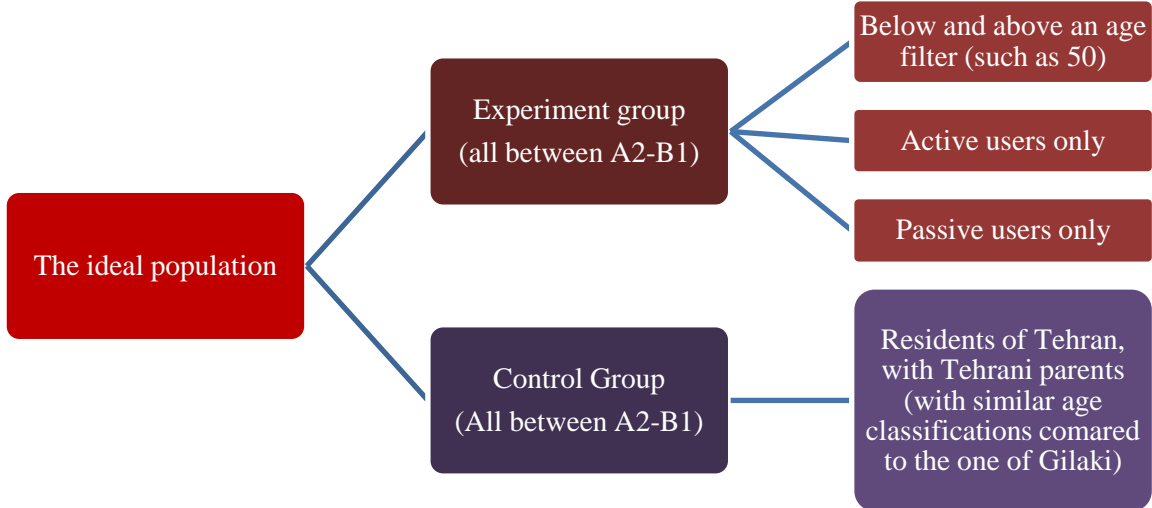


Figure 17 - The ideal population for the study

In the background questionnaire, participants chose from the three proficiency levels elementary, intermediate and upper-intermediate. This classification was subsequently converted into a numeric level scale of 1 to 3. The mean proficiency of the two groups, 2.09 and 2.44 for the experiment and control group respectively, can be observed in figure 18.

The participants of the experiment group aged from 18 to 40 none of whom, except one, were active users of Gilaki. The Age range of the control group (2- to 41) was close to the experiment

group. In addition, the participants from the control group reported that their parents came from a mixed variety of linguistic (mostly Iranian languages) backgrounds.

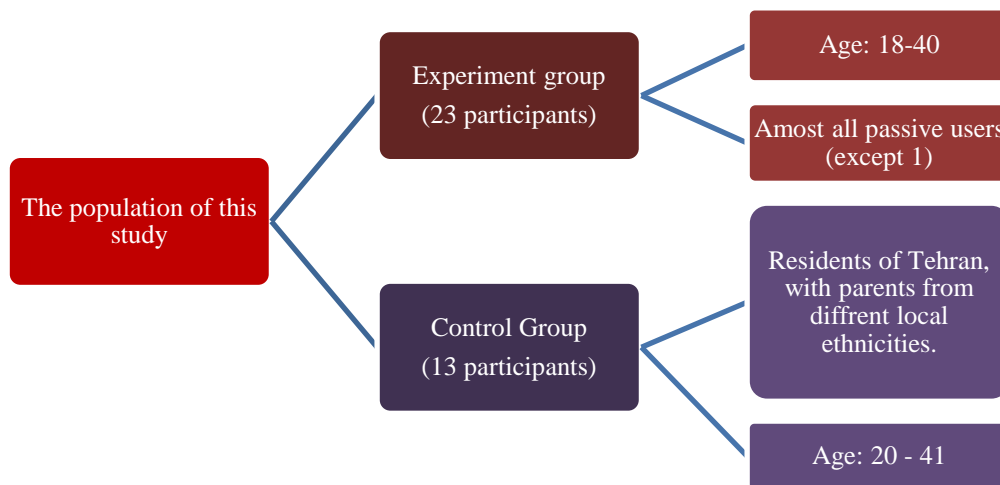


Figure 18 - The population of this study.

In the next chapter, I will discuss the results of the experiment and explain how the present population might have affected the results in one way or another.

4 Data Analysis

4.1 Introduction

Initially, I had proposed that there are differences between the two closely related varieties of Standard Farsi and Gilaki, vowel wise. The proposition was that Gilaki possesses a larger repertoire of vowels and is more marked, including the lax vowels of /ɪ/, /ʊ/, /ʌ/ and /ə/, which might manifest themselves in the performance of the participants, from the experiment group. It was also mentioned that, in Standard Farsi, the same lax vowels above are assimilated into their tense counterparts as in /ɪ/ to /i:/, /ʊ/ to /u:/, /ʌ/ to /ɑ:/ and for schwa it is assimilated into /æ/. These differences naturally led to a comparison between the lax vs tense in the produced target words/sounds in the test. One, and probably the most significant feature making a differentiation between the lax and tense vowels is the duration of their production. Therefore, one way to evaluate the performance of each group was to measure the duration of each target sound produced by the participants and compare them to the one of the other group, to see if there are any considerable differences.

However, this would not be sufficient for three major reasons:

- 1) The duration of an utterance can be influenced by:

- a) The proficiency of the reader/participants – lower proficient participants tend to read with hesitation and thus more slowly, especially in case of new words.
 - b) The nature of the word itself, whether it carries a positive or negative emotional valance.
- 2) If both groups in the experiment produce similar durations they might still have differences in their F1 and F2 values.
 - 3) Only the duration of a vowel cannot determine whether a vowel is lax or tense, especially when produced by non-native speakers. Therefore, for a more precise comparison, a second criterion needed to be applied in: Quality difference.

Quality of a vowel can be defined as the general position of the tongue, jaw, lips etc. and the voice setting that leads to the production of a certain sound, which is different from another sound regardless of its duration. For Example, the vowels /I/ and /i:/, short and long respectively, both can be intentionally produced at one unique duration. In other words, *bin* /bIn/ and *bean* /bi:n/ can be intentionally uttered in a way that both would take the same amount of milliseconds, whether for emphasis or some other reason. However, in a more natural setting, *bin* is expected to bear a shorter vowel duration. This prolonging of the vowel /I/ in *bin* does not necessarily call for a conversion into /i:/sound. In fact, the short and long characteristic of these vowels are features of that sound which seem to always accompany the quality of that very vowel when produced by a native speaker (table 5).

Vowel	Quality	Duration
I	X	Short
i:	Y	Long

Table 5 - Quality of vowels

As you can see in the table above, one can conclude that:

- 1) If the target words/sounds are produced according to a standard variety (or by a native speaker for example), one can easily differentiate between the lax and tense vowels based on quality or duration since in a standard variety these two features accompany each other according to an expected pattern. In fact, /i:/ is expected to be both tense and long so if it is long it is expected to be /i:/.

- 2) On the other hand, even if there is some degree of modification in an utterance, such as for emphasis, which might affect the duration of that utterance/vowel, there are still two other features that are more resistant to change: the F1 and F2 values.

4.1.1 Assessing the duration and vowel quality

Once the responses have been collected, there could be two ways how to evaluate the production of each group and then compare the two groups with each other. One could choose between:

- 1) Native speakers' judgment (NSJ).
- 2) A software-aided evaluation (SAE).

Both evaluation procedures have their own pros and cons (table 6). For instance, NSJ can be less time-consuming which could be an option for a rather immediate result in a limited amount of time. However, it can be less objective since judgments have to be made based on listeners' personal interpretations. To avoid error rate of subjectivity and/personal interpretations, there may be a need for a considerable number of listeners' so that a more reliable common ground can be extracted out. In addition, the judgments the listeners provide can also be influenced by other features of the produced words, not necessarily vowels, as well as familiarity (or unfamiliarity) with the L1 of the participants.

On the other hand, an evaluation done by a computer software leads to a more objective and interpretation-neutral outcome. At the same time, any software needs to be provided with compatible raw material and a set of guidelines and standards to function accordingly. This may not always turn out to be very straightforward particularly if the data contains some degrees of variation. To avoid errors requires a considerable amount of time devoted to the modification and preparation of the data. I will get into more details in the next sections since semi-automatized analysis was the method applied for the data analysis in this study.

Pros	Most objective	Provides explicit detail	Performance does not change due to external factors
ASE			
Cons	Can be time-consuming (depending on the data)		Needs to be modified to be software compatible

Pros	Less objective	Cannot provide explicit detail	Needs a large population of listeners	Easily affected by external factors (personal interpretations, bigger number of languages involved, etc.)
NSJ				
Cons	Less objective	Cannot provide explicit detail	Needs a large population of listeners	Easily affected by external factors (personal interpretations, bigger number of languages involved, etc.)

Table 6 - NSJ vs SAE.

4.2 Coding the data with Praat

As mentioned in the previous section, in order for the data to be processed objectively, it had to undergo a cleaning process. There are three different types of software I used in this study:

- 1) BAS Web Services (BWS)
- 2) Praat
- 3) R Studio

In the upcoming sections, I am going to explain each of these three steps.

BWS (BAS 2021) provides a series of services for phonological analyses. In this study, I have used it as a tool to provide the necessary input, the textGrid files, for Praat. TextGrid files are simply text files with grids specified to different segmentation such as sentences, words, vowels, etc. (figures 19 and 20). The textGrid files are subsequently imported into R for further analysis (figure 22).

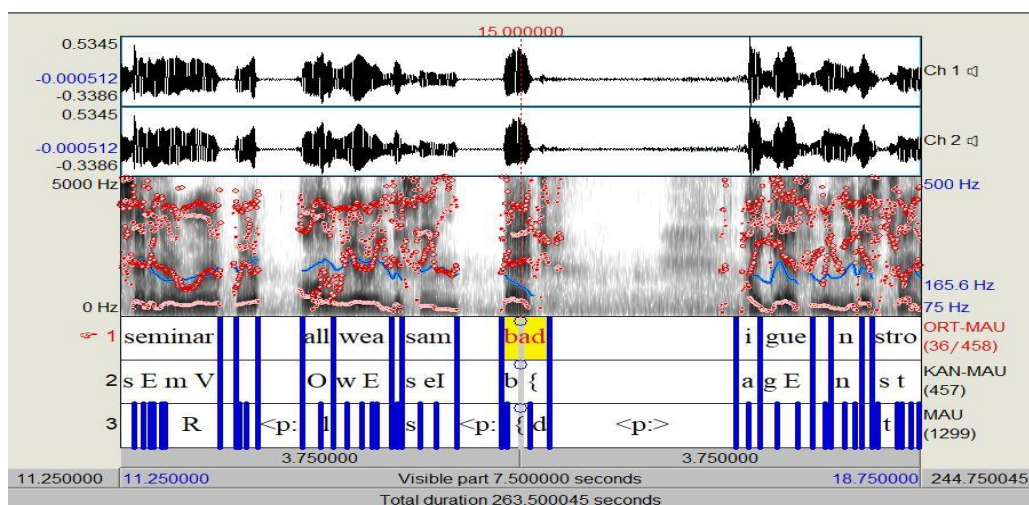


Figure 19 –Praat showing segments at sentence level.

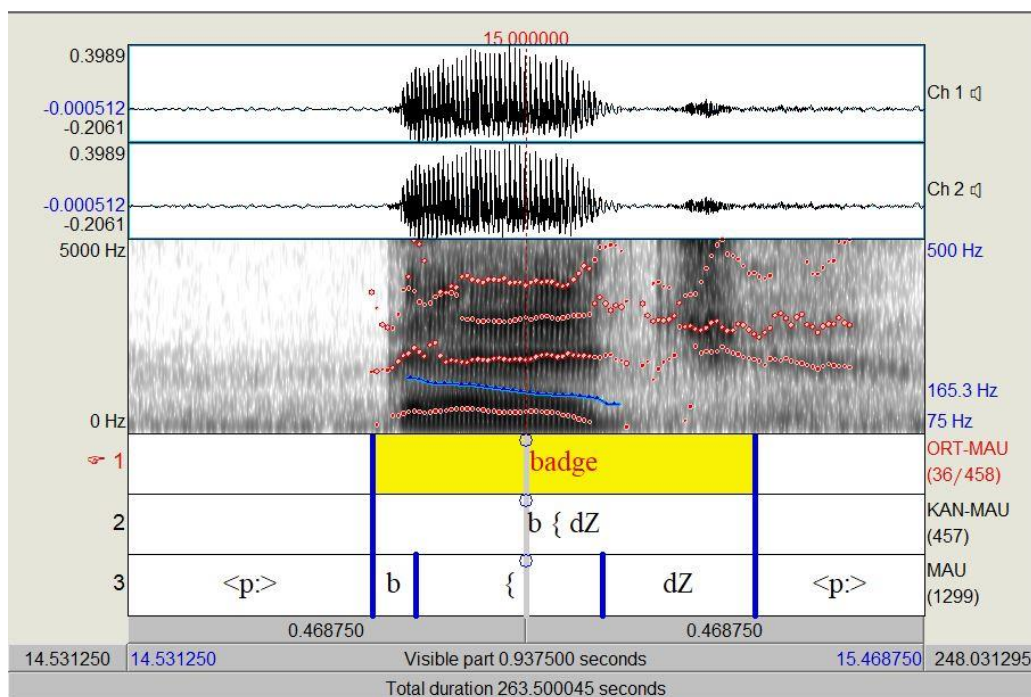


Figure 20 – Praat showing segments at word and vowel level.

TextGrid files can be opened by text editing apps, however, opened by Praat, they can be aligned with their audio file where each grid/segmentation is in alignment with its utterance in the audio file.

To have textGrid files, one needs to upload the audio (preferably WAV) and the text file (excluding certain characters such as apostrophe, comma, etc.) to BWS. Secondly, BWS asks you to select the language used in the file (figure 21). In this study, I set the language on American English, since there were no participants with either British or Australian Englishes and based on my intuition that American English is slightly more similar to Standard Farsi (and Gilaki) in terms of phonology. Finally, the software produces a textGrid to be download.

WebMAUS Basic

(allowed formats are: aiff, au, avi, tiac, tiv, mpg, mp3, mpeg, mp4, nis, nist, ogg, snd, sph, wav, doc, docx, odt, pdf, rtf, txt) or multiple signals all to be paired with the same annotation file _TEMPLATE_FILE_.doc|docx|odt|pdf|rtf|txt.

Service options

Language English (US)

Output format Praat (TextGrid)

When selecting 'emuDB' (EMU-SDMS) as output format, the service will pack the resulting EMU-SDMS database into a ZIP file, which can be retrieved by clicking on the 'Download as ZIP-File' button.

Figure 21 - Uploading files to BAS Web Services in order to download textGrid files.

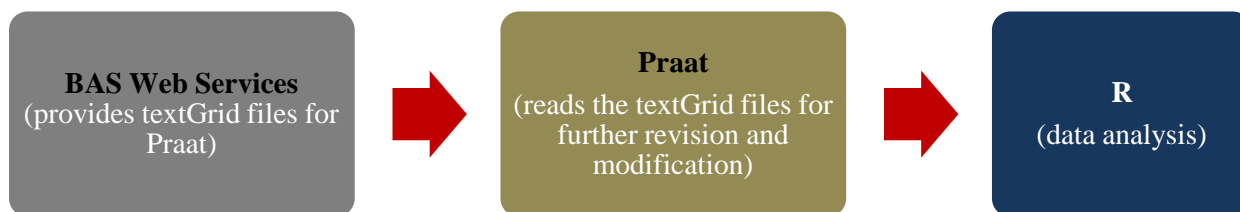


Figure 22 - Three main stages of data analysis from coding the files to analysing them in the R environment.

Praat requires two main input data before any further progression:

- 1) textGrid file.
- 2) WAV audio file.

