Variation in loanword phonology: The case of /v/ and /tʃ/ in English loanwords into Saudi Arabic

Areej Alenazi

Doctor of Philosophy

University of York

Language & Linguistic Science

March 2023

Abstract

The thesis investigates variation in English loanwords into Saudi Arabic, triangulating across the results of three experimental tasks [non-word perception (oddity task), real loanword production, non-word production] and a short language attitudes survey. This multi-faceted approach allows exploration, for the first time, of the extent to which a range of competing factors - whose individual impact is known from prior research - interact and/or combine to influence how a novel target segment is realized in a loanword context, addressing: input modality (audio and/or written stimulus), participant's level of exposure to English, target consonant word position and participant's gender. The empirical focus is on two target consonants that are present in English but absent from most varieties of Arabic including the Saudi variety, but for which there is a near-equivalent in Arabic, and which have been reported to display variability in English loanwords into Arabic: /v/ mapping to [v]~[f], and /tʃ/ mapping to [tʃ]~[ʃ]. The reasons for this variability have not been clearly explored or explained in previous studies. Due to Covid-19, data was collected online with 67 participants, stratified by gender and expected level of exposure to English.

Results from the three tasks converge in suggesting that loanword adaptation is not a unitary phenomenon; various factors combine in different ways in their relative effect on adaptation for each of the two target contrasts, supporting a dynamic model which is sensitive to individual properties of both speaker-hearers

Table of Contents

List of Tables

List of Figures

Acknowledgements

First, words cannot express my gratitude and appreciation to my supervisor, Sam Hellmuth, for her invaluable guidance and support. My study journey had many challenges, Sam has been always supportive. For this, I am truly thankful.

I would also like to extend my appreciation to the members of my Thesis Advisory Panel, Eleanor Chodroff and Catherine Laing, who offered me guidance, feedback and suggestions along the way. I also would like to thank Vincent Hughes for helping me to sort out some troubles with statistical analyses.

Finally, I would be remiss in not mentioning my family. I would not have been able to complete this thesis without my husband, Adel, who has been a great support to me. For this, I am deeply grateful. I also would like to thank my children, Tarik, Ghaith, Deema and Reema, for their patience, and for putting up with me being a part-time mum for such a very long time! I also wish to thank my dad, sisters and brothers for their emotional support and encouragement.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

I declare that part of the results of my thesis was presented in these conferences:

Alenazi, A. & Hellmuth, S. (2021, April 4-6). *The Variable adaptation of /v/ and /tʃ/ in English loanwords in Arabic* [Presentation] Experimental Arabic Linguistics, Sharjah, UAE.

Alenazi, A. & Hellmuth, S. (2022, April 4-8). *Saudi Arabic Speakers' Perception of Non-native Contrasts* [Poster presentation]. Colloquium of the British Association of Academic Phoneticians, York, UK.

Alenazi, A. & Hellmuth, S. (2022). Variable adaptation of /v/ and /tʃ/ in English loanwords in Arabic. In A. Botinis (Ed.), *Proceedings of the 13th International Society of Experimental Linguistics*, *13* (pp. 9-12), Paris, France.

1 Introduction

1.1 Aims of this thesis

In research on loanword phonology, it has been observed that borrowed words are not always fully adapted to the native phonology of the borrowing language; adaptations can take different shapes when a borrowed word contains a novel sound. When one language borrows a word from another language two commonly observed outcomes are that a novel sound may be either preserved in its non-native form or substituted by a sound in the L1. As will be explained in the next chapter, we refer to the process of borrowing in which a novel sound is preserved as *importation* (Kang, 2011). That is, the novel sound that is retained in a loanword is referred to as *imported*. Studies of various languages have found variation in the production of novel sounds in loanwords (e.g., Adler 2006; Huang & Lin, 2016). For example, coronal stops do not exist in Hawaiian. When an English word is borrowed, a coronal stop /t/ is often replaced by a plosive /k/ as in 'trap' [kəla:pə], but it is sometimes allowed as in 'truck' [təlakə] (Adler, 2006). These *variable adaptations* raise questions about the factors that lead to preserving or substituting novel sounds. It is important to understand these factors because the frequent use of a novel sound can change the phonology of the borrowing language. For example, large-scale borrowing of Italian words by the speakers of Faetar, an isolated dialect spoken in southern Italy, led to nativising of geminate consonants (Nagy, 1994). In Korean, [s] and [ʃ] are allophones of a single phoneme; [ʃ] occurs in the context of the front vowel [i] or glide [j]. However, due to the influx of loanwords, the two sounds have become contrastive (e.g., 'show' [ʃo] vs. 'cow' [so]) (Lee, 2013).

The interest of this thesis as a whole is the research gap regarding a comprehensive understanding of the factors governing variable adaptation of novel structures in loanwords. More specifically, my interest lies in factors affecting the variable adaptation of novel sounds in Arabic. The popularity of English has led to a large number of English words being used in Arabic. Many of these words contain novel sounds. Prior studies on English loanwords into Arabic have already established the existence of variation in the production of novel sounds (e.g., Aloufi, 2016; Sa'aida, 2015). We know that when an English word containing a novel sound is borrowed into Arabic; the novel sound is either retained or replaced by an Arabic sound (e.g., Abu Guba, 2016). However, the reasons for this variability have not yet been clearly explored or explained in previous studies.

The present study seeks to fill this gap, and thereby intends to develop our understanding of mechanisms that govern variability in producing novel sounds in loanwords, in general. The present study will be different from these previous studies in both methodology and object of inquiry. First, most of the prior studies on English loanwords into Arabic sought to offer a grammatical explanation for the adaptation process with a particular attention to syllabic and morphological modifications. Additionally, most of these studies were based on real words collected from the Internet, newspapers, and magazines (e.g., Al-Athwary, 2017; Hafez, 1996). Lev-Ari and Peperkamp (2014) claim that studying the process of loanword adaptation experimentally can help in extending our understanding of the mechanisms underlying when, why and how existing sounds are changed or new ones are added. As will be summarised in 1.3, both real words and non-words were used in this study to elicit participants' productions of the target sounds in different conditions. Second, this study explores this phenomenon from different angles. Previous studies on loanwords, particularly in Arabic, did not provide a full explanation of loanword adaptation as a multifaceted process that is influenced by different factors. To the best of our knowledge, no study has yet explored a full range of factors affecting variable adaptations of novel sounds. The present study will examine the impact of input modality on the production of the target sounds in interaction with a range of other factors (level of exposure to the source language, gender and word position).