Subsequently, Praat opens a window with three tiers which, in this study, have been specified as (i) sentences/words, (ii) sounds (for each word) and (iii) the individual sound level. As evident in figure 19, tier 3 demonstrates the individual sounds (in SAMPA), tier 2 provides the SAMPA phonetics for every word, and tier 1 provides information about the sentences with isolated words. In the middle of the window, right above the tier, you can see the formant visualisations and finally on the top there are the soundwaves of the audio files.

To listen to an individual sound, for example the /i:/ sound in *people*, one needs to click on the segment associated with that very sound in tier 3, play the sound and hear it (figure 23). However, this may not turn out to be as straightforward as it seems. The accuracy of the final segmentation and in other words the segment specified to every single sound can be influenced by three main factors:

1. The performance of the participant/reader:
 - As I mentioned earlier, BWS requires to choose the language of the input data which also means that it only recognises a given language, English in this case, based on comparing it to the standard varieties available on the website. Consequently, if the recorded utterances contain certain local qualities that are different from the chosen standard, the final textGrid is very likely to contain inaccuracies. In other words, the more native-like the input, the more precise the textGrid file. In a study involving intermediate users of non-native speakers of English we could expect such inaccuracies.

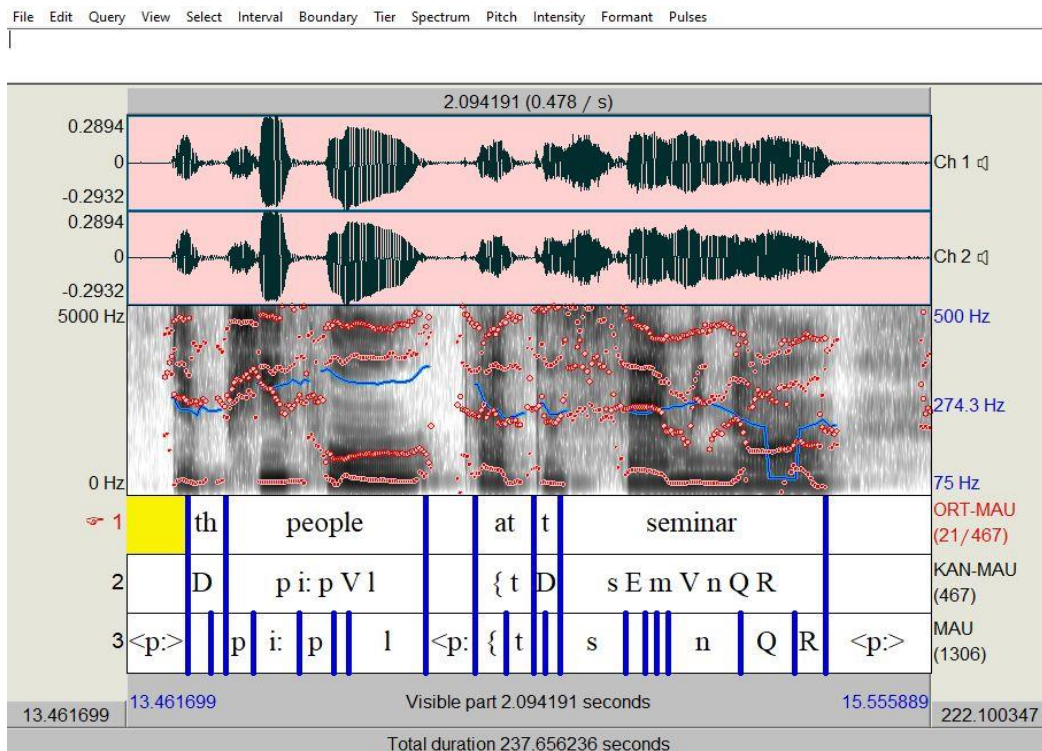


Figure 23 – listening to and adjusting the vowels in Praat.

2. The quality of the audio file:

- The text and the audio file need to match in every detail. If there are background noises or randomly uttered words or sounds the alignment between the tiers and the tiers with the audio file can be easily distorted.

3. Random and unpredicted factors:

- There have been cases in this study where the alignments were substantially distorted and out of order for no clear reason.

As a result, almost all the files from this study had to undergo a second manual alignment, either for some of the sounds or for all of them. Nevertheless, this must also be highlighted that a manual modification of alignments is generally inevitable at this stage, for the sake of the nature of the sound analysis itself because:

1. The textGrid files on Praat only provides a general segmentation of sounds and sound analysis is based on milliseconds. Therefore, Praat cannot precisely decide which spectrum of the sound one is looking for or if the alignment between the two sounds of /p/ and /r/ for instance is accurate. The latter can become really complicated with nasals,

rhotics either before or after a vowel since they gradually get merged into one another. Consequently, one needs to review and modify the alignments manually again.

2. There can be, and have been, inconsistency in the symbols used for one unique sound. For example, Praat uses SAMPA codings (not IPA) and as an example in SAMPA the symbol @ stands for /ə/ however, there were countless cases where another SAMPA symbol was set in that position, for instance { for /ə/. This also had to be manually corrected.

Once the alignments and sound symbols have been modified and corrected, the file was ready to be used for the next step, in the statistical software R.

4.3 Data analysis with R

As mentioned previously, the analysis was done in the R statistical programming environment. The analysis was comprised of a few sections which I have listed below as an overall preview. Each section will be explained in detail. This does not include the data preparation commands discussed earlier. The following commands were applied into R:

1. Extract five formants with formant ceilings of 5000 and 5500 for men and women respectively regarding only the 54 target words (and disregarding the rest).
2. Extract F1, F2 and F3 at midpoint (although only the first two were used in the analysis)
3. Load and merge the extracted data from both groups into one table.
4. Convert SAMPA encodings into IPA.
5. Detect and display the vowels only in the target words.
6. Model the data
7. Add data from native speakers' (reference data)
8. Model the data with the reference data added⁵

The total number of files (participants), excluding the corrupted files, was 36 with 23 and 13 for the test and the control group respectively. The proficiency of the participants was measured based on self-assessment and/or reporting if they were currently studying at a specific level of a language school. Participants chose between elementary, intermediate, and upper-intermediate levels which was subsequently converted into a numeric scale from 1 to 3. As a result, the average means of proficiency for the two groups were 2.09 and 2.44 for the

⁵ Stages 7 and 8 were subsequently excluded from the analysis. This will be explained in the discussion section.

experimental Gilaki (test) and control Farsi group respectively (table 7) which indicates a significant difference between the proficiency of the two groups.

	group	mean_prof
1	control	2.443992
2	test	2.096512

Table 7 - Means of proficiency of the two groups.

4.4 Formant Ceiling

At this level, we have detailed data from every individual's productions in the prepared textGrid files. However, the focus of the study is only on certain sounds within certain target words embedded in sentences. Therefore, the next step was to extract the appropriate formant from each audio file and from the target words only (figure 24).

Men and women are typically associated with lower and higher pitch respectively. For this reason, to extract the right formant, the formant ceiling needs to be set to a value that suits the (gender of the) speaker. For women, the average formant ceiling is 5500, which covers a higher register, and for men it is 5000. Therefore, files from both groups of test and control had to be processed according to a gender-register classification. At this stage, every file was processed individually, and the same function was applied to every single one (figure 24).

It should be clarified that these two formant ceilings only represent the average value and may not always succeed to extract the most informative segment of the audio file. In other words, for a more in-depth analysis, one may also try different formant ceilings especially in cases where a participant has a vocal register significantly different from the average criteria. For example, there can be speakers from either (any) sexes who demonstrate a distinctively lower or higher register compared to the average register they may generally belong to. In this study, there were no participants who did not fit into either group, pitch-wise.

```

75- ```{r}
76 # extract formants from file
77 Pathwav = paste0(absdir, mydir[1])
78 PathTextGrid = str_replace(Pathwav, ".wav", ".TextGrid")
79 TextGridInfo = tg.read(PathTextGrid)
80 T1 = data.frame(tmin = TextGridInfo[1][[1]]$t1,
81               word = TextGridInfo[1][[1]]$label)
82 T3 = data.frame(tmin = TextGridInfo$MAU$t1,
83               tmax = TextGridInfo$MAU$t2,
84               label = TextGridInfo$MAU$label) %>%
85   dplyr::filter(label != "")
86 TextGrid = dplyr::left_join(T3, T1, "tmin") %>%
87   tidyr::fill(word)
88
89 target <- c("did", "some", "clean", "badge", "glue", "fee", "accident", "see", "seems", "humid",
"through", "dust", "accurate", "book", "increase", "rdiculous", "much", "sadnwich", "bank",
"harmony", "gig", "food", "show", "bullet", "citizens", "doughnut", "grown", "understood",
"decode", "accept", "alone", "looked", "account", "loan", "car", "umbrella", "large", "park",
"band", "ring", "darling", "dark", "soup", "article", "push", "crook", "teens", "stupid",
"cook", "practicing", "such", "sugar", "live", "clue")

```

Figure 24 - Extracting the appropriate formants from each file.

The extracted formants from men and women were put together in groups of test and control group saved as text documents (figure 25).

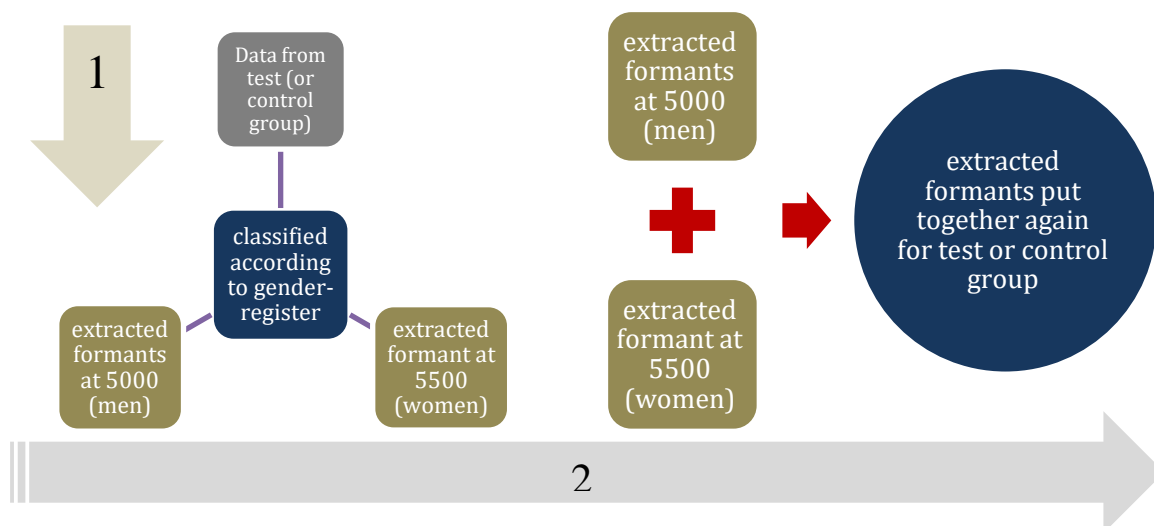


Figure 25 - Process of extracting formants according to gender-register classification in the test and control groups.

```

ae5jul_f5.txt - Notepad
File Edit Format View Help
"tmin" "tmax" "label" "word" "file" "midpoint" "f1_mdpt" "f2_mdpt"
"1" 2.48857110638604 2.56745977125859 "I" "did" "AE5JUL"
"2" 3.61 3.72823223267287 "V" "some" "AE5JUL" 3.6691161163
"3" 8.24 8.35 "i:" "clean" "AE5JUL" 8.295 317.81 2437.03 3249
"4" 13.0154031384363 13.2912984147628 "{" "badge" "AE5JUL"
"5" 15.768644008609 15.91 "u:" "glue" "AE5JUL" 15.8393220043045

```

Figure 26 - an example of a file with extracted formants.

The final output was saved under the participant's code and the same process was repeated for all the 36 participants separately.

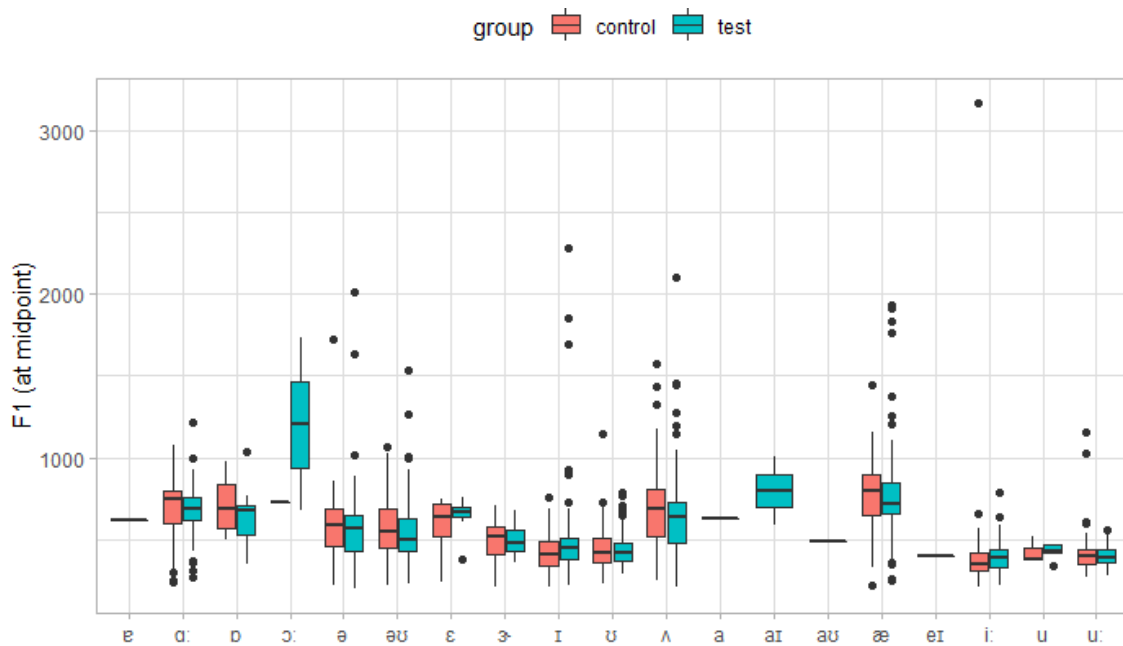


Figure 27 - an example of all the vowels and filler at their F1.

Figure 27 represents the boxplots of the vowels, including both target and fillers, from both groups. The fillers were initially included in the data for two main reasons:

1. Mainly to have a wider distribution of vowels (and diphthongs) to avoid repetition of the target vowels.
2. To analyse the fillers to see if there are any substantial differences in the pronunciation of non-target vowels between the two groups.

However, I decided to postpone the second objective to another study since this was not the focus of this study. The study became narrowed down to the target vowels only.

Subsequently, the target words/vowels were classified according to the criteria in table 8.

Traditionally, the tense vowels on the right are also called the long vowels, only taking their duration into account. However, this classification fails at the last row of the vowels, /ə/ vs /æ/, since neither could stand as the shorter or longer form of the other. In fact, /ə/ is the problematic vowel since it does not seem to have a 'long' counterpart. Nevertheless, the /ə/-/æ/ contrast was interesting to include in this study. As I have already mentioned in the literature review, /ə/ gets quite often assimilated into /æ/ in Standard Farsi. Therefore, it may be safe to say that they could be considered as counterparts for the purposes of this study. These two vowels were labeled X and analysed separately from the main (traditional) lax/tense group of vowels.

Classification 1		Classification 2			
Lax	Tense	Lax	Tense	Lax (x)	Tense (x)
/ʌ/	/ɑ:/	/ʌ/	/ɑ:/	/ə/	/æ/
/ɪ/	/i:/	/ɪ/	/i:/		
/ʊ/	/u:/	/ʊ/	/u:/		
/ə/	/æ/				

Table 8 - Classification of the vowels.

5 Results

The comparison between the two groups could be observed from two points of view:

1. Whether the two groups have demonstrated a similar or different performance between the lax vs tense vowels **in general**.
2. Whether the similarity or the difference is present at the vowel pair level.

Therefore, a preview of the models that I have constructed are provided below:

1. Model 1. (all vowels except for /ə vs /æ/)
 - Model 1-1. A general comparison of the **duration** of the lax and tense vowels between the two groups.
 - Model 1-2. A general comparison of the **F1** of the lax and tense vowels between the two groups.
 - Model 1-3. A general comparison of the **F2** of the lax and tense vowels between the two groups.
2. Model 2 (model x). Comparing /ə/ vs /æ/
 - Model 2-1. This model examines the **duration** between /ə/ vs /æ/ between the two groups.
 - Model 2-2. This model examines the **F1** between /ə/ vs /æ/ between the two groups.
 - Model 2-3. This model examines the **F2** between /ə/ vs /æ/ between the two groups.

3. Model 3

- Running the same model for individual vowel pairs regarding their duration, F1 and F2 (similar to Model 2)

The most immediate observation is that the two groups demonstrated almost equal means of duration, F1 and F2 in uttering the target vowels, with an average duration of (at) less than 0.125 ms (figure 28), and an average of nearly 500 hz for their F1 and about 1500 hz for their F2 for both groups (figure 29 and figure 30).

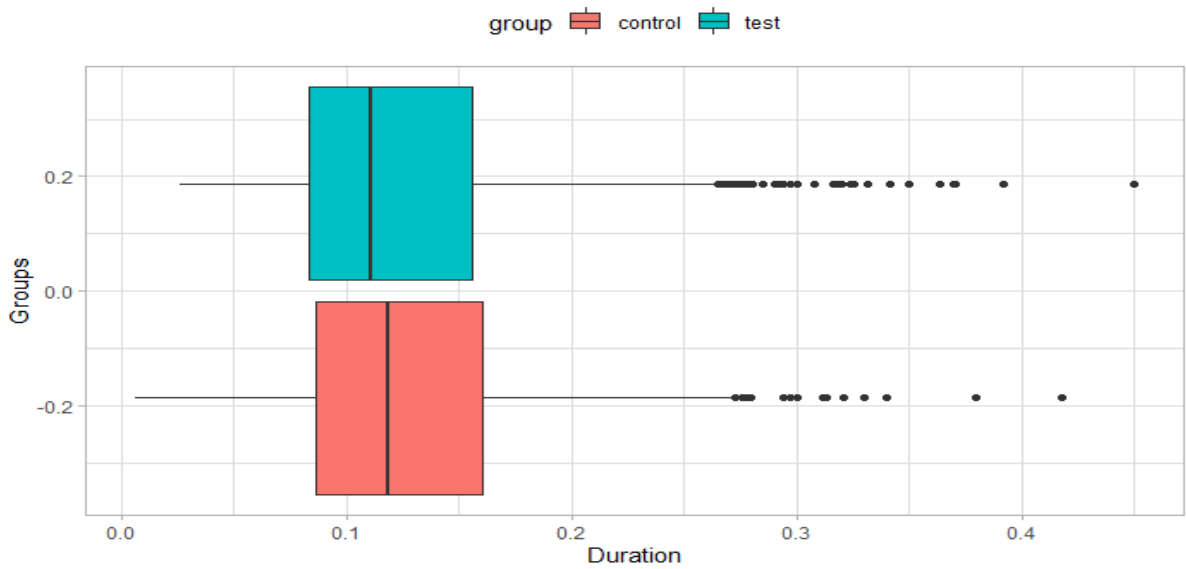


Figure 28- general comparison of the duration between the two groups.

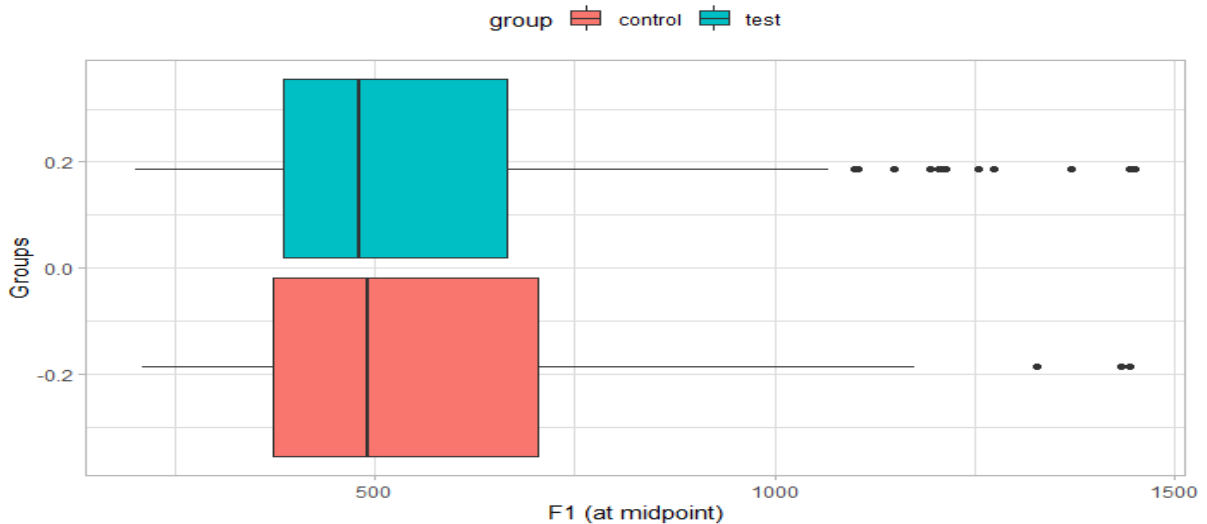


Figure 29 - general comparison of the F1 between the two groups.

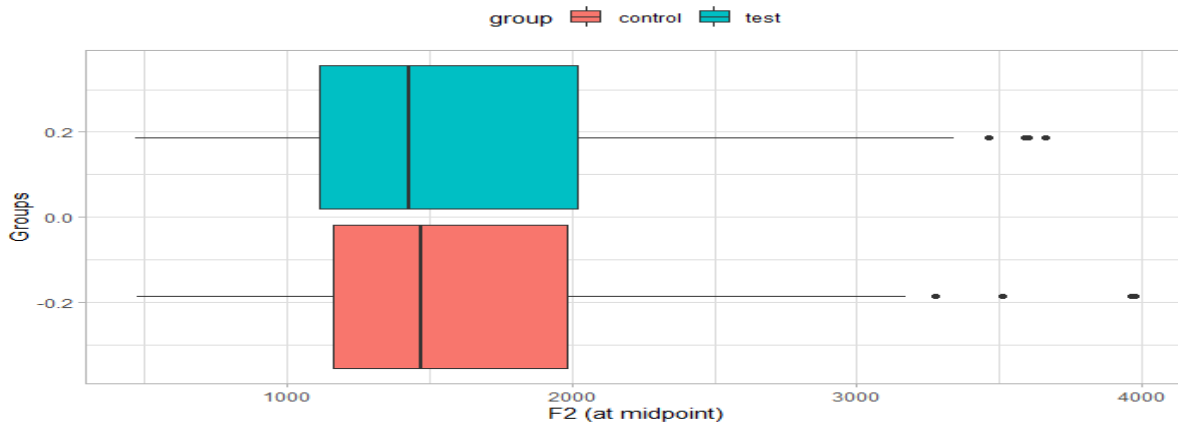


Figure 30 - General comparison of the F2 between the two groups.

Model 1-1

For the first model, I fit a linear mixed effect model, taking the duration as the dependent variable using the proficiency and the word and file (participants) as the random effect. Results demonstrate that there is significant effect of the condition ($\beta = 0.04$, $t\text{-value} = 3.7$) and proficiency ($\beta = -0.01$, $t\text{-value} = -2.7$). However, the results show no significant interaction between groups and conditions, ($\beta = -0.003$, $t\text{-value} = -0.8$) suggesting that both, the test and the control group have demonstrated a very similar performance, and produce tense and lax vowels differently in terms of duration.

Model 1-2 and 1-3 (Traditional lax and tense vowels)

A similar result was obtained on the comparison between the F1 and the F2 for the tense and lax vowels (figures 32 and 33). Both groups demonstrate a very close performance distinguishing between lax and tense vowels in terms of the fundamental frequencies quantified as F1 and F2.

I used as similar R syntax for the second and the third model as in model 1 discussed above, but with the F1_mdpt (model 1-2) and F2_mdpt (model 1-3) as dependent variables respectively, to see if the two groups demonstrated any differences in their first and second formants dependent on the group and condition. The model with F1 as the DV (1-2) demonstrated a significant effect of the condition ($\beta = 34$, $t\text{-value} = 0.8$). However, there was no significant interaction between the group and condition ($\beta = -14.01$, $t\text{-value} = -0.1$). Similar to model 1, there was a significant effect of proficiency ($\beta = 28$, $t\text{-value} = 2.1$).

Finally, the next model (1-3), compared the two groups across the two conditions (lax and tense) at their F2, and the results were as follows. The model revealed a significant effect of condition ($\beta = -57$, t -value = -0.5) and proficiency ($\beta = 126$, t -value = 3.1). The interaction between group and condition was not significant ($\beta = -16.4$, t -value = -0.3). Table 9 below summarizes model output for models 1-1, 1-2 and 1-3, and figures 31-33 provide a visual illustration of the obtained data.

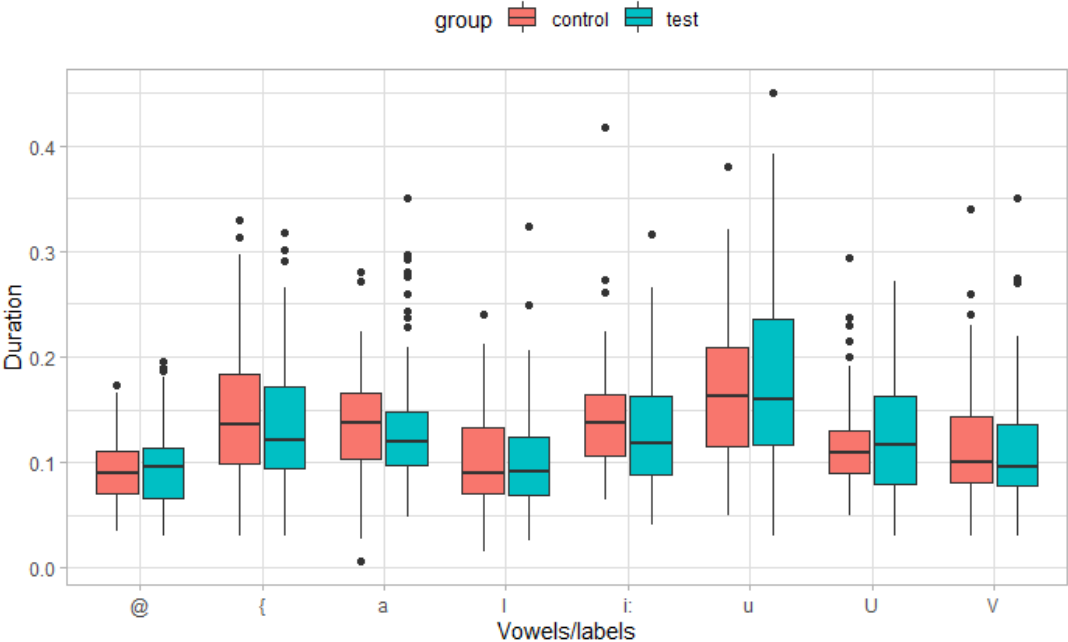


Figure 31 - comparison of the duration of the vowels between the two groups.

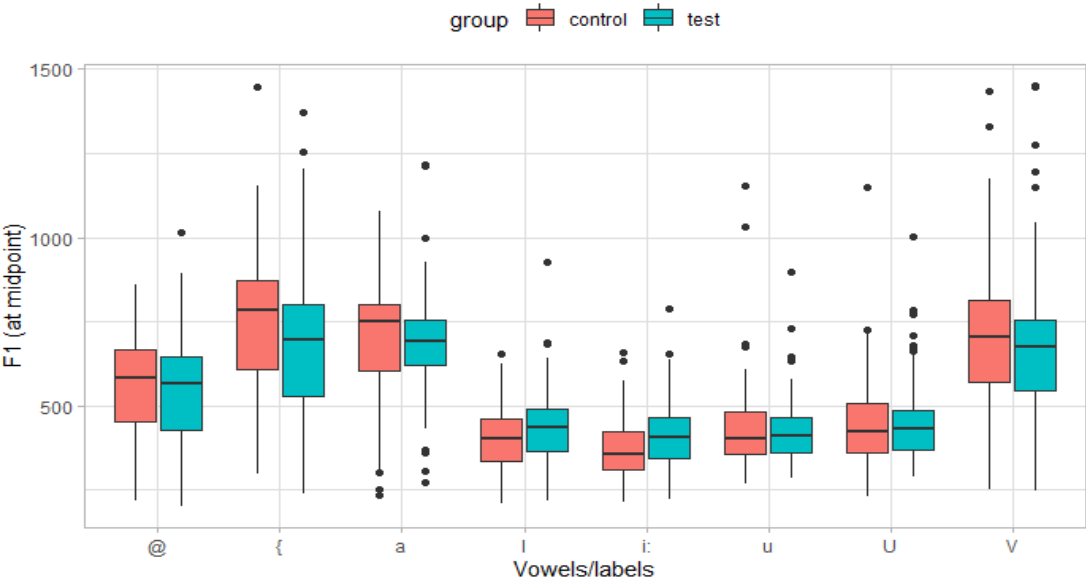


Figure 32 - comparison of the F1 of the vowels between the two groups.

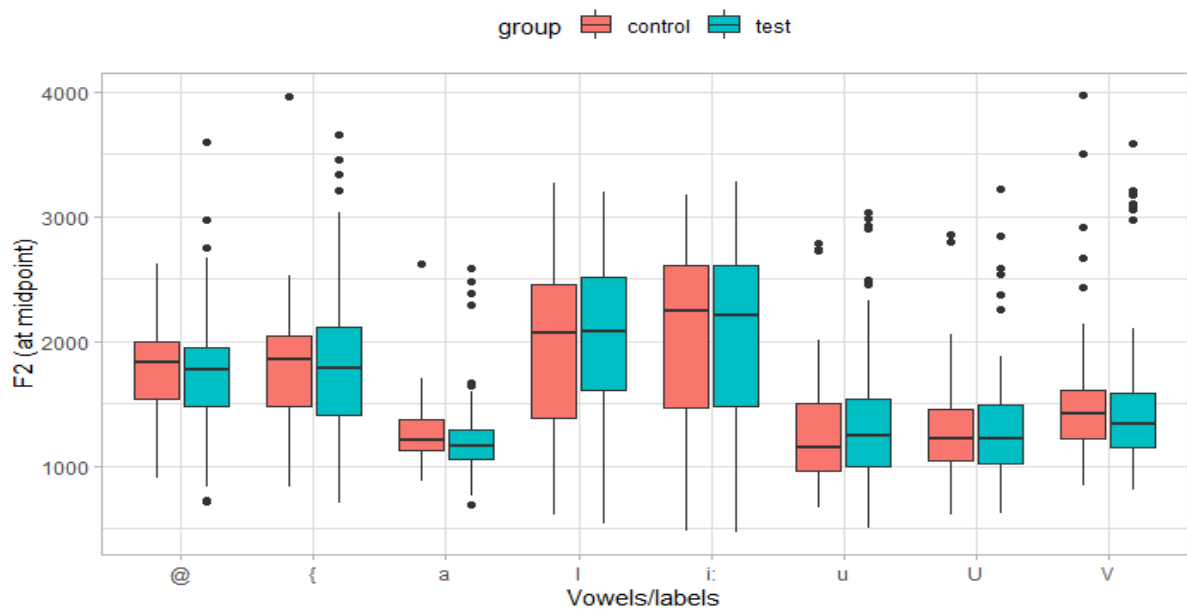


Figure 33 - comparison of the F2 of the vowels between the two groups.

Table 9 demonstrates the summary of the results of the model 1 series.