To better understand the design and measures involved in this study, it is important to discuss in detail our factors of interest. The first factor is input modality comprising auditory and orthographic information. As will be discussed, in detail, in Chapter 2, both perception and orthography play key roles in facilitating or impeding the production of novel structures in loanwords (e.g, Boersma & Hamann, 2009; Vendelin & Peperkamp, 2006). Numerous L2 speech production studies suggest that the modality in which a lexical item is initially or most frequently encountered can influence the production of novel sounds, especially if the sounds are realized differently, or do not occur, in the L1 (e.g., Bassetti & Atkinson, 2015; Bassetti, 2017). These studies suggest that production patterns of novel sounds may vary depending on how borrowers are first exposed to a source word. For this reason, the study aims to examine *the perception of the target novel sounds.* Additionally, the study aims to give insights into *the extent to which selection of the imported variant is influenced by input modality (auditory vs. orthographic)* and

develops our understanding of *the extent to which the availability of the two input modalities interacts.*

The second factor is level of exposure to the source language. Several studies on loanword phonology have provided evidence that level of exposure to the source language plays an important role in the production of novel sounds in loanwords (e.g., Kang & Schertz, 2021; Poplack, 2018). L2 research suggests that the effect of acoustic and orthographic information obtained from input modalities on L2 production can vary for individuals with different levels of L2 exposure: the more experience in the L2, the easier it becomes to perceive and produce novel structures (e.g., Kwon, 2017; Mok et al., 2018). That is, the production patterns may vary depending on borrowers' level of exposure to the source language. Consideration of language exposure is thus a central concern in this study. One of the aims of the study is to examine *the extent to which selection of the imported variant is influenced by level of exposure to the source language.* To do that, the study will involve individuals expected to have three different levels of English exposure: high, medium, and low.

The third factor is gender. The selection of one linguistic variant may depend on sociolinguistic factors. For example, women may tend to use more prestigious variants than men, or vice versa. Although an effect of gender has been reported in literature on sociolinguistic variation in Arabic, it is yet to be explored with regards to the production of non-native structures in loanwords. Previous sociolinguistic studies on Arabic showed that prestigious forms are retained more frequently by women than men (e.g., Al-muhanadi, 1991; Assiri, 2014; Omari & Van Herk, 2016). Sociolinguistic factors that are relevant to the L1 may also come into play when producing L2 forms (e.g., Adamson and Regan, 1991). For this reason, the final aim of this study is to examine *the extent to which selection of the imported variant is influenced by gender.* To test the effect of gender, an almost equal number of male and female participants will be involved in this study.

The fourth and final factor is word position. Some previous studies have shown that the position of a novel sound within a word in the source language may affect how it is produced (Huang & Lin, 2016; Kubozono, Itô & Mester, 2008). For this reason, it is important to consider the role of word position. One of the aims of this study is to examine *the extent to which selection of the imported variant is influenced by word position.* To do that, the target segments in the stimuli will be embedded in three different positions: initial, intervocalic, and final.

The investigation in this study will inform our understanding of how novel sounds are perceived and produced by borrowers of different genders and levels of exposure to the source language. In what follows, we will discuss the motivation for choosing Arabic as the borrowing language, explain the overall study design, and present the structure of this thesis.

1.2 Why English Loans into Arabic as a Test Case?

Arabic is investigated as the borrowing language in this study for two main reasons. First, as explained in the previous section, studies of English loanwords into Arabic where attention was paid mainly to the effect of native phonology, in fact showed variation between speakers in the production of novel sounds. (e.g., Al-Athwary, 2017; Abu Guba, 2016; Aloufi, 2016; Jarrah, 2013; Saaida, 2015). A research gap exists because the reasons for this variation have not yet been clearly explored or explained. Examining sources of individual variation is important to explain why and how sound changes occur. Previous literature on loanword phonology showed that the frequent use of a novel variant may cause a change in the L1 phonological system (e.g., Nagy, 1994). Second, Arabic is distinct from English in many ways. Arabic has more consonant phonemes than English, but there is still a considerable set of English consonant phonemes that do not exist in Arabic. In addition, the ways in which English phonemes are written are very different from the ways in which Arabic phonemes are written, and the correspondence between graphemes and phonemes is less direct in English than in Arabic. These phonological and orthographic differences may contribute significantly to the variable adaptation of novel sounds.

It is important to note here that, in this study, we focus specifically on speakers of Arabic in Saudi Arabia, a country where the interest in learning English has notably increased among young people, as it is seen as a passport to better education and employment. As will be discussed in Chapter 2, English enjoys a high status in Saudi Arabia. This high status and intense language contact are expected to invite more lexical borrowings to the native Arabic spoken in the country; hence, more novel structures. According to Haspelmath and Tadmor (2009), borrowing is motivated by the need to fill a lexical gap in the borrowing language, the prestige of the source language or by extensive exposure to the source language and its culture.

In this study, the focus on Saudi university students facilitates investigation of the impact of level of English exposure on the production of target novel sounds. In Saudi Arabia, university students' exposure to English can differ considerably based on their academic programs due to differences in the medium of instruction. For example, students studying

English typically have the most frequent exposure to the language in their classes, while students studying Arabic have the least exposure. Students studying medicine and science fit in the middle of the spectrum; medical and science-related curricula are usually delivered to students in English, whereas practical training (at hospitals, labs or elsewhere) is generally conducted in Arabic. These differences are exploited in this thesis such that choice of academic program is used as a proxy for level of exposure to English. Under Covid-19 conditions, we were not able to assess participants' proficiency in English because assessments happened in unsupervised settings.