Model 1-1, 1-2 and 1-3

	Duration		F1		F2	
	Beta	t value	Beta	t value	Beta	t value
Intercept	0.1	9.15	457.3	9.965	1302.6	9.8
grouptest	-0.004	-0.6	8.9	0.4	42.4	0.6
conditiontense	0.04	3.7	34.02	0.8	-57	-0.5
prof	-0.01	-2.7	28.08	2.1	126.7	3.1
Grouptest:prof	-0.003	-0.8	-14.01	-0.1	16.4	0.3

Table 9 - Summary of the models 1-1, 1-2 and 1-3 (fixed effects).

Model 2-1 (comparing /ə/ vs /æ/)

I ran the same models for the vowel pair of @ and { (/ə/ and /æ/) labeled pair x. As it can be observed, the two groups demonstrated a similar performance in their production of the pair x vowels. Both groups produced an average of about 0.1 ms for @ (/ə/) and an average of 0.130 for { (/æ/). Although, in the latter, the difference is slightly more pronounced: the test group has demonstrated a slightly shorter duration (0.125) compared to the one of the control group

(almost 0.140) which is slightly larger than their difference in performing @ (see Appendix B for figure 34).

The same pattern could be observed in the difference between the F1 and F2 for the two vowels (figure 35 and 36 in Appendix B). Figure 35 shows that the two groups may show a difference in their F1s for their /æ/).

However, the models for the pair x vowels demonstrated that there was a significant effect of condition for the duration ($\beta = 0.06$, t-value = 2.3) and the F1 ($\beta = 190.2$, t-value = 3.6). Proficiency had a significant effect on the F1 ($\beta = 53.7$, t-value = 2.9) and F2 ($\beta = 110.8$, t-value = 2.08), but not duration ($\beta = -0.006$, t-value = -1.1). There was a marginally significant trend in the interaction of group and condition for duration ($\beta = -0.01$, t-value = -1.8), however, the difference didn't reach real significance. Table 10 summarizes the results of the model 2 series.

	Pair x					
	Duration		F1		F2	
	Beta	t value	Beta	t value	Beta	t value
Intercept	0.1	4.2	431.2	7.1	1479.5	9.5
grouptest	0.004	0.4	7.4	0.2	7.8	0.08
Conditiontense	0.06	2.3	190.2	3.6	-11.3	-0.12
prof	-0.006	-1.1	53.7	2.9	110.8	2.08
Groupptest:conditiontense	-0.01	-1.8	-46.2	-0.983	38.4	0.3

Table 10 - Summary of the model x series (only fixed effects).

Although the results so far have revealed enough about the two groups, their performance and whether there is any interaction between them, I took the analysis further into detail to compare the two groups at the vowel pair level, similar to one of the pair x (Appendix C and D)

6 Results from the attitude tests

Recall that in the background section, I argued that to obtain a more comprehensive picture of a multilingual language profile, it is necessary to have some information about the participants' attitudes and language backgrounds. I reviewed such parameters as un/intentionality, attitude,

willingness to use, formal training and specific characteristics of L/H varieties which have been found important when analysing the effects of CLI in multilingual productions. In this chapter, I present data from two attitude tests (AT1 and AT2). AT1 was conducted as an independent questionnaire, before the current study and was subsequently revised and expanded to include more questions. AT2 was conducted as an addition to the background questionnaire, and was administered only to the Gilaki group.

In what follows I will present the results of both ATs. There are two reasons for this:

1. AT1 managed to recruit a significant number of participants. This is why the results are more informative (and statistically reliable) than a substantially smaller powered AT2.
2. The independent nature of the AT1, which was not a part of a formal study, conducted on a user-friendly platform (Instagram), might have led to more spontaneous responses from the participants.

6.1 AT1

Questions:

1. Do you use Gilaki for your daily conversations in general? (full sentences in Gilaki and not just individual words incorporated in Farsi speech)
2. Are you more likely to use Gilaki for joking, mockery, comedy, etc?
3. Do you use Gilaki to talk about serious and important issues in life?
4. If your answer to the previous question is NO, do you think Gilaki is incapable of conveying such concepts?⁶
5. Do you, subconsciously, tend to think that Gilaki is inferior and unstylish?
6. Do you intentionally try not to have a Gilaki accent in certain situations?
7. Do you find the current situation of Gilaki worrying (Yes) or a natural stage or process of any language (No)?

	No	No (%)	Yes	Yes (%)	Total
1	101	79%	27	21%	128
2	31	25%	92	75%	123

⁶ The ambiguity of this question was later compensated with further explanation to the participants.

3	95	76%	30	24%	125
4	98	88%	13	12%	111
5	102	82%	22	18%	124
6	66	54%	56	46%	122
7	23	19%	81	81%	104

Table 13 - Results from the AT1.

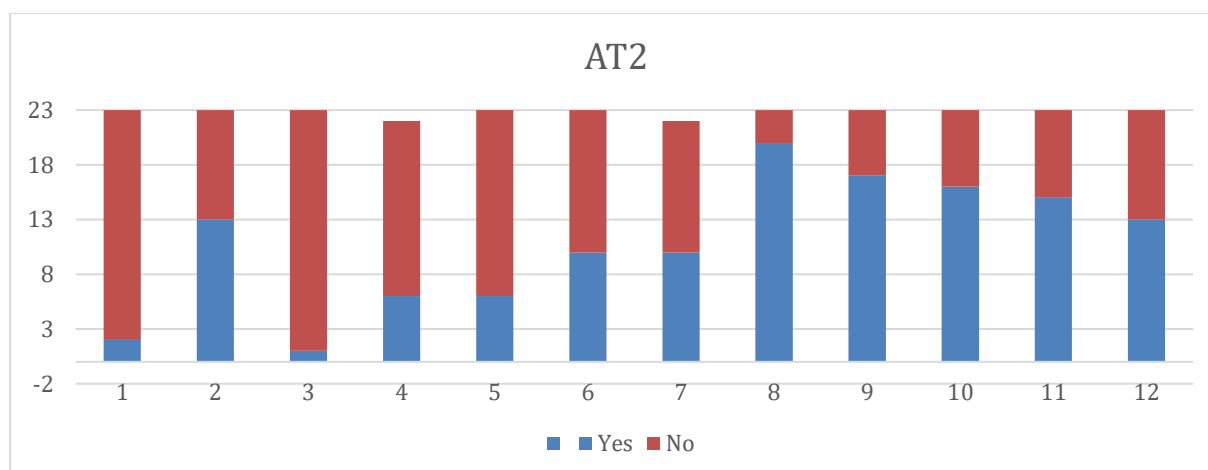
6.2 AT2

In the literature review, I briefly described the sociolinguistic situation of Gilaki. I also theorised that the modern state of Gilaki can be generally captured in three main stages:

1. Before the standardisation: used naturally by the locals (before 1921)
2. Standardisation of Farsi as the national language: decline in use by the new generations (between 1921 until ca 2000)
3. Revitalisation of Gilaki: increase in use but change in function (ca 2000 onwards)

Therefore, it was necessary to add a few questions to the AT1 which could target participants' attitudes more directly. The following five questions were added to AT1 so that the total number of questions was 12 for AT2.

8. Do you think it is worth preserving Gilaki?
9. If possible, are you willing to learn and use Gilaki?
10. Do you think Gilaki is a dialect but not a language?
11. Do you think a dialect, in general, has a lower status compared to a standard language (for example Gilaki compared to Farsi)?
12. Do you think Gilaki is one of the dialects of standard Farsi?



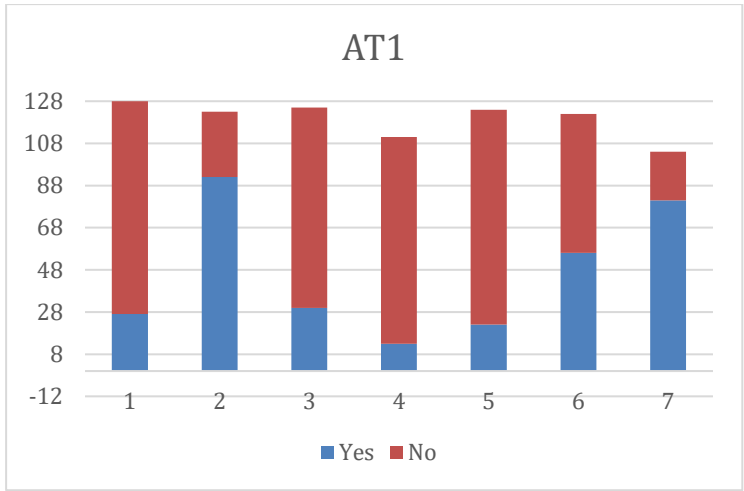


Figure 34 - AT1 and AT2 by the number of responses.

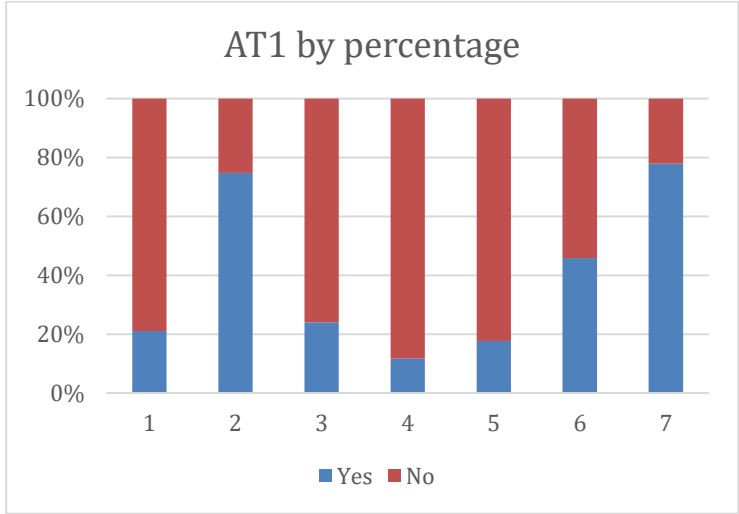
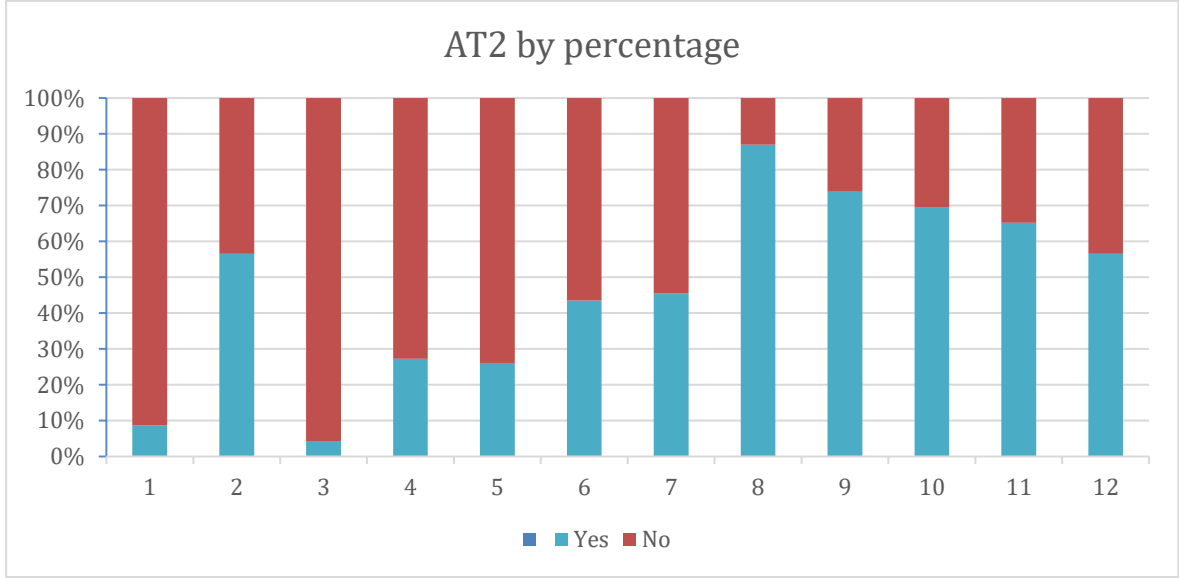


Figure 35 - At1 and AT2 by the percentage of answers for each question.

Although the two attitude tests had strikingly different population in terms of size, they seem to have a very similar pattern for the first seven questions they had in common. This may not be easily expandable to the five added questions in AT2, but the similar results across the first seven questions make AT2 more reliable in my view.

To sum up the core results from AT1 and AT2 :

1. The majority (79%) of respondents do not use Gilaki in daily conversations.
2. The majority (75%) of respondents use Gilaki for mockery and joking.
3. The majority (76%) of respondents do not use Gilaki to talk about serious issues.
4. The majority (88%) think however that Gilaki **is not incapable** to convey serious issues.
5. The majority (82%) think Gilaki is inferior and unstylish.
6. Around half of the participants (54%) try not have a Gilaki accent in certain situations.
7. The majority (81%) find the current situation of Gilaki worrying.
8. The majority (85%) think it is worth preserving Gilaki.
9. Most of the respondents (75%) are willing to learn and use Gilaki.
10. The majority (70%) think Gilaki is a dialect (not a language).
11. Most participants (65%) think Gilaki has a lower status compared to Farsi.
12. Around half (55%) think Gilaki is one of the dialects of Standard Farsi.

Based on the results of the survey, we can conclude that the attitudes of the Gilaki users is generally very similar to what I described in the background section. The users consider Gilaki a low status variety, use it mostly for a specific register (mockery and joking), try not to have a Gilaki accent in Farsi and overall consider it to be just a provincial (inferior) dialect of Farsi. At the same time, the users agree that Gilaki can be used across different contexts (also for serious conversations), and are willing to preserve, learn and use it. At the present moment, the majority of respondents are not actively using Gilaki in daily conversations (even in the province where this variety is mostly spoken).

In the next section, I am going to review the results from the main experiment and the attitude tests and try to create a meaningful link between the two.

7 Discussion

In this chapter, I am going to discuss the results of the experiments conducted for this Thesis. I start with repeating the research questions and predictions and explaining what the results have

and have not demonstrated in regards with the questions and predictions. Next, I discuss the findings in light of the theories presented in the literature review.

Research questions:

RQ 1: Do Gilaki speakers demonstrate a more pronounced differentiation in duration of tense vs lax vowels in English compared to the Standard Farsi speakers?

RQ 2: Do Gilaki speakers demonstrate a more pronounced differentiation in F1 and F2 values for the tense-lax contrast compared to the Standard Farsi speakers?

Predictions:

Pr 1: Gilaki speakers are likely to demonstrate a more pronounced differentiation in duration between tense and lax vowels.

I mentioned in the literature review that Gilaki, though more marked, is phonologically dominated by the Standard Farsi. Therefore, there is still a probability for Gilaki users (the test group) to utter similar vowels. However, my prediction is that there will be a difference.

Pr 2: Yes. Gilaki speakers are likely to demonstrate different F1 and F2 values.

Due to the phonological differences that still exist between Gilaki and Standard Farsi, I predicted that Gilaki users are going to demonstrate different spectral dimensions.

However, the results indicated that that we couldn't corroborate neither of the predictions of the study. It should be noted that null results do not indicate that the predictions were wrong, simply that the predictions could not be supported given the data that have been collected.

	Predictions	Results
Research Question 1	Yes	inconclusive
Research Question 2	Yes	inconclusive

Table 14 - Predictions and the results of the study.

7.1 Duration

There are not many studies on the Standard Farsi phonological system comparing it to the other (Iranian) varieties spoken in the country. Therefore, it appeared to be a challenge to say whether Gilaki or Standard Farsi speakers differ significantly in the duration of their vowels. However, it might also be a point of interest to know that Gilaki speakers are mostly known as louder and faster speakers especially by the people living in Tehran, where the Standard Farsi is most associated with. Nevertheless, I predicted that there should be significant differences between the groups in terms of their vowel durations of the English vowels, however, the models indicated a significant effect of condition for both groups, but no interaction. In other words, apart from some minor differences, the two groups demonstrated similar differences in durations for the critical vowels⁷.

This, however, raises a few more questions:

1. Is this because of the dominance of the Standard Farsi as the H variety especially on the younger generation that must have led to an amalgamation of phonological features and characteristics of Gilaki and Farsi?
2. Do Gilaki and Standard Farsi have significant differences in their vowel durations at the present moment?
3. Have both groups successfully managed to acquire the English vowels and what we observe is a ceiling effect?

7.2 F1 and F2

As I have already mentioned in the preceding chapter, the lax-tense distinction is typically differentiated in terms of vowel duration (long vs short). However, theoretically it is possible to utter a lax vowel longer than its normal duration. Therefore, the underlying characteristic responsible for the differentiation of that very vowel may exist independently of its duration. That characteristic is in fact the combination of its front/backness and open/closeness which can be summarised in terms of the spectral dimensions of the vowels. In addition, natural

⁷ I may also highlight that this prediction was initially made considering a population of lower proficiency who could have demonstrated a stronger transfer from their L1s.

duration of vowels can be easily distorted when uttered by elementary-intermediate L2/L3 learners because:

1. They may have not acquired the target like features of those vowels in addition to non-facilitative influence from their own L1/L2 phonological repertoire. Therefore, they are likely to transfer vowel features from their L1 or L2 (Gilaki or Farsi)
2. The very nature of the experiment and spontaneous exposure (and reading out loud) to the sentences can create hesitation or uncertainty which in return can lead to longer durations.

Consequently, it appeared to be essential to have a comparison of the first and second formants (F1 and F2) of the vowels between the two groups. The second prediction was that there would be an interaction between the spectral differences of the vowels in the two groups. Although both groups produced English vowels with comparable durations, I predicted that there still could be differences in the height-backness (F1 and F2) of the target vowels. This prediction was put forth based on the observation that Gilaki, with a larger repertoire of vowels, overlaps with English more than Farsi does. Therefore, Gilaki participants would be more likely to produce the vowels they have in common with English differently from Standard Farsi participants.

The second and third statistical models, constructed to test the second prediction, demonstrated that the two groups performed very similarly in terms of their F1 and F2 too. Once more, it was revealed that there was no significant interaction between condition and group in terms of the first and/or second formants.

F1 and F2 are typically associated with the general position of the tongue with F1 and F2 indicating the height and backness of the vowel respectively. F1 has an inverse relation with the vowel height, meaning that the higher the formant the lower the vowel height (2.2. Formants of Vowels – Phonetics and Phonology 2021). For F2, the relation is straight, meaning that the higher the formant the more front the vowel. So in summary, both groups articulated the vowels very similarly not only in terms of the duration but also in terms of the characteristics that every vowel possesses such as how high, low, front and/or back the vowel is articulated in the oral cavity. (Acoustic Phonetics: Formants 2021)

Although, pair x, indicated an exception with a strong effect of condition at the F1. This could be the result of the natural difference between the two vowels /ə/ and /æ/ which, in fact, do not form a real pair and have more distinct spectral differences. Compared to the other pairs, which are real minimal pairs, the differences between /ə/ and /æ/ are more than minimal in terms of how high and front they are (for a comparison see figure 39). Therefore, this could be said that both groups, the Standard Farsi group in particular, have managed to differentiate the spectral differences between the /ə/ and /æ/.

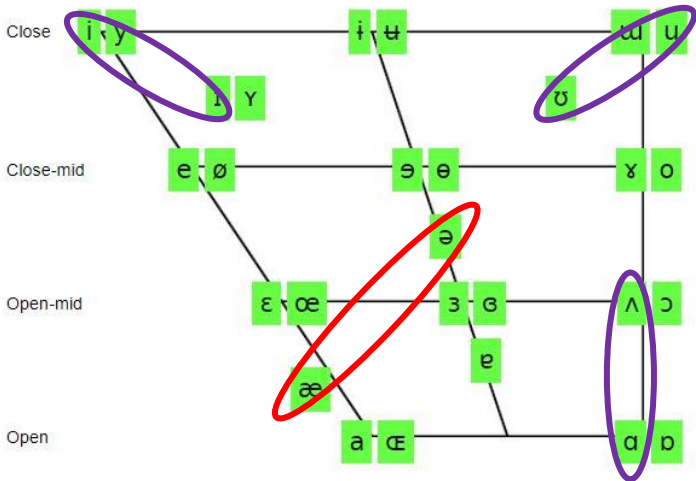


Figure 36 - spectral differences at pair x compared to the minimal pairs.

To reiterate the results, the models, in general, revealed two important results:

1. the significant effect of condition (lax vs tense)
2. the effect of proficiency

7.3 Effect of Condition

The first model revealed that despite the lack of a significant interaction between condition and group, there was a significant effect of condition. This is rather important since it indicates that both the test and the control group differentiated between the lax and the tense vowels in terms of duration in production. Therefore, it can be said that it is not likely that the participants have produced vowels, in particular, lax vowels, longer than the normal duration.

However, for the F1 and F2, there was no effect of condition which could mean that the two groups have only observed the duration of the vowel as the criterion for the lax-tense

distinction. If the Gilaki groups is considered to possess vowels with spectral differences, this could be interpreted as follows:

- They might have produced one almost unique vowel for both lax and tense condition but with only longer and shorter durations. This could be justified by accepting that the Gilaki speakers used the vowels from their Standard Farsi repertoire which only has tense vowels. Therefore, they should have classified them by adjusting the duration to compensate for the differentiation, but not reaching the target-like F1 and F2 differences.

The vowel-pair analysis also indicates that despite significant effect of condition in general, the effect is smaller for some pairs. For instance, for pair i, and pair x, the effect is at 2.1 and 2.3 respectively, however, for pair a and pair u, the figures are 1.6 and 1.7 which are close to significant, but not yet significant. This could mean that there is less differentiation for pair a and pair u in terms of their length (duration). In other words, /ʌ/ and /ɑ:/, and /ʊ/ and /u:/ are more assimilated into one vowel for each pair. However, for the other two pairs of i and x, the differentiation is more distinct.

At F1, there is a similar pattern of results. Pair i (-1.8) and pair x (3.6) are either significant or close to that whereas the other two pairs of a and u the figures are 0.1 and -0.1 respectively which indicates more differentiation of both groups in general for pair i and pair x and more assimilation of the vowels for pair a and pair u.

At F2, the results are somehow the opposite where the only significant effect of condition belongs to pair a with the t value of -3.4. The other three pairs are not close to significant differences at F2.

7.4 Effect of Proficiency

In the literature review, I referred to two studies (Wrembel 2019 and Lee 2020) which demonstrate that instructed training and proficiency have significant impact on the pronunciation of L2 learners. In other words, through instructed training which leads to higher proficiency with time, learners manage to acquire a more target-like pronunciation. Therefore, effects of CLI from L1 to L2 (or to L3) should be first and foremost observed at a lower level of proficiency.

As described in the methodology section, most of the participants in this study were an intermediate level of proficiency in English. The Farsi group was significantly more proficient than the Gilaki-Farsi group, with the means being 2.44 and 2.09 (out of 3), respectively. It may have been the case that we captured participants at the time when they have already overcome non-facilitative CLI from Farsi and have already acquired the English vowel system. At the same time, this should also be noted that the instruction itself could determine the way learners may acquire language features (Wrembel 2019), such as pronunciation. I believe that this is another reason why elementary students could have been a more suitable population for this study. For example, I know from personal teaching experience that more often than not, the only differentiation criterion Iranian students are given for lax vs tense vowels is the duration. I also referred to a study by Seddighi (2012) which reports that differentiating minimal pairs is a significant challenge for Iranian students. Therefore, we do not have enough information about the behaviour of the participants from the Gilaki-Farsi group, whether they had unlearned the F1 and F2 differentiation for the English vowels, or they never distinguished these values to any different degrees before the training.

The model 1 series indicate that there is a significant effect of proficiency for all vowel pairs, except for pair x, for their durations. Furthermore, the models for the vowel pairs indicate that the height (F1) of the target vowels, except for pair x, were not significantly influenced the proficiency, whereas for the F2 the effect of proficiency was significant effect for all pairs except for pair u. Mildner and Horga (1999) found similar patterns in a study on the relation between proficiency and vowel formants. They found that it is in fact the F2 which had a high correlation with proficiency of their L2 English learners. For F1, the correlation was insignificant. Therefore, it might be that the F1, the height of the vowel, may not be significantly influenced by proficiency. Further research is needed to provide further empirical support to this hypothesis.

7.5 Attitude, Intentionality and Transfer

Earlier in the literature review, I reviewed the “Characterisation of CLI Types Across Ten Dimensions” by Jarvis and Pavlenko (2008, 24). I referred to one of the dimensions called “intentionality” and discussed that learners can intentionally influence the strength of CLI based on a number of extralinguistic factors. This is interesting and quite unconventional for traditional approaches to CLI since according to this point of view any transfer or absence of

transfer may be modulated by a (semi)conscious decision which itself could be influenced by other external factors.

In the case of the present study, I argued that Gilaki possesses some of the prerequisites for its users (even if they use it at all) to intentionally inhibit the transfer of certain phonological features into their speech. The results from the models tell us that the two groups in general do not differentiate between the lax and tense vowels in terms of the spectral dimensions (F1 and F2 values). This result would be in line with transfer from L1 to L2 for the control Farsi group since their variety lacks that differentiation, whereas for the test Gilaki-Farsi group who have a more marked variety this looks like an absence of transfer. With the results from the attitude tests, this could be interpreted in the following ways:

1. The dominance and predominance of the Standard Farsi has already filtered out the more-markedness of Gilaki, which is then transferred into their English utterances.
 - 1-1. Half of the present population reported that they try not to have a Gilaki accent. In other words, it means that they prevent themselves from producing certain vowels. This figure could also be slightly higher since 82% of the participants from the test group reported that they find Gilaki inferior and unstylish. Therefore, it raises a question whether they might have made a conscious effort to unlearn those 'marked' (i.e., absent in Farsi) vowels.
 - 1-2. The process of (sub) conscious unlearning could be more probable since none of the participants of the test group, except two, reported that they actively used Gilaki on a daily basis. This is in line with the results from the AT1 where 79 % of participants reported the same.
- 2- The absence of the transfer could be due to explicit training and the way the lax-tense differentiation has been instructed.
 - 2-1. Participants of the test group may differentiate between the lax and tense in terms of the F1 and F2 values in producing Standard Farsi. However, they might have been taught to differentiate the English minimal pairs by duration only.

According to Bohn's Desensitising Hypothesis (1995, 294): "[W]henver spectral differences are insufficient to differentiate vowel contrasts because previous experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast". Overall, there are broadly three ways in which the results of this study can be linked with this hypothesis:

1. The Gilaki-specific vowels have died out at production level (for the test group in this study)
2. The Gilaki-specific vowels still exist but are suppressed by other extralinguistic factors.
3. My assumption on Gilaki possessing minimal pairs was wrong.

It does not seem straightforward to come to a conclusion whether the participants of the test group have unlearned, or simply lost, the ability to produce these vowels differentiating their F1s and F2s, if their variety possesses those features. The majority of participants (75%) reported that they predominantly use Gilaki for mockery and joking. We can hypothesise that in this particular function they should probably be able to both differentiate (perceive) and produce certain vowels they associate with Gilaki. In other words, they should be able to produce certain Gilaki vowels at a conscious level yet fail to do so in an English production test, and make them different with respect to F1 and F2 values.

The observed pattern raises two important questions:

1. If the participants are able to produce tense and lax vowels with different F1 and F2 values in Gilaki, why isn't this differentiation translated to their English productions?

The answer to this question could be associated with the Feature Hypothesis (McAllister, Flege and Piske 2002) that an L2 phonological contrastive feature that is not "exploited" in the learners' L1 turns out to be more difficult to learn. For instance, the use of minimal pairs in English leads to meaning change, which is not the case in Standard Farsi. The question that needs further investigation is whether these minimal pairs still exist in present-day Gilaki.

I had initially claimed that Gilaki possesses minimal pairs similar to English. However, I believe further research is needed (along the lines of the analysis conducted in this thesis regarding English) to properly investigate whether Gilaki possesses those spectral distinctions between lax and tense vowels in the actual production of the current generation of Gilaki speakers. In other words, even if the assumption that Gilaki differentiates between lax and tense vowels phonologically, it remains to be tested whether the spectral characteristics of the lax and

tense vowels are well-exploited (de-exploited) in the test group. I leave this question for future research.

2. If the lack of spectral contrasts between the lax and tense vowels in L2 English is due to training or the dominance of Standard Farsi phonetics, can the test group benefit from a training that intentionally aims to transfer these features from their variety to their English productions? This could be another case for further investigations. For instance, an intervention study could be designed to test whether a reference and/or a comparison of English vowels to Gilaki vowels can be exploited to the benefit of this group of learners.

In the literature review, I referred to the book *Learner English* (Swan and Smith 2001), as an example and how cases of positive and negative transfer could be implemented in teaching methodologies. I also suggested that a probable positive transfer for the test group could be used as a facilitator in the training of Gilaki speakers. The results of the current study do not allow us to conclude that there is a sign of positive transfer from Gilaki vowels for their speakers in this study. Therefore, it remains unclear whether access to a larger vowel repertoire in a previously acquired language can become beneficial for Gilaki speakers acquiring L3 English with a comparable vowel system, considering the associated low prestige of the L1 variety. This certainly calls for more studies involving low prestige varieties in the future. I will get back to this once more in the conclusion section.

7.6 Crosslinguistic Influence

The two theories of CLI important for the discussion of my results are the Full Transfer Potential (proposed within the Linguistic Proximity Model; see Westergaard 2021 and Westergaard et al. 2017) and the Full Transfer (taken up in L3 Acquisition by the Typological Primacy Model; Rothman 2015, Rothman et al. 2019). As already reviewed in the background section, L2 acquisition has been traditionally associated with the FT, which is not surprising given that the L1 is the only language that can exert influence on a newly acquired L2. In fact, it is the L3 (Ln) acquisition that has led to theories such as – inter alia – the LPM or FTP. More and more evidence is now being brought to the table that is in line with the LPM/FTP rather than the FT, suggesting that all previously acquired languages interact and can be the source of CLI in additional language acquisition. However, there is still a scarcity of research that investigate the role of language distance in L3/Ln acquisition. It is thus unclear how similar or different the L3A situation is when the L1 and the L2 are two closely related varieties (CRV)

versus when the L1 and the L2 are typologically more distant languages. It is still more of a question whether speakers of two closely related varieties can be considered bilinguals or not. Consequently, should we consider one's learning process an L2 or L3 acquisition? It remains to be clarified by future research if the results of a study involving closely related varieties can be comparable to studies investigating typologically more distant ones, and how the results can be explained by the existing theories of L3A.

In addition, the terms monolingual, bilingual and/or multilingual need to be clarified when applied to situations involving CRVs. For instance, Lundquist and Vangsnes (2018) investigated two CRVs and found that one group could act like bilinguals, while the behaviour of the second group was more monolingual-like. Therefore, it can be concluded that it is important to investigate all languages involved in a study, and maybe even especially so when the multilingual situation involves closely-related varieties.

In the case of the present study, there is tangible scarcity of studies on Iranian varieties and Gilaki in particular. There is hardly any information whether Gilaks (people of the Guilan province) behave like monolinguals or bilinguals in certain situations when encountered with Standard Farsi in one way or another, neither do we know if these differences can promote bilingual behaviour in Gilaks when acquiring an L3 (Ln). Gilaki has at least three other properties that differentiates it from Standard Farsi, but which make the variety structurally closer to English.

Below I list several grammatical properties of Gilaki which could be used by future studies to design experiments testing if Gilaki speakers can indeed benefit from facilitative CLI of these properties to an L3 English and/or Norwegian.

1. The negative present continuous tense

Gilaki:

- Mæn daræm nemiram.
- I have.1sng Neg.go.1sng.
- I am NOT going.

Standard Farsi

- Mæn nemiram.
- I Neg.go.1sng.
- I do not go.