Noting that novel sounds are variably adapted in Arabic, and due to the need for a comprehensive study on loanword phonology that explores different factors that affect variable adaptation of novel structures, this study will fill this gap by focusing specifically on novel /v/ and /tʃ/ found in English loanwords. Although Arabic is described as lacking /v/ and /tʃ/, the two sounds have each emerged in loanwords. Previous studies have revealed that target /v/ and /tʃ/ in loanwords are either preserved as [v] and [tʃ] or replaced by Arabic [f] and [ʃ], respectively (e.g., Abu Guba, 2016; Aloufi, 2016; Saaida, 2015). The present study focuses on these sounds for two reasons. The first reason is that there are many English loanwords into Arabic containing these two sounds (e.g., chocolate, chimpanzee, virus and vitamins). The second reason is their orthographic representations and expected perceptual mappings to Arabic. /tʃ/ exists in some Arabic dialects spoken in Gulf countries (Bahrain and Kuwait) while /v/ does not. Accordingly, /tʃ/ is expected to be easier to perceive than /v/; hence, auditory information may be more helpful to Arabic speakers in respect of [tʃ] than [v]. However, /tʃ/, /ʃ/ and /k/ can be represented by the same grapheme <ch> in English. Accordingly, the orthographic representation of /v/ is more transparent than /tʃ/; hence, orthographic information may be more helpful to Arabic speakers in respect of [v] than [tʃ].

1.3 Overall Study Design

In the present study, as summarised in Table 1.1, Saudi speakers of Arabic performed three tasks (one perception task and two production tasks), they also completed a short survey on attitudes towards the English language and American culture. Figure 1.1 shows the order of these tasks. The same set of participants performed all the tasks in the same order to allow triangulation across patterns of behaviour in each task.

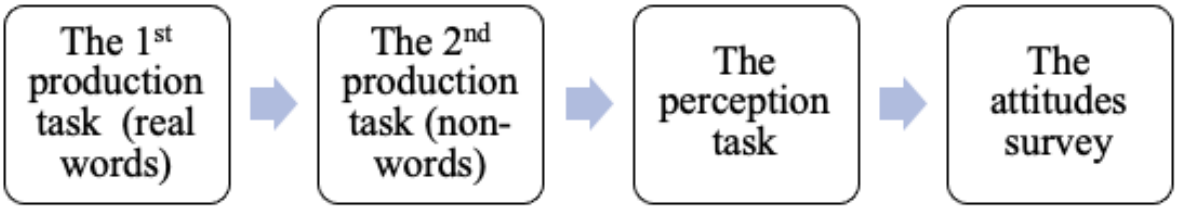


Figure 1.1: The order of tasks

During data collection, the production tasks preceded the perception task to prevent participants becoming familiar with auditory inputs, since the same non-words were used in the perception task as in the second production task. The perception task is however presented first in this thesis because its results were analysed first, and used as predictions for analysis of the production task. As noted above, both perception and production performance were tested in the same individuals to explore the relationship between the two.

The three tasks and attitudes survey were conducted online, because in-person data collection was made impossible by the strict ban on in-person gathering at the time of study due to the COVID-19 pandemic. The production and perception tasks were not performed on the same day. A link to the perception task was sent to participants two days after completion of the production tasks. The attitudes survey was then completed by all the participants after performing the three tasks.

Table 1.1: Production and perception tasks

Order	Task	Procedure	Stimuli
1	1st production task	Participants were asked to read sentences and fill in the blanks with target words represented in pictures.	Real words
2	2nd production task	Participants were tested in three conditions: aural-only (auditory inputs), written-only (orthographic inputs), and aural-written (auditory-orthographic inputs). In aural-only condition, they heard non-words pronounced by an English speaker. In written-only condition, they read non-words. In aural-written condition, they heard non-words while viewing their written forms.	Non-words
3	Perception task	Participants were asked to decide which stimulus is different, with the option to say that the three auditory stimuli are the same if they do not hear an odd stimulus.	Non-words

1.4 Structure of the Thesis

This thesis comprises seven chapters. This first chapter presents the aims of the thesis, discusses the motivation for choosing Arabic as the borrowing language, outlines the overall study design, and provides a roadmap for the following chapters.

Chapter 2 begins by reviewing the key terms in the relevant literature and providing a survey of instances, in various languages, in which novel sounds have been retained or substituted in loanwords. Additionally, the chapter reports the status of English in Saudi Arabia. The chapter also reviews previous studies concerned with factors affecting the variable production of novel sounds in loanwords in general, and discusses previous studies concerned specifically with English loanwords into Arabic. Furthermore, this chapter reviews background literature on how the input modality (acoustic and orthographic information) in which a lexical item is initially encountered can influence the production of L2 forms. The chapter concludes with a summary of how the previous literature provides important insights into the different factors that may influence the production of novel structures in loanwords which still leaves a research gap which this thesis aims to fill; namely, a detailed study of the interaction of these factors.

Chapter 3 presents the perception task that aimed to determine participants' discrimination accuracy of the target contrasts, /tʃ-ʃ/ and /v-f/. The Perceptual Assimilation Model (PAM) framework (Best, 1995) and the Perceptual Assimilation Model of Second Language Learning (PAM-L2) framework (Best and Tyler, 2007) were briefly introduced in this chapter, along with different examples of each assimilation type. The chapter provides detailed information about the participants, stimuli, and procedures employed in the task. The main findings of the task are then presented and discussed.

Chapter 4 presents the production task intended to elicit the production of target /v/ and /tʃ/ in non-words. The task considers the possible impact of input modality (i.e., aural-only, aural-written, and written-only) and a range of other factors (word position, gender, and level of English exposure) on participants' production of the two target sounds. The chapter spells out the stimuli, methods, the manual classification of target /v/ and /tʃ/ realizations and their acoustic measurements. The chapter also presents and discusses the main findings of the task.

Chapter 5 presents the production task intended to elicit the production of target /v/ and /tʃ/ in real words. As in the production task reported in Chapter 4, the production task in this

chapter considers the impact of level of English exposure, word position, and gender on the production of the two target sounds. The chapter also includes a presentation of the stimuli, methods, the acoustic measurements of target /v/ and /tʃ/ and results, as well as a discussion of the main findings of this task.

Chapter 6 presents a short survey that aimed to examine attitudes towards the English language and American culture. The analysis in this chapter is specifically intended to explore the potential effects of gender. The chapter presents, in detail, the design of the questionnaire, the results, and discussion of the main findings of the survey.

Finally, Chapter 7 offers a summary of the key findings in this thesis, discusses the impact of the different factors and their interaction, presents a k-means clustering analysis across all perception and production tasks, provides the conclusion of the entire study and ends by presenting the limitations of the study and suggestions for future research.

2 Background

This thesis aims to examine factors affecting variable adaptation of two target sounds, /v/ and /tʃ/, by Saudi Arabic speakers. Examples from various languages are discussed in this chapter to illustrate the range of possible sources of variability in the production of novel sounds in loanwords. Section 2.1 reviews definitions of terms that are commonly used in this thesis. Section 2.2 provides an overview of the status of English in Saudi Arabia. Section 2.3 discusses the internal factors and external factors that affect the production of novel sounds in loanwords. Section 2.4 offers a brief review of studies conducted on English loanwords into Arabic. Section 2.5 discusses prior research on the impact of input modality on L2 production. Section 2.6 summarises and discusses how the literature review on loanwords helped to shape the research questions and the methodology of this study.

2.1 Definitions of Terms

This section defines and discusses terms that are commonly used in this thesis. Subsection 2.1.1 addresses lexical borrowing and loanword criteria while subsection 2.1.2 discusses importation and adaptation and provides a survey of their incidence in various languages.

2.1.1 Lexical borrowing

Lexical borrowing refers to the process by which a word is transferred from one language into the lexicon of another language as a result of contact (Winford, 2010). Poplack (2018) defines borrowing as "the process of transferring (Clyne, 2003) or incorporating (Thomason & Kaufman, 1988) lexical items originating from one language into discourse of another" (p.6). The language that provides the lexical item is identified as the *source language*, whereas the language into which the lexical item is hosted is called the *borrowing language* (Poplack, 2018). It is generally agreed that nouns are the most commonly borrowed category. Gardner-Chloros (2009) indicates that nouns, syntactically, have fewer restrictions than other word-classes and that they are more accessible to individuals with any degree of bilingualism in the language from which loans are taken. Borrowing can be motivated by the need to fill a gap in the lexicon of the borrowing language, by the prestigious status of the source language or by extensive exposure to the source language and its culture (e.g., Haspelmath & Tadmor, 2009). Additionally, it has been suggested that borrowing can be driven by a desire to express affiliation with the source

language (Lev-Ari & Peperkamp, 2014, Paradis & LaCharite; 2012) or to avoid homonymy or taboo words (Winford, 2003).

Myers-Scotton (2006) distinguishes between types of borrowings: *cultural borrowing* and *core borrowing*. Cultural borrowing involves foreign words for objects or concepts that are new to the borrowers' culture, such as the borrowed English word 'fax' into Arabic. By contrast, core borrowing refers to foreign words that are borrowed despite the fact that equivalent words exist in the borrowing language. For example, the English word *job* is borrowed into French despite the fact that there is a French word, *boulot*, with the same meaning. Another classification is provided by Coetsem (1988) who distinguishes between *borrowing* and *imposition* based on linguistic dominance. The direction of transfer is often from the source language to the borrowing language by an agent who is linguistically dominant in a language in which he/she is more proficient (the source language or the borrowing language). *Borrowing* involves the transferring of a foreign word from the source language by a speaker who is linguistically dominant in the borrowing language, as in the case of an Arabic speaker using an English word while speaking Arabic. On the other hand, *imposition* involves the introduction of a foreign word into the borrowing language by a speaker who is linguistically dominant in the source language, as in the case of an English speaker using an English word while speaking Arabic.

To avoid any confusion, it is also important to clarify the distinction between two language contact phenomena: *borrowing* and *code-switching*. In an effort to disambiguate the two, Poplack (2018) argues that *borrowed words* are clothed with the morpho-syntactic structure of the borrowing language, while single words or multiword sequences that do not show syntactic, morphological or phonological integrations are identified as *code-switching* (Poplack, 2018). However, Gardner-Chloros (2009) argues that *borrowing* can be distinguished from *code-switching* only in diachronic terms; non-native words start as code-switches and finally end up as loanwords. According to Gardner-Chloros (2009), it is a misconception to view synchronic *borrowing* and *code-switching* as distinct processes. She claims that morpho-syntactic integration is not a reliable way to distinguish between borrowing and code switching because some individuals can morphologically integrate both. However, Poplack (2018) defines foreign words that show syntactic, morphological or phonological integrations, but which people use

infrequently, as a separate category of *nonce borrowing*. Poplack's criteria are shown in the table below.

Table 2.1: Poplack's (2018) criteria of loanwords, nonce borrowing and code switching

	Phonological Integration	Morphological Integration	Syntactic Integration	High Frequency
Loanwords	✓	✓	✓	✓
Nonce borrowing	✓	✓	✓	×
Code-switching	✓	×	×	×

The production task in Chapter 5 is primarily interested in loanwords that meet Poplack's criteria, as many words have been borrowed very recently from English, especially in the fields of fashion, science and technology, by either bilinguals or monolinguals who may have heard or read the words in the media. It is difficult to use a diachronic criterion (Gardner-Chloros, 2009) to classify the recently borrowed words. The selected borrowed words in the production task in Chapter 5 show morphosyntactic integration and are frequently used by individuals across the community (see 5.3).

2.1.2 Loan Phonology: importation vs. adaptation

The existing literature on loan phonology suggests that novel segments in a loanword may be either retained or replaced. That is, borrowers may change or preserve a novel segment in loanwords. Haugen (1950) provides one of the earliest discussions of borrowing, distinguishing between importation and adaptation. *Importation* is a process in which a novel segment from the source language is retained (Haugen,1950). Bator (2010) notes "importation is the process of the adaptation of the thing to be borrowed in an unchanged way, as it is in the original language" (p.40). Kang (2011) defines *importation* as "a situation where a structure not attested in native phonology is exceptionally allowed in loanwords" (p. 2260). Previous studies have demonstrated that the importation of a non-native segment is associated with bilingualism, attitudes and prestige (e.g., Lev Ari et al, 2014; Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2012). In *adaptation*, on the other hand, a segment from the source language is replaced with a segment in the borrowing language (Haugen, 1950). Kang (2011) defines *adaptation* as a situation where a non-native structure is changed, so as to make it more similar to the borrowing language. It has

been suggested that adaptation rates are higher in loanwords that are frequently used over a long period of time (Haugen, 1950; Poplack & Sankoff 1984). An enormous amount of research has supported the view that adaptation takes place when a novel segment does not exist in the borrowing language (e.g., Burenhult, 2001; Kim, 2009), but unnecessary adaptation may also occur when the same segment exists in the two languages (e.g., Yang & Golston, 2001; Kang, 2011). For example, /ʒ/ in French loanwords in Hmong is substituted with /j/ despite the fact that /ʒ/ exists in Hmong (Yang & Golston, 2001). According to Peperkamp (2004), unnecessary adaptations may occur as a result of the phonetic decoding process that occurs during speech perception.