Although syntactically possible, the example above is considered “wrong language” by the Standard Farsi people who use a negative present simple tense instead. This, however, leads to two different meanings for a Gilak.

2. Use of the infinitive “dashtan” which means “have” to make present perfect tense (as an additional option).

Gilaki

1st person singular

Mæn æ filmə bideh dərəm

I this film.obj. see.PP. have.1sng

“I have seen this film”

Standard Farsi

1st person singular

Mæn film ra **dideham**

I film.OM..see.pp.1sng

“I have seen the film”

3. Putting the adjective before the noun

Gilaki:

Ite pile gav

/itə pɪlə gɑ:v/

Det. adj.enclitic obj

A big cow

Standard Farsi:

Yek gav-e bozorg

Det. obj.enclitic obj

A big cow

There are also other features in Gilaki that resemble other languages such as Norwegian. For instance:

- Using the possessive pronoun *shin* after the possessor

Gilaki:

Æ ketab Mehrsa shine

/æ kɪtʌb meɦɾʌ ʃiːn-ə/

This book.obj Mehrsa.PN.(F) shin.Poss e.to be. 3rd.sng

(This book Mehrsa 's is)

Denne boken er Mehrsa sin

Standard Farsi:

In ketab e Mehrsa ast

This book.obj e.enclitic Merhsa.PN(F) to be.3rd.sng

(This book Mehrsa 's is)

Denne boken er Mehrsa sin

Summing up, I suggest that these interesting non-trivial differences between two closely related varieties can be fruitfully exploited in future studies to investigate the source and degree of CLI in additional language acquisition. It is however important to conduct testing in all involved varieties to establish if linguistic properties in question are in fact employed in a bilingual way by speakers of closely related varieties, or if the minoritised variety is approximated to the more dominant one.

Turning back to the results of the present study, the results are generally inconclusive to the effect whether the test group act any differently from the control group, with respect to producing English vowels.

In other words, the results can be compatible with a proposal that both groups have transferred Standard Farsi phonological features into the English L3 acquisition, and have both the learned to differentiate between lax and tense vowels by means of duration.

On the other hand, taking the results of the attitude tests into consideration, the FT proposal may be challenged by the observation that we do not know whether the lack of CLI from Gilaki could be a result of an intentional suppression of Gilaki being a Low variety or not.

Furthermore, we do not know whether or not participants have received appropriate phonological input during their training or not. It may be the case that the learners are trying to approximate the pronunciation of their English teacher who may be L1 Farsi speaker.

However, having taken all the aspects into account, it could still be the case that there is full transfer of the phonological features from the Standard Farsi into English as the L3 for the Gilaki/test group, even if Gilaki possesses English-like minimal pairs of lax-tense vowels. How can this interpretation be accounted for?

1. Whether on account of a lower status, or intentional inhibition of the low variety, unacute instruction, etc, Gilaki users do not seem to show any sign of facilitative CLI from their L1. In other words, Gilaki users act like Farsi monolinguals in not differentiating between the F1 and F2 values of the English vowels.

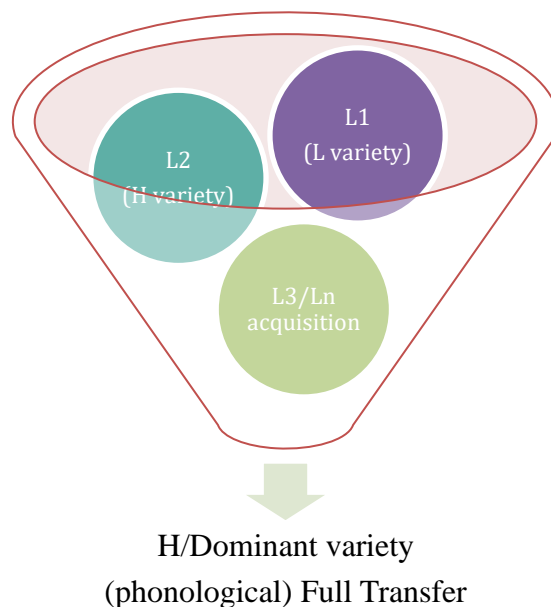


Figure 37 - Full Transfer from the dominant variety.

2. Gilaki and Standard Farsi resemble each other in regards to their vowels durations and spectral dimensions. Therefore, transfer may be associated with both Gilaki and Farsi. As a result, my assumption on the Gilaki minimal pairs is challenged, if not rejected.

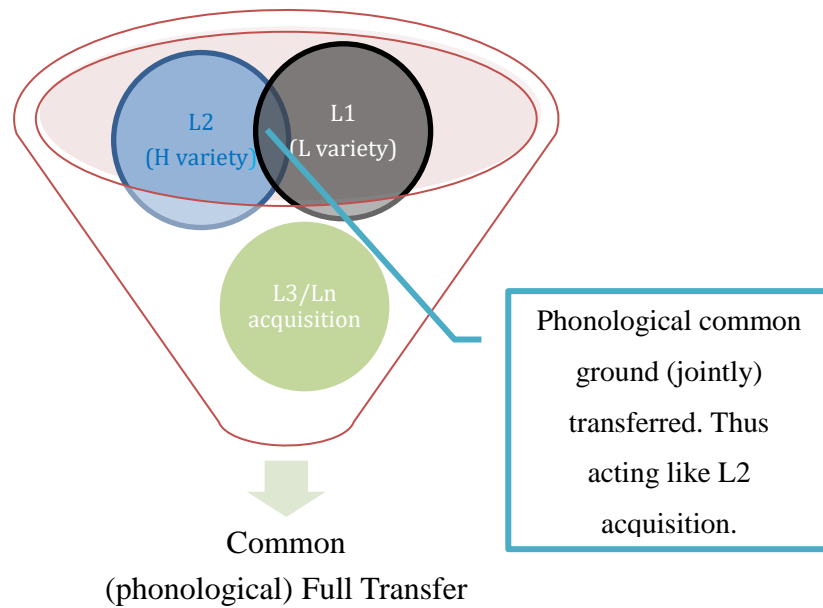


Figure 38 - English receiving phonological transfer as an L2 not L3.

It remains unclear whether the Gilaki group benefited from having the schwa sound /ə/ in their phonological repertoire, or if the Standard Farsi group benefited from more proficiency in English. As I discussed in the literature review, Archibald and Yousefi (2018) reported how Persian English learners benefited from a superficial dissimilarity to successfully convert the right-edge cc clusters into English onset clusters. It remains to be seen how abstract the structural similarity between languages has to be for the CLI to obtain. In other words, CLI with respect to a given property in L2 or L3 might be more abstract than a direct one-to-one correspondence, thus allowing facilitative CLI even from a structurally more distant variety with the potential to adapt to the new feature. Future investigation should clarify and ideally quantify the role of the degree and abstractness of linguistic/structural proximity in addition to the role of linguistic distance between the involved languages. I will look forward to such research.

8 Conclusion

The aim of this study was to investigate if the speakers of Gilaki could benefit from a larger vowel repertoire in their L1 and use it to facilitate their production of English vowels compared to the Standard Farsi speakers. The initial hypothesis was founded on account of an assumption that Gilaki vowel system is phonologically more marked than the Farsi one, and shares similar minimal pairs with English in terms of height & backness characteristics. The one existing previous study that I referred to stated the presence of such lax and tense vowels in Gilaki, and

also the existence of the /ə/ sound which is absent in Standard Farsi. It also refers to the /o/ sound as close-mid and back vowel that is absent in Farsi. However, it argued that the pair i and pair u only differ in their durations. Based on these assumptions, I proposed that there are spectral differences within these pairs. As a summary, I assumed that Gilaks should benefit from the more-markedness of their variety and the mentioned spectral dimension of the vowels to produce English vowels.

In addition, I also suggested that any positive transfer from Gilaki could be reinforced as a learning facilitator. To this end, I brought examples of similar classic works that have focused on negative and positive transfer from various L1s (or L2s). I also highlighted that I, by no means, promote the radical native-likeness paradigm. However, I argued that it might still be more practical to have a standard version model, since individuals still make their own modifications during the learning process and the fact that the proposed alternative paradigms, such as the Lingua Franca Core, have not yet succeeded to provide a practical substitute.

To investigate if my predictions were true, I designed a production test accompanied by a background questionnaire. The Gilaki group took an additional attitude test where they reflected on their relation to their local variety.

The production test was a set of 54 sentences that participants had to read spontaneously. Each sentence had a previously chosen target word embedded. There were four pairs of sound and a group of fillers. For every single sound, 6 sentences were designed. Each sentence, out of the total 54, only contained 1 target word that was subsequently extracted and analysed. The participants took the test offline but were given proper instructions at the beginning of the test.

The results revealed that both groups pronounced the words similarly, observing the duration to differentiate between the minimal pairs. This result did not corroborate my assumption that the Gilaki group should have benefited from their more-marked variety. Furthermore, it challenged my claim that Gilaki would transfer minimal pairs which have spectral differences. However, it is unclear if the absence of the hypothesised spectral dimensions was due to the dominance of Standard Farsi, quality of training, unlearned features or not existing of those features at all. Thus, the results are compatible with multiple theories, including full transfer of the Standard Farsi into English as the L3. Otherwise, there would have been (cumulative) transfer of the phonological features from both Gilaki and Standard Farsi into English.

However, I would rather interpret the full transfer here as **one** of the possible outcomes rather than an absolute and destined result. In other words, according to Westergaard's Full Transfer Potential (2019), full transfer is one of the possible outcomes, not the only one. So far, in this study, we can interpret the results as full transfer, should we consider English as an L3. Overall, to see the full transfer as one potential result provides us with an outlook in which the other contributing factors such as intentionality, personal preferences and H/L associations are also important. In fact, there is also a potential of a mixed transfer given a different combination of factors. The more we learn about L3A, the more obvious seems the need consider a broader range of factors that can influence the type and direction of transfer, factors that can potentially undergo changes to varying degrees throughout time, such as revitalisation movements that affect the social status of a language.

I reviewed several theories and models which propose how crosslinguistic influence obtains. I have argued that the outcome may be more complicated and difficult to predict when it comes to situations when the L1 and the L2 are two closely related varieties. However, I have also argued that the Full Transfer, as a rather mechanistic process, does not comply with a very nuanced and multifactorial picture of L3A that emerges from the recent studies, and that it is very challenging to integrate multiple contributing factors within a FT model.

Furthermore, I discussed that the two varieties of Gilaki and Standard Farsi co-exist in a rather uneasy and unbalanced social situation which can influence the activation of the two varieties in a formal education situation and affect the direction of and the reasons for transfer, especially when compared with two varieties with almost equal status. Subsequently, I referred to intentionality, arguing that such unbalanced social status is likely to induce intentional and (semi)conscious intervening of the speakers to hinder certain features from being transferred into the new language being acquired. The results from the attitude test indicated a likelihood of inhibition of Gilaki pronunciation (in particularly, vowels) by the Gilaks.

Overall, despite the (annoyingly) inconclusive results of the current thesis, I have learned a lot throughout the process of working on this project, starting from the ways how to design and conduct a psycholinguistic experiment, to practical skills of coding and analysing phonological data, to the importance of considering multiple factors in theorizing and modelling multilingual acquisition, ranging from the importance of a detailed analysis of similarities and differences between specific linguistic properties, the role of linguistic and extralinguistic factors in specific

combinations of the involved varieties. I can now see many possible ways to explore these questions in the future.

8.1 Extra Note

Earlier in the analysis section, where I reviewed the stages of the data analysis, I mentioned that the two stages were finally excluded from the analysis section. These last two stages were initially included to compare the results of the test and control group with the data from native speakers of English (i.e., **the reference group**) but were subsequently excluded for the following three reasons:

1. My original idea was to compare the results with the data from a study (Yang 1995) that compared 13 American vowels with 10 Korean vowels. The data did not provide any information on the vowel /ə/. Therefore, another data from a study on the “phonetics of the schwa vowel” (Flemming 2009) was added to the native speaker’s data. However, the second data only provided the means for F1 and F2. As a result, the reference data became inconsistent and uneven.
2. The first reference data (Yang 1995) was conducted on a different set of words. As I have already mentioned in the methodology section, the word itself, its position in the sentence and the form of the test can all affect the results of the production. Coupled with the previous reason, it seemed even further from an appropriate reference data.
3. The test group performed similarly to the control group, generally not showing any significant effect of spectral dimensions. In case of any opposite results, it would have been necessary to compare the test group’s performance with the one of the native speakers to see if the F1 and F2 values are similar to ones of the English. But given the lack of differences, such a comparison seemed superfluous.

Disregarding the results from the comparison with the reference data, I decided to leave this for further studies in which a population of native speakers would take the same test and read the same sentences. In addition, the present study itself has certain drawbacks that need to be compensated for (more) reliable results. For instance, the population that took part in the experiment was small, especially in the control group. The small number of participants grew even smaller because of random and unexpected errors in the files they sent back. Furthermore, the imbalance and highness of the proficiency level between the two groups, both of which were already at (upper) intermediate general. Overall, there seemed to be enough reasons to

postpone any further comparison and analysis to a time when more robust data are at hand to achieve more reliable results.

As mentioned earlier, there is a tangible scarcity of studies on the Gilaki language and especially on its relation to other varieties in terms of crosslinguistic influence. Therefore, there is some novelty in both the results and the present study itself.

The present study was initially designed to be the first stage of a three-phase-experiment which will be postponed to further studies in the future. The initial idea was to conduct a production task, as phase 1, to investigate if there are any significant traces of positive transfer in the experiment group, however the results of this study did not reveal anything in that regard. For phase 2, both the experiment and the control group will receive pronunciation instruction. During the instruction, examples from Gilaki will be used to see if participants find it practical and facilitative to associate English vowels with those of the Gilaki, and thus produce them more easily. This stage is supposed to be followed by asking the participants if they found the Gilaki examples helpful in finding the closest pronunciation in English, for the sounds they were assimilating with Standard Farsi. In addition, a second attitude test will be conducted to see if they have any new thinking about their local variety, especially in terms of its capabilities and its status.

Phase 3 will be another production task (post-test) to see if they succeed in maintaining the learnt instructions and try to use them consciously after a certain course of time.

Phase 1	Phase 2	Phase 3
Production Task 1 (pre-test)	Pronunciation Instruction	Production Task 2 (post-test)
Questionnaire	Attitude Test 2	Attitude Test 3
Attitude Test 1		

However, with the results of the current study, I have decided to conduct a preliminary experiment prior to any other. The results have raised questions about my assumption on the existence of English-like minimal pairs in Gilaki. Also, I have proposed that, according to the

results from the attitude tests, participants might have ‘unlearned’ certain Gilaki vowels, at least in production. Therefore, it is essential to investigate what vowels, with what durations and spectral dimensions, Gilaks use when speaking both in Gilaki and Standard Farsi. It will be both necessary and revealing if the vowel pool of the Gilaki population has undergone any changes during the last 50 years. The results of that study can pretty much determine if there will be any grounds for the 3-phase experiment to be conducted.

Works cited

Archibald, J., (2018) Advanced level phonology. In: Malovrh P and Benati A (eds) The Blackwell handbook of advanced proficiency in second language acquisition. Oxford: Blackwell, pp. 241–63.

Archibald, J., & Yousefi, M. (2018). The redeployment of marked L1 Persian codas in the acquisition of marked L2 English onsets: Redeployment as a transition theory. In *Conference on Central Asian Language and Linguistics. Bloomington, Indiana.*

BAS | web service interface | tools | speech | automatic segmentation and labelling | grapheme to phoneme conversion | syllabification | text-to-speech | Bavarian Archive for Speech Signals. (2021). Retrieved 31 October 2021, from <https://clarin.phonetik.uni-muenchen.de/BASWebServices/interface/WebMAUSBasic>

Balas, A., Kopečková, R., & Wrembel, M. (2019). Perception of rhotics by multilingual children. In *Proceedings of the 19th International Congress of Phonetic Sciences. Edited by Sasha Calhoun, Paola Escudero, Marija Tabain and Paul Warren. Melbourne: Australasian Speech Science and Technology Association* (pp. 3725-29).