As an example of adaptation and importation in English loanwords into Arabic, consider the adaptation of /v/ and /tʃ/. English /v/ and /tʃ/ are sometimes imported (e.g., 'cover' [kavar] and 'chips' [tʃibs]) and sometimes replaced by /f/ and /ʃ/, respectively (e.g., 'microwave' [maykru'weef] and 'chimpanzee' [ʃam'baanzi]) (Abu Guba, 2016; Aloufi, 2016; Saaida, 2015). Some instances of segmental importation and adaptation in other languages are listed below.

(1) Importation

a) Russian (Holden, 1976): palatalised consonants always occur before the front vowel /e/, but non-palatalised consonants are sometimes tolerated in loanwords, as in 'hotel' /hotɛl/ –> /otel/.
b) Japanese (Itô & Mester, 1995): Japanese has three lexical strata—native, Sino-Japanese (old borrowings from Chinese) and foreign loanwords. Coronal stops are prohibited before /i/ in native Japanese words, but are common in foreign loanwords, as in 'party' [paatti].
c) Hawaiian (Adler, 2006): /t/ was imported in the English word 'truck' [təlakə] although coronal stops do not exist in the native phonology.
d) Mexicano (Lev-Ari et al., 2014): Spanish segments / d b f g ɲ x r/ are more likely to be imported in domains in which Spanish has more prestige than Mexicano, such as in technology and education, as in the importation of word-initial /f/ in 'fotografia' –> [fotoɣaβjaʔ].
e) Hebrew (Lev-Ari et al., 2014): Hebrew speakers import novel segment /w/ in some loanwords, as in the Arabic interjection 'walla' [walla].

f) Quebec French (Hsu & Jesney, 2017): the affricate /tʃ/ is imported in words borrowed from English even though it does not exist in the L1 inventory, as in 'batch' [batʃ].

(2) Adaptation

a) Jahai (Burenhult, 2001): in words borrowed from Malay, /k/ is replaced with /ʔ/ in word final position, as in 'lake' /tasik/ –> [taseʔ].

b) Burmese (Chang, 2009): in words borrowed from English, labiodental fricatives /v/ and /f/ are substituted with /pʰ/, as in 'café' /kæfeɪ/ –> [kəpʰi].

c) Some dialects of Quebec French (Paradis & LaCharité, 2008): in words borrowed from English, the postalveolar fricative /ʃ/ is replaced with /h/, as in 'shop' /ʃɔp/ –> [hɔp].

d) Thai (Kenstowicz & Suchato, 2006): in words borrowed from English, /v/ is replaced with /w/ in onset position and with /p/ in coda position, as in 'conservative' [khɔ̄nsə̄ə̄wēētip].

e) Korean (Kang, 2008): in words borrowed from English in the 1930s, /b d ɡ/ are replaced with the tense stops / p' t' ɡ'/ in word-initial position, as in 'double' [t'əpɨl'], and with the lax stops /p t k/ elsewhere, as in 'board' [p'otɨ].

f) Korean (Kim, 2009): in words borrowed from English, /z/ is replaced with /ts/, as in 'zoom' [tsum].

g) Arabic (Abu Guba, 2016): in words borrowed from English, /p/ is replaced with /b/, as in 'clip' [klibb] and 'piano' [byanu:].

2.2 English in Saudi Arabia

Before turning to the wider literature on loanword phonology, this section presents a brief historical overview of the status of English in Saudi Arabia as this study focuses specifically on speakers of Arabic in Saudi Arabia.

There is no consensus on the exact date when English was introduced into Saudi education. For example, Niblock (2006) claimed that English was introduced in education in 1932 while Al-seghyer (2014) believes that it was introduced after the establishment of the General Directorate of Education in 1924. After the discovery of oil, English attained a high status in the economic sector, which in turn had a great influence on English teaching in the country (Al-Johani, 2009). According to Al-Braik (2007), in 1978, 90% of workers in certain business domains, such as hospitals and restaurants, were foreigners. For this reason, one of the main reasons to teach English in the country at the time was to enable Saudis to communicate

satisfactorily with foreign workers. However, despite this early introduction in education and the economy, English, for decades, received little attention from Saudis. There was little interest in learning English because it was seen as a threat to the native language, culture, and even religion (Alsharhani, 2016).

In the early 2000s, significant efforts were made by the Saudi government to promote English, which resulted in significant changes in the status of English in Saudi Arabia (Elyas, 2008). The Ministry of Education has funded hundreds of thousands of Saudis to learn English and pursue their studies in English-speaking countries, such as the UK, USA, Australia, and New Zealand. Additionally, English has become a compulsory subject in both private and public schools and is used as the medium of instruction to teach students in medical and engineering departments (Faruk, 2013).

Today, there is no doubt that English enjoys a high status within Saudi society. English is now necessary if one desires a career in a high-status company or to obtain a scholarship. English is the medium of communication and training in major companies in the country, such as Aramco and Saudi Airlines (Mahboob & Elyas, 2014). Scholarships in English- speaking countries are now contingent upon getting unconditional admission which requires fulfilling the necessary English language requirement. That is, in order to get a scholarship, one must achieve a required language score in IELTS or TOEFL (or equivalents) to apply to foreign universities.

2.3 Factors Affecting the Variable Production of Novel Sounds in Loanwords

In this section, I discuss in detail internal factors such as native phonology, perception and orthography (subsection 2.3.1) and external factors such as bilingualism, attitudes and prestige (subsection 2.3.2) that could possibly affect the production of non-novel sounds in loanwords.

2.3.1 Internal Factors

Numerous studies have yielded important insights into how variable production of novels sounds is conditioned by internal linguistic factors, such as the L1 (e.g., Huang & Lin, 2016; Kubozono, Itô & Mester, 2009), perception (e.g., Boersma & Hamann, 2009) and orthography (e.g., Hamann & Colombo, 2017; Kang, 2009; Vendelin & Peperkamp, 2006).