Bardovi-Harlig, K., & Sprouse, R. A. (2017). Negative and positive transfer. In J. Liontas (Ed.), *The TESOL encyclopedia of English language teaching*. Hoboken, NJ: Wiley. doi.org/10.1002/9781118784235 .eelt0084

Bohn, O. S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. *Speech perception and linguistic experience: Issues in cross-language research*, 279-304. Jennifer Cabrelli Amaro & Magdalena Wrembel (2016) Investigating the acquisition of phonology in a third language – a state of the science and an outlook for the future, *International Journal of Multilingualism*, 13:4, 395-409, DOI: [10.1080/14790718.2016.1217601](https://doi.org/10.1080/14790718.2016.1217601)

Cabrelli Amaro, J., & Wrembel, M. (2016). Investigating the acquisition of phonology in a third language—a state of the science and an outlook for the future. *International Journal of Multilingualism*, 13(4), 395-409.

Cook, V. (2000). Comments on Vivian Cook's "Going beyond the Native Speaker in Language Teaching". The Author Responds. *TESOL Quarterly*, 34(2), 329. doi: 10.2307/3587957

Dauer, R., 2005. The Lingua Franca Core: A New Model for Pronunciation Instruction?. *TESOL Quarterly*, 39(3), p.543.

Derwing, T. (2003). What Do ESL Students Say About Their Accents?. *Canadian Modern Language Review*, 59(4), 547-567. doi: 10.3138/cmlr.59.4.547

Derwing, T. and Munro, M., 2008. Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), pp.476-490.

Eckman, F. (1977). MARKEDNESS AND THE CONTRASTIVE ANALYSIS HYPOTHESIS. *Language Learning*, 27(2), 315-330. doi: 10.1111/j.1467-1770.1977.tb00124.x

Edwards, J., & Zampini, M. (2008). *Phonology and second language acquisition*. Amsterdam: J. Benjamins Publishing.

Flemming, E. (2009). The phonetics of schwa vowels. *Phonological weakness in English*, 78-98.

Fuster, C., & Neuser, H. (2020). Exploring intentionality in lexical transfer. *International Journal of Multilingualism*, 17(4), 516-534.

Hammarberg, B., & Hammarberg, B. (2005). Re-setting the basis of articulation in the acquisition of new languages: A third-language case study. In B. Hufeisen & R. Fouser (Eds.), *Introductory readings in L3* (pp. 11–18). Tübingen: Stauffenburg Verlag.

Ionin, T. (2021). Property-by-property acquisition of a second language: What transfers, under what conditions, and how does recovery from transfer proceed? *Second Language Research*, 37(3), 435–439. <https://doi.org/10.1177/0267658320934523>

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

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

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979-1000.

Kuitert, S. R. (2013). *English Loanwords in Norwegian: A Loanword Processing Study in Young Speakers of Norwegian* (Master's thesis, Norges teknisk-naturvitenskapelige universitet, Det humanistiske fakultet, Institutt for moderne fremmedspråk).

Lee, B., Plonsky, L., & Saito, K. (2020). The effects of perception-vs. production-based pronunciation instruction. *System*, 88, 102185.

Lundquist, B., & Vangsnes, Ø. A. (2018). Language separation in bidialectal speakers: Evidence from eye tracking. *Frontiers in psychology*, 9, 1394.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics*, 30(2), 229-258.

Mildner, V., & Horga, D. (1999). Relations between second language proficiency and formant-defined vowel space. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1455-1458).

Munro, M. and Derwing, T., (1999). Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, 49, pp.285-310.

Munro, M. and Derwing, T., 1995. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. *Language Learning*, 45(1), pp.73-97.

Paraskiewicz, K. (2015). *English loanwords in New Persian*. FOLIA ORIENTALIA VOL. LII — 2015

Paul, L. (2010). Iranian Language Reform in the Twentieth Century: Did the First Farhangestān (1935-40) Succeed?. *Journal of Persianate Studies*, 3(1), 78-103.

Perry, J. (2005). p.97 *Lexical areas and semantic fields of Arabic loanwords in Persian and beyond*. – Csató É.Á. (ed. et al.). *Linguistic convergence and areal diffusion: case studies from Iranian, Semitic and Turkic*. London – New York: 97–109.

Podberesky, R., Deluty, R., & Feldstein, S. (1990). EVALUATIONS OF SPANISH- AND ORIENTAL-ACCENTED ENGLISH SPEAKERS. *Social Behavior And Personality: An International Journal*, 18(1), 53-63. doi: 10.2224/sbp.1990.18.1.53

Ramirez, N. F., & Kuhl, P. (2017). The brain science of bilingualism. *YC Young Children*, 72(2), 38-44.

Rastorgueva, V. S., Kerimova, A. A., Mamedzade, A. K., Pireiko, L. A., Edel'man, D. I., & Lockwood, R. M. (2012). *The Gilaki Language*. Acta Universitatis Upsaliensis.

Rastorgueva, V., & Lockwood, R. (2012). *The Gilaki language*. Uppsala: Acta Universitatis Upsaliensis.

Rothman J, González Alonso J, and Puig-Mayenco E (2019) Third language acquisition and linguistic transfer. Cambridge: Cambridge University Press. P 157.

Sato, C. J. (1991). Sociolinguistic variation and language attitudes in Hawaii. In J. Cheshire (Ed.), *English around the world: Sociolinguistic perspectives* (pp. 647–663). Cambridge: Cambridge University Press.

Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second language research*, 12(1), 40-72.

Seddighi, S. (2012). An account of Iranian EFL pronunciation errors through L1 transfer. *Iranian Journal of Applied Language Studies*, 2(2), 197-214.

Slabakova, R. (2021). Does Full Transfer Endure in L3A?. *Linguistic Approaches to Bilingualism*, 11(1), 96-102.

Smith, M. S., & Kellerman, E. (1986). Crosslinguistic influence in second language acquisition: An introduction. *Crosslinguistic Influence in Second Language Acquisition*, New York: Pergamon.

Sung, C. C. M. (2016). Does accent matter? Investigating the relationship between accent and identity in English as a lingua franca communication. *System*, 60, 55-65.

Trude, A. M., & Tokowicz, N. (2011). Negative transfer from Spanish and English to Portuguese pronunciation: The roles of inhibition and working memory. *Language Learning*, 61(1), 259-280.

Trudgill, P. (2005). Native-speaker segmental phonological models and the English lingua franca core. *English pronunciation models: A changing scene*, 77-98.

Unsworth, S. (2013). Current issues in multilingual first language acquisition. *Annual Review of Applied Linguistics*, 33, 21-50.

Wardhaugh, R. (2010). *An introduction to sociolinguistics*. Chichester, West Sussex, U.K: Wiley-Blackwell. Wardhaugh, R. (2010). *An introduction to sociolinguistics*. Chichester, West Sussex, U.K: Wiley-Blackwell.

Westergaard, M. (2021). L3 acquisition and crosslinguistic influence as co-activation: Response to commentaries on the keynote 'Microvariation in multilingual situations: The importance of property-by-property acquisition'. *Second Language Research*, 02676583211007897.

Westergaard, M. (2021). Microvariation in multilingual situations: The importance of property-by-property acquisition. *Second Language Research*, 37(3), 379-407.

Westergaard, M., Mitrofanova, N., Mykhaylyk, R., & Rodina, Y. (2016). Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal Of Bilingualism*, 21(6), 666-682. doi: 10.1177/1367006916648859

Weth, C., & Juffermans, K. (Eds.). (2018). *The tyranny of writing: Ideologies of the written word*. Bloomsbury Publishing. p 6.

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in systems neuroscience*, 7, 90.

Wrembel, M. (2015). In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology.

Wrembel, M. (2021). Multilingual acquisition property by property: A view from a wider perspective. *Second Language Research*, 0267658320935167.

Wrembel, M., Gut, U., Krzysik, I., Lewandowska, H., & Balas, A. (2019). ACQUISITION OF RHOTICS BY MULTILINGUAL CHILDREN. In *Proceedings of the 19th International Congress of Phonetic Sciences. Canberra: Australasian Speech Science and Technology Association* (pp. 954-58).

Wrembel, M., Marecka, M., & Kopečková, R. (2019). Extending perceptual assimilation model to L3 phonological acquisition. *International journal of multilingualism*, 16(4), 513-533.

Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of phonetics*, 24(2), 245-261.

Acoustic Phonetics: Formants. (2021). Retrieved 16 October 2021, from <https://home.cc.umanitoba.ca/~krussll/phonetics/acoustic/formants.html#:~:text=The%20frequency%20of%20the%20first%20formant%20is%20mostly%20determined%20by,frequency%20F1%20%3D%20high%20tongue%20body>).

2.2. Formants of Vowels – Phonetics and Phonology. (2021). Retrieved 16 October 2021, from https://corpus.eduhk.hk/english_pronunciation/index.php/2-2-formants-of-vowels/

Appendix A

The production task, background questionnaire, attitude test and R scripts:

https://osf.io/eydvs/?view_only=2fed9b67ccff47108f85eeca2109dd6e

Appendix B

Figures 34, 35 and 36 for the pair x duration, F1 and F2.

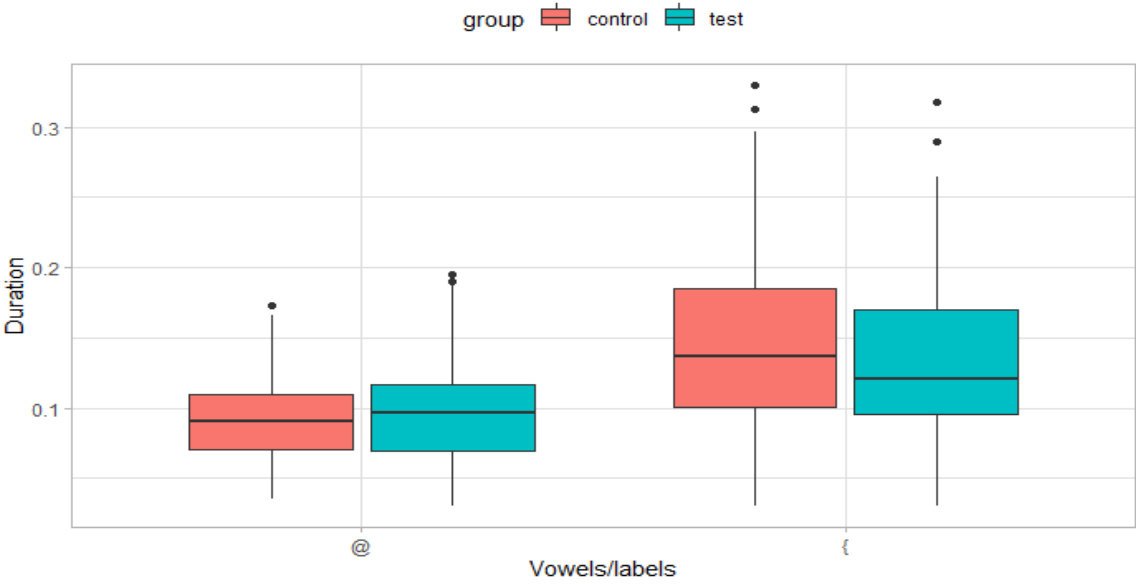


Figure 39 - Duration of the pair x vowels.

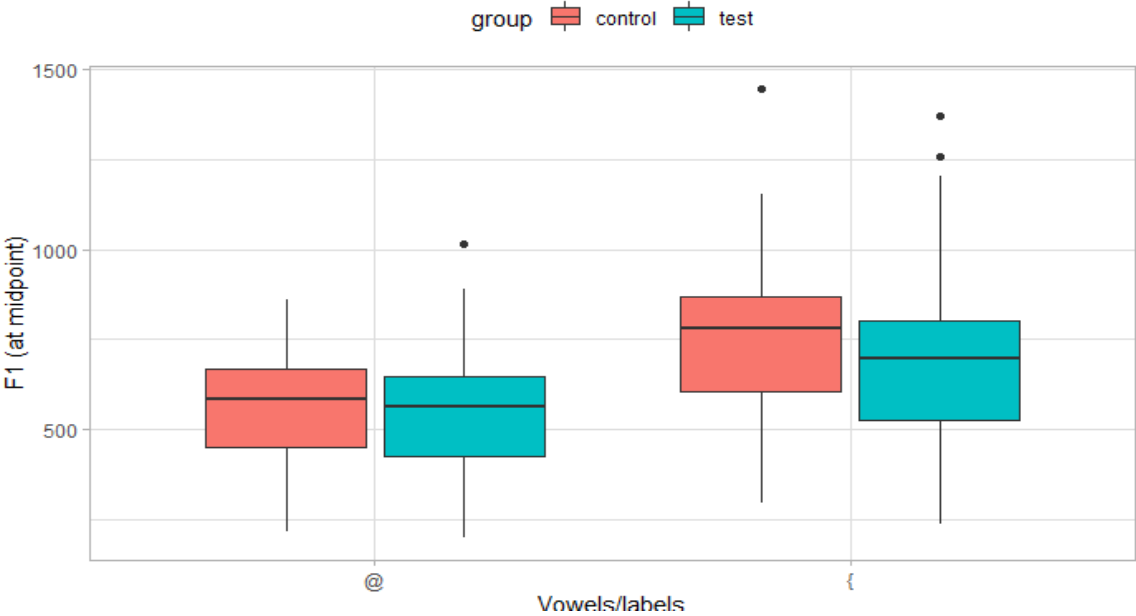


Figure 40 - F1 of the pair x vowels.

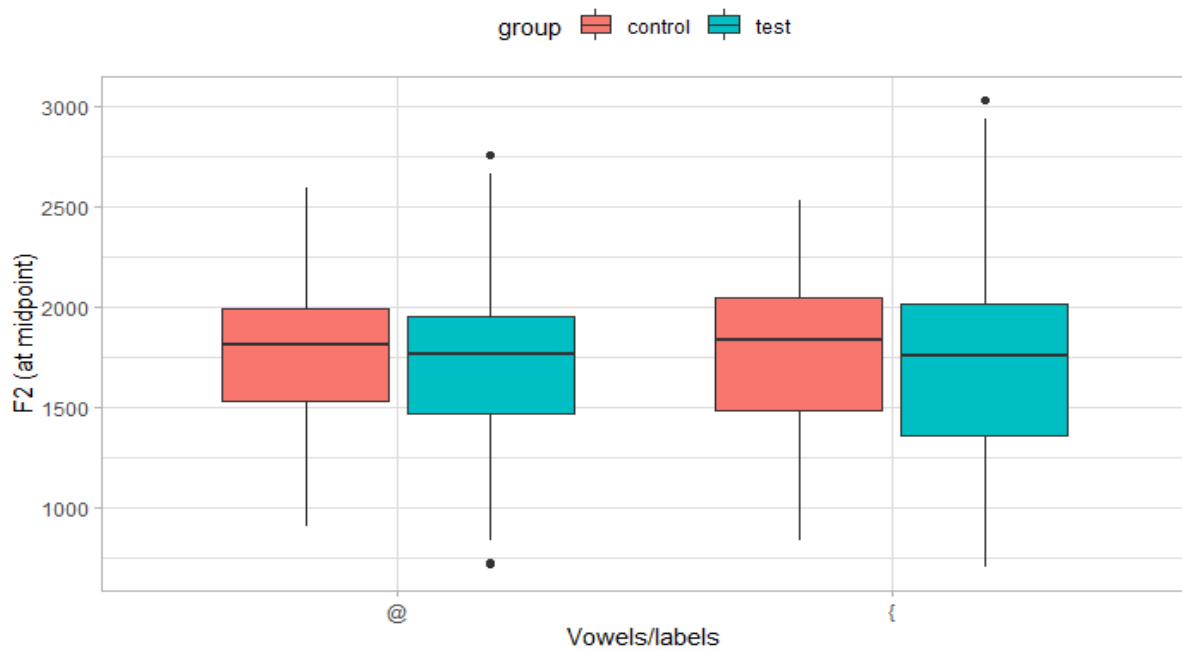


Figure 41 - F2 of the pair x vowels.