Considerable research attention has been paid to the impact of the L1 on the production of novel sounds using the framework of constraint based models, such as the Theory of Constraints and Repairs Strategies (Paradis, 1987) and Optimality Theory (Prince and

Smolensky, 1993). Optimality Theory (OT) (Prince & Smolensky, 1993) posits that the surface form is optimal if it incurs the fewest violations of a set of ranked constraints. Speakers may use phonological processes, such as epenthesis or deletion, to satisfy high-ranked constraints in their native language. The constraints are divided into two types: markedness constraints and faithfulness constraints. The notion of markedness is grounded on universal principles of speech perception and production. For example, consonant clusters are marked because they are, cross-linguistically, less common. Markedness constraints penalise surface forms (outputs) with specific properties. For example, the constraint "OCP" (Obligatory Contour Principle) is violated by any sequence of two or more adjacent identical elements (e.g., adjacent instances of a particular phonological feature); another example is the constraint "*COMPLEX" which is violated by consonant clusters. On the other hand, faithfulness constraints require outputs to match inputs. For example, the faithfulness constraint "DEP" (i.e., so-called because of the 'dependency' between inputs and outputs) penalises the insertion of material not in the input. Another example is the constraint "MAX" which prohibits deletion (i.e., outputs should be maximal). The optimal surface form is determined by the interaction of faithfulness constraints and markedness constraints. Marked properties are observed in surface forms in cases where faithfulness constraints outrank markedness constraints.

Theory of Constraints and Repairs Strategies (TCRS) is also a constraint-based approach that was devised to account for loanword adaptations. A repair is defined as a change to the form of a word to meet the phonological requirements of the borrowing language. TCRS states that a repair must be done when a constraint is violated. According to TCRS, preserving or deleting a segment is determined by how costly a repair strategy will be. TCRS claims that segmental information should be maximally preserved (Preservation Principle). However, deletion, as a repair strategy, takes place when preserving segmental information exceeds the threshold. TCRS claims that there is a limit on the number of permitted repairs, known as the "Threshold Principle", with the limit set at two mechanisms. This threshold is argued to explain why one repair strategy is observed rather than another. For example, if the borrowing language does not permit consonant clusters, the illicit cluster can be adapted either by epenthesis (inserting a vowel to break up the cluster) or deletion (deleting one consonant in the cluster). In Marshallese, which prohibits consonant clusters, the final cluster in the English loanword ‘pump’ was adapted by deletion [bam]. According to the Preservation Principle, the final cluster can be easily adapted

by epenthesis. However, the sound /p/ should be replaced by a native sound because it does not exist in Marshallese which raises the cost. Therefore, the final cluster was adapted by deletion rather than epenthesis (Brasington, 1997).

LaCharité and Paradis (2005) indicated that segmental mapping is based on the phonological categories of the L1 and L2. That is, the target L2 sound is not produced as the L1 sound that is acoustically identical or closet but as a sound that is phonologically in the same category (in terms of features). Kubozono et al. (2009) examined variation in the production of English words borrowed into Japanese. It was found that gemination in English loanwords is sometimes tolerated, e.g., 'pick' [pikku], and sometimes blocked, e.g., 'peak' [piiku]. The researchers indicated that these patterns are compatible with the fact that Japanese phonology favours heavy-light syllable sequences more than light-heavy syllable sequences. In order to improve the prosodic structure, gemination is allowed to occur in the first syllable only if the coda consonant is preceded by a short vowel. Huang and Lin (2016) investigated the variable production of the English coda nasal /m/ in Modern Standard Mandarin. In coda position, /m/ was adapted variably, either by inserting a vowel, e.g., 'rum' [lɑŋ.muː] or by changing its place of articulation, e.g., 'camp' [k^han.puː]. Huang and Lin (2016) demonstrate that these patterns are attributed to the Mandarin phonology because only /n/ and /ŋ/ are allowed in coda position. Thus, a vowel is often inserted after /m/. However, they noted that /m/ cannot be preserved by inserting a vowel if it is followed by an obstruent that has the same place of articulation as in 'camp' [k^han.puː].

In addition, perception is also reported to influence the likelihood of adaptation. Dupoux et al. (1999) provided one of the earliest discussions of the effect of the borrowing language on the perception of non-native structures. An experimental study was carried out by Dupoux et al. (1999) to examine how Japanese and French speakers perceive consonant clusters. The findings reported that Japanese listeners had difficulty in distinguishing between VCCV (e.g., ebzo) and VCuCV (ebuzo) due to the fact that consonant clusters are prohibited in Japanese. On the other hand, French listeners had no difficulty in identifying the illusory vowel (i.e., an epenthetic vowel within consonant clusters) as French allows consonant clusters.

To account for variable production of English word final stops in Korean, Boersma and Hamann (2009) proposed an Optimality-Theoretic model that involves both perception and production. As the researchers noted, in Korean, stops are of three types: lax, aspirated and fortis.

All three types are produced as lax in coda position. In English loanwords, vowel insertion often occurs after final stops; that is, codas are perceived as onsets. However, there is variability in the production of English loanwords with the final velar stops /g/ and /k/; a vowel is variably inserted after these two stops. Boersma and Hamann (2009) attributed this variability to perception. They demonstrate that vowel insertion occurs because Korean listeners interpret the presence of a release burst as a vowel. Acoustic variability in the burst release influences whether an epenthetic vowel appears after a velar stop: vowel insertion does not occur if there is no audible velar release. Final velar stops are often pronounced without an audible release after lax vowels. Thus, 'spike', for example, was produced as [si.pi.ki], while 'quick' was produced as [kʰwik]. This finding is consistent with another study conducted by Peperkamp et al. (2008) in which they examined the production of English and French word final /n/ in Japanese. In English loanwords, /n/ was adapted as a moraic nasal consonant, as in 'pen'. In contrast, in French loanwords, /n/ was always followed by an epenthetic vowel. The strong release of the French word final /n/ was perceived as [ɯ] by Japanese speakers as in 'customs' douane −> duannu.

Kim (2021) investigated vowel insertion following word-final stops in English words borrowed into Korean. Of particular interest was whether release, voicing, place of final stops, tenseness of the vowel preceding final stops, final stress, and word length trigger vowel insertion. Thirty Korean native speakers performed a similarity judgement task. The task involved 132 non-words (84 monosyllabic, 24 disyllabic and 24 trisyllabic words). Each non-word had three forms: an English form (e.g., [kʰɛt]) and two Korean forms (e.g., [kʰɛtˀ] and [kʰɛtʰɨ]). The Korean participants were asked to listen to the three forms and to determine whether the second form (the English form) sounds more similar to the first form or the third form (the Korean forms). Kim (2021) found that release, place of stops and word length were significant factors. The English form was more likely to be heard as CVCV than CVC when the English form was monosyllabic, and the final stop was released and dorsal or coronal.

In contrast to the previous studies, however, a recent study by Martin et al. (2022) examined the perception and production of the /k-g/ contrast by Dutch speakers. In Dutch, /g/ and /k/ are not contrastive. However, this contrast has emerged recently in Dutch due to the heavy borrowing from English (e.g., 'mango' vs. 'manco'). Of particular interest was whether Dutch speakers' individual productions of the imported sound /g/ would correlate with their ability to discriminate English /g/ from their L1 /k/ in perception. Dutch /k/ is produced with a

shorter VOT than that of English /k/. In initial position, English /g/ is often produced with a short-lag VOT which makes it very similar to Dutch /k/.

The findings of this study revealed no link between perception and production. Dutch speakers had no difficulty in perceiving the /g-k/ contrast but were still less likely to produce /g/ in words borrowed from English. Martin et al. (2022) indicated that there are two possible explanations for these results. First, prevoicing, in non-intervocalic position, is more difficult to produce during velar constriction compared to labial and coronal constrictions. Second, the study did not examine the social associations that participants have with the two variants; the selection of one variant may depend on social factors. For example, some speakers may tend to use one variant more often in informal speech than in formal speech.

Several studies provide evidence for the impact of orthography on variable production of novel sounds. Vendelin and Peperkamp (2006) examined the production of English vowels by French bilingual speakers. Participants were asked to produce non-words in oral-only and oral-written conditions. In the former, they listened to non-words pronounced by a native speaker of English. In the latter, they listened to the non-words while viewing the written forms on a screen. The results showed differences in the response patterns of the two conditions; that is, the production of the English vowels was significantly influenced by the availability of written representations. For example, English vowels /ʊ/ and /u/ were adapted as /o/ in target non-words when these two vowels were represented by the grapheme <oo>.

Kang (2009) examined loanwords in 1930's Korean in which the consonant /s/ was produced either as lax /s/ or tense /s*/. The text data for the loanwords in this study was drawn from the Modern Chosun Loanword Dictionary which contains more than 18,000 loanwords appeared in written sources in the 1930's. Kang (2009) noted that whether the English /s/ was spelled with a single or two letters had a significant impact on selections between the two variants; if the consonant is spelled with two identical letters, it was often adapted as a tense (e.g., 'miss' /mis*ɨ/).

Finally, Hamann and Colombo (2017) showed that the production of English intervocalic consonants in Italian often depends on how words are spelled. Intervocalic consonants that are spelled with two identical letters are often adapted as geminates, such as 'banner' [ban.ner] and 'hobby' [ɔb.bi]. However, the study found some exceptions, such as 'puzzle' [pa:.zel] and 'fashion' [fɛʃʃon]. The consonant /ʃ/ in 'fashion' was adapted as a geminate although it is spelled

with two different letters (i.e., a diagraph <sh>). On the other hand, the consonant /z/ in 'puzzle' was adapted as a singleton [z] although it is spelled with two identical letters. The authors demonstrate that L1 phonology outweighs the impact of orthography in these exceptions: in Italian, /ʃ/ is always a geminate in intervocalic position while /z/ is always an intervocalic singleton.

2.3.2 External Factors

Intra-speaker and inter-speaker variability in the production of loanwords have been attributed to different external factors, such as bilingualism, attitudes and prestige. Bilingualism (i.e., native-like knowledge of L2) is significantly correlated with degree of adaptation; the higher one's level of bilingualism is, the more likely one is to preserve L2 segments (e.g., Paradis & LaCharite, 2011; Kwon, 2017; Kang & Schertz, 2021). As an example, consider a study by Kwon (2017) in which he showed variation in the perception of English word-final stops in novel borrowed words by Korean speakers with different degrees of bilingualism. Unlike early bilinguals, near-monolinguals and late bilinguals tended to insert a vowel after English word-final stops. The presence of release in English after word-final stops was perceived by near-monolinguals and late bilinguals as a vowel. Further, consider Lev Ari et al (2014)'s study in which speakers were influenced by the degree of bilingualism of their interlocutors. Lev Ari et al (2014) demonstrate that an interlocutor's level of bilingualism can influence a fellow speaker's likelihood of adaptation. In analysing the production of loanwords from Spanish into Mexicano, they found that speakers accommodated to their interlocutors' likelihood of adaptation. Speakers with low levels of proficiency in the L2 adapted less often when their interlocutors were more proficient in Spanish.

Prestige and attitudes are also reported to influence the likelihood of adaptation (e.g., Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2012). Paradis and LaCharité (2012) demonstrate that the retention of non-native forms reflects the prestigious status of the source language. However, borrowers may attempt to preserve novel segments due to the high prestige of the source language but fail to master articulations as a result of *flawed importation.* For example, French adaptations of the English interdentals /θ/ and /ð/ did not reflect the impact of the L1 phonology which would result in /t/ and /d/, or perception which would result in /f/ and /v/. Instead, English interdentals were adapted as /s/ and /z/ by French speakers, indicating that their attempts to preserve English segments were not successful. For instance, 'big brother' was

adapted by French speakers as [bɪgbʀɔzœʀ]. Paradis and LaCharité (2012) noted that these unexpected adaptation patterns elude explanations based on the L1 phonology and faulty perception. They explained that adaptation /t/ and /d/ to /s/ and /z/ is likely due to the high prestige of English in France. French borrowers' attempts to preserve novel segments resulted in flawed importations.

In a study conducted by Lev-Ari et al. (2014), it was found that there was a connection between the relative prestige of Spanish and the production of loanwords into Mexicano. In order to examine the impact of prestige of Spanish, the researchers classified the Spanish loanwords into categories of prestige. Spanish has a high prestige level in the domains of education and technology but has a low prestige level in the social domain. It was found that novel segments were more likely to be preserved in domains in which Spanish has more prestige than Mexicano.