Appendix C

Analysis of the vowel pairs:

As a reminder, table 11 illustrates the pairs in both SAMPA and IPA.

	Vowel Pairs SAMPA – IPA			
	SAMPA	IPA	SAMPA	IPA
Pair i	I	ɪ	i	ɪ:
Pair a	V	ʌ	a	a:
Pair u	U	ʊ	u	u:
Pair x	@	ə	{	æ

Table 11 - vowel pairs in SAMPA and IPA

Vowel-Pair analysis

The proceeding models have been constructed with the same formula in terms of the dependent variable, mixed and random effects. Table 12 demonstrates the summary of the pair level

analysis. Models indicate results similar to the general analysis (Model 1), revealing that condition has significant effect (or is close to significant) at duration but turns out to be insignificant at F1 and F2. There are still exceptions such as pair x which shows very significant effect of condition at F1 (3.6) too, in addition to pair a having a pronounced condition effect at F2 (-3).

Duration

	Pair i		Pair a		Pair u		Pair x	
	Beta	t value	Beta	t value	Beta	t value	Beta	t value
Intercept	0.1	7.5	0.1	7.9	0.1	6.4	0.1	4.2
grouptest	-0.006	-0.7	-0.01	-1.4	0.001	0.1	0.004	0.4
Conditiontense	0.03	2.1	0.02	1.6	0.04	1.7	0.06	2.3
prof	-0.01	-2.3	-0.01	-3	-0.01	-2.3	-0.006	-1.1
Groupptest: conditiontense	-0.007	-0.96	0.001	0.1	0.006	0.5	-0.01	-1.8

F1

	Pair i		Pair a		Pair u		Pair x	
	Beta	t value	Beta	t value	Beta	t value	Beta	t value
Intercept	359.3	9.3	617.5	8.5	416.5	8.5	431.2	7.1
grouptest	36.1	1.7	-18.2	-0.4	5.6	0.2	7.4	0.2
Conditiontense	-38.34	-1.8	7.1	0.1	-4.05	-0.1	190.2	3.6
prof	19.6	1.4	33.5	1.3	10.7	0.6	53.7	2.9
Groupptest: conditiontense	3.1	0.2	9.5	0.2	-21.2	-1.1	-46.2	-0.983

F2

	Pair i		Pair a		Pair u		Pair x	
	Beta	t value	Beta	t value	Beta	t value	Beta	t value
Intercept	1315.8	6.1	1272.6	10.9	1174.42	6.8	1479.5	9.5
grouptest	122.6	1.05	11.9	0.1	32.1	0.3	7.8	0.08
Conditiontense	58.6	0.3	-237.1	-3.4	-21.6	-0.2	-11.3	-0.12
prof	256.3	3.6	86.09	2.09	45.23	0.7	110.8	2.08
Grouptest:	-36.1	-0.2	-27.5	-0.3	65.1	0.8	38.4	0.3
conditiontense								

Table 12 - Summary of the results of the fixed effects at vowel pair level analysis.

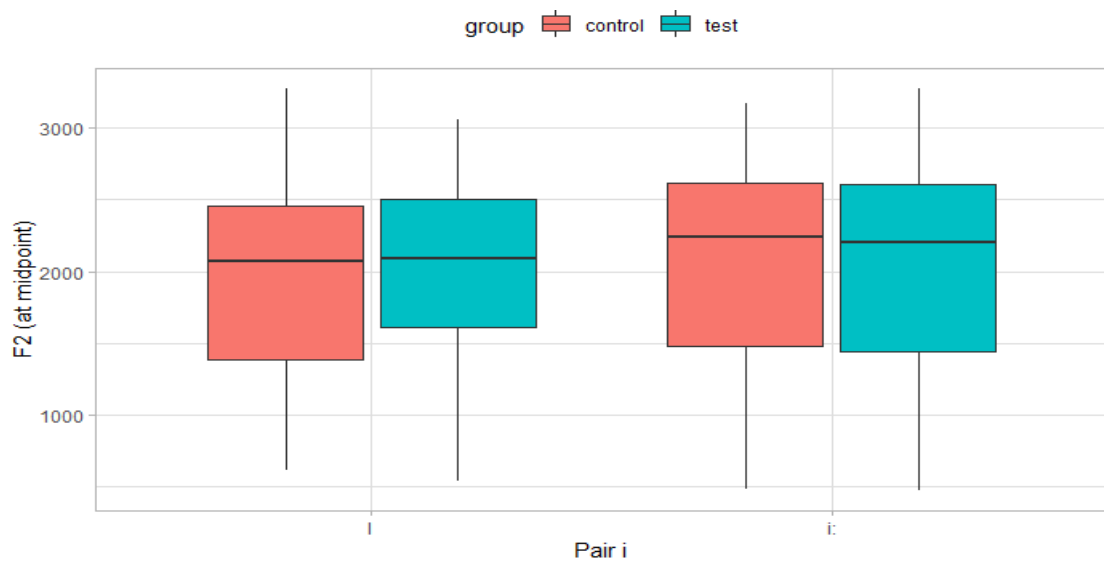
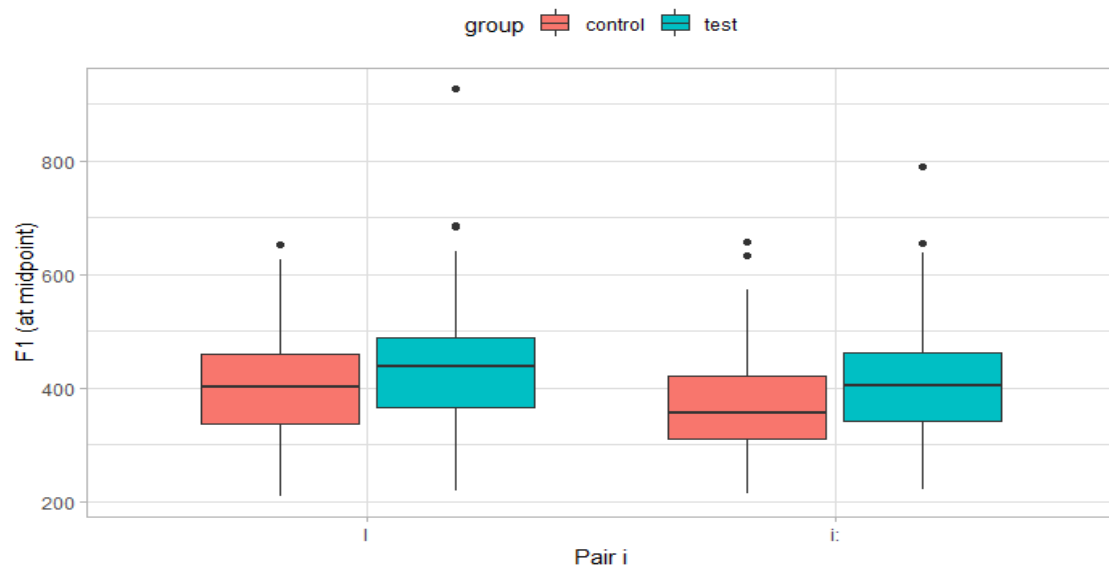
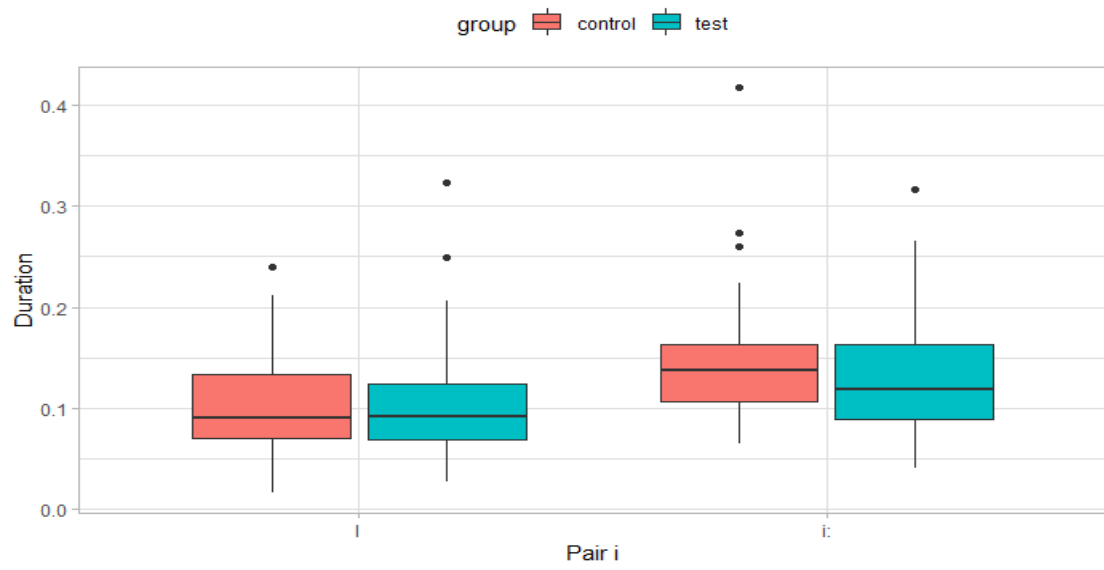
In terms of effect of proficiency, except for pair x, all three pairs demonstrate a significant effect at duration with ($\beta = -0.01$, t-value = -2.3), ($\beta = -0.01$, t-value = -3), ($\beta = -0.01$, t-value = -2.3) and ($\beta = -0.006$, t-value = -1.1) for pair i, pair a, pair u and pair x respectively. Neither of the pairs, except for pair x, show an effect of proficiency at their F1. However, there is again effect of proficiency at their F2, except for pair u with ($\beta = 256.3$, t-value = 3.6), ($\beta = 86.09$, t-value = 2.09), ($\beta = 45.23$, t-value = 0.7) and ($\beta = 110.8$, t-value = 2.08) for pair i, pair a, pair u and pair x respectively.

And finally, there are only two interactions which are at F1 pair i ($\beta = 3.1$, t-value = 0.2) and F1 pair a ($\beta = 9.5$, t-value = 0.2).

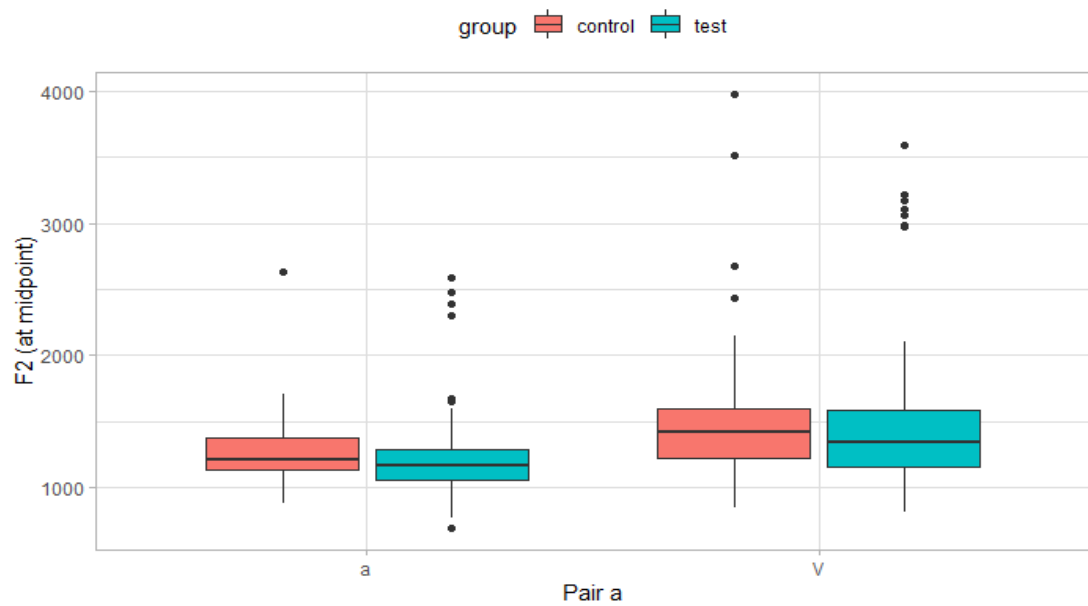
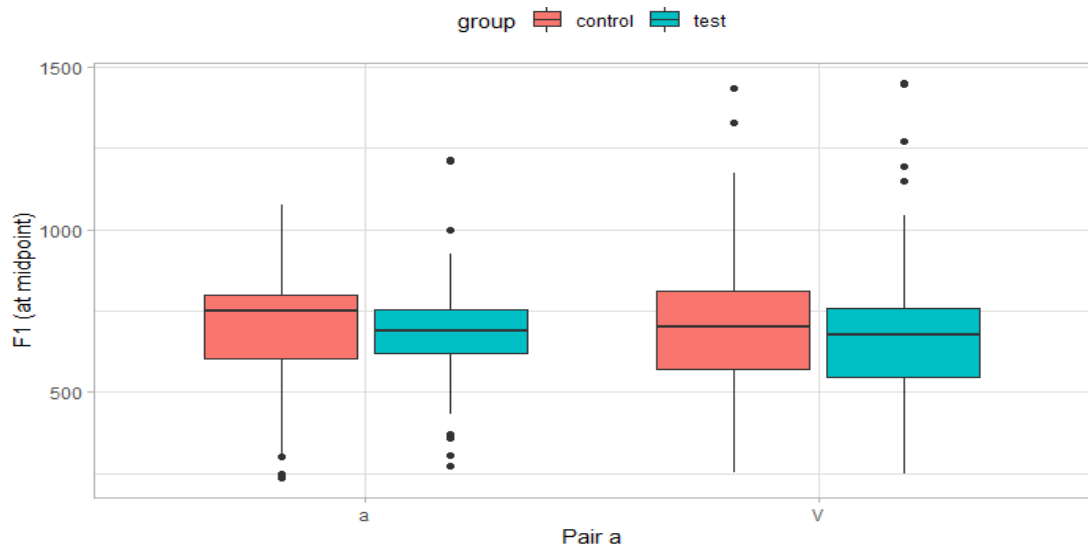
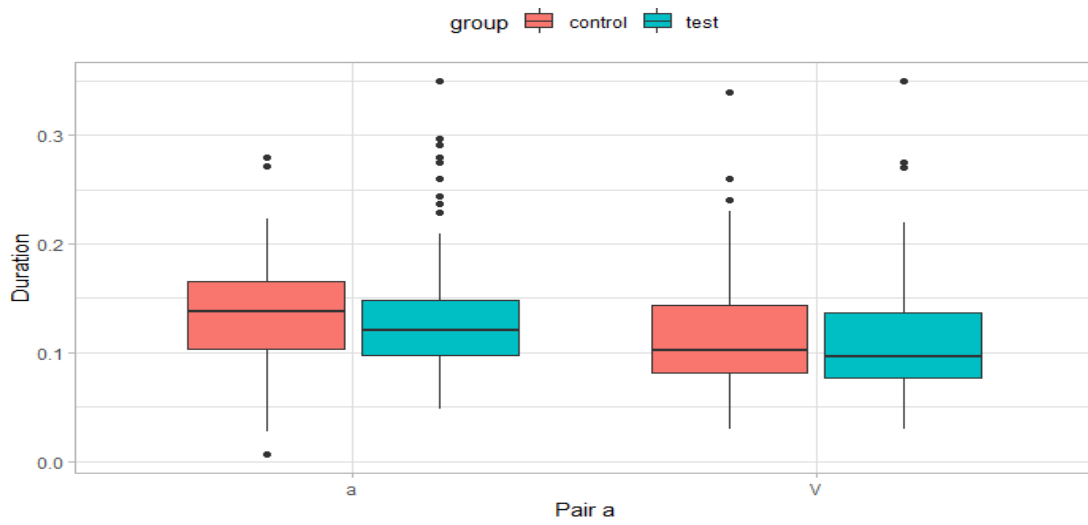
Appendix D

Figures for the vowel pairs

Pair I vowels



Pair a vowels



Pair u vowels