Similarly, Lev-Ari and Peperkamp (2014) examined the probability that speakers' adaptations would be affected by the relative prestige of a donor language. In the study, French speakers were encouraged to pronounce an Italian product in spontaneous conversations using the game Go Fish. The product had a novel name, *genna*, that contains an initial affricate [dʒ] that does not exist in French. The product was used as a between-participant variable in the game as either 'an ice cream' or 'a beer'. The two products show Italy's relative prestige in regard to that of France; Italy's prestige is high for ice cream but low relative to France for beer. The findings showed that French speakers were less likely to produce [dʒ] when the word referred to Italian beer. In contrast, French speakers were more likely to retain [dʒ] when the word referred to Italian ice cream. The study also found that the more self-reported motivation French speakers had to speak a foreign language without a foreign accent, the less likely they were to make adaptations.

In a recent study, Bańko et al. (2022) explored the relationship between purism (i.e., speakers' tolerance for accepting the presence of loanwords in their native language) and the preference of a particular loanword adaptation technique. 213 Polish speakers were asked to complete two surveys. In the first survey, participants were asked to choose one among five potential names that had different degrees of adaptation for an invented loanword. In the second survey, participants were asked explicitly about their attitudes to the presence of English loanwords in Polish. It was found that the less puristic attitudes the participants had, the more often they accepted unadapted names.

2.4 English Loanwords into Arabic

This section reviews previous studies specifically on English loanwords into Arabic. The process of lexical borrowing has been studied in many Arabic dialects such as Jordanian (Abu Guba, 2016; Sa'aida, 2015), Egyptian (Galal, 2004; Hafez, 1996), Hadhrami (Alsaqqaf, 2006; Bahumaid, 2015) and Hijazi (Aloufi, 2016; Jarrah, 2013). Most of the previous studies attempted to provide a grammatical explanation of this phenomenon with particular attention to the morphological and phonological adaptations.

Segmental and syllabic adaptations of English loanwords into Arabic have been mostly attributed to the L1 phonology (e.g., Abu Guba, 2016; Aloufi, 2016, among others). For example, Aloufi, (2016) examined the phonology of English loanwords into Hijazi Arabic using two theoretical frameworks: Theory of Constraint and Repair Strategy (TCRS; Paradis & LaCharite, 1997) and Optimality Theory (OT; Prince and Smolensky, 1993). The findings of this study showed variation in the production of novel English sounds. For example, the affricates /dʒ/ and /tʃ/ are either imported or substituted with the fricatives /ʒ/ (e.g., jeans −>/ʒinz/) and /ʃ/ (e.g., chat −> /ʃa:t). The findings also showed modifications (e.g., epenthesis and deletion) in syllabic structure of English loanwords in Arabic. Aloufi, (2016) concluded that OT is better than the TCRS in explaining modifications; however, neither theory fully accounts for the variation found in the study. Similarly, Abu Guba, (2016) examined phonological adaptations of English loanwords into Ammani Arabic using the OT framework. The findings showed that English loanwords were modified by a number of phonological processes, such as deletion, epenthesis, assimilation and gemmination. In addition, the study showed variation in the production of English consonants /ʒ tʃ p v ŋ / that do not exist in Ammani Arabic. English consonants were sometimes imported (retained) and sometimes replaced with their counterparts /dʒ ʃ b f n/ in Ammani Arabic. These segmental and syllabic adaptations were considered to be phonological because Ammani Arabic phonology accounted for the majority of cases.

While a few prior investigations have examined the role of perception in the adaptation process in Arabic, these have only focused on the perception of coronal consonants. No attention has been made to the perception of other consonants. Previous studies suggest that variable adaptations of the coronal consonants /t/ and /s/ into Arabic loanwords can be predicted from the linguistic environment. Although /t/ and /s/ exist in Arabic, they are adapted as emphatic pharyngealised coronals (e.g., ‘tarzan’ [ṭɑrazan]) in the presence of the back vowel /ɑ/ (Naim,

1998; Louriz and Kenstowicz, 2009; Alzaaq, 2017). Alzaaq (2017) examined the effect of the back vowel /ɑ/ on the perception of English /t/ and /s/ by Saudi speakers of Arabic. The target items were non-words containing /t/ or /s/ followed by the back vowel /ɑ/. Participants were asked to listen to non-words and then write what they hear. The findings showed that /s/ was more likely to be adapted with pharyngealization than /t/.

Additionally, few studies examined the impact of orthography on the adaptation of English loanwords. Hamdi (2017) indicated that the effect of orthography is salient in adapting English vowels. He designed a questionnaire that contained multiple-choice questions. The participants were given two Arabic orthographic forms and asked to select the form they thought best represented the loanword. One form corresponds orthographically to the English word more than the other form. The findings showed that the selected Arabic orthographic forms are influenced by English orthography. For instance, the participants preferred <بسكويت > /biskawi:t/ instead of <بسكوت > /bisku:t/ as it is orthographically more similar to the original word 'biscuit'. <بسكويت > /biskawi:t/ contains /w/ representing <u> and /i:/ representing <i> while <بسكوت >/bisku:t/ contains only /u:/ representing <u>.

Abu Guba (2016) referred to some cases in which the English consonants written with two identical letters were adapted as geminates, such as 'roll'[rull] and 'drill' [drill]. However, the researcher argued that orthography does not always play a role in gemination because many words contain a consonant written with two identical letters but which is pronounced as a singleton, such as 'million' [malyoon]. Additionally, there were cases where gemination took place even when a consonant is written with one letter, such as 'net' [nitt]. Abu Guba (2016) argued that gemination is caused mainly by the native phonological system that requires the prosodic word to have at least two moras. To satisfy this requirement, either gemination or vowel lengthening is used to repair a monomoraic word because the single consonant in word-final position in Arabic is extrametrical.

Finally, relatively little is understood about the impact of social factors on variable adaptations of English loanwords into Arabic. Sa'aida (2015) examined phonological adaptations of English loanwords into Jordanian Arabic. The findings showed variation between participants in the production of loanwords. These findings were attributed to frequency of English use. It was reported that participants who did not use English frequently were more likely to repair English words that are not compatible with the L1 phonology. Additionally, in a study carried

out by Alnamer and Alnamer (2018), the findings showed that using English loanwords in Emirati Arabic was influenced by gender, age and level of education. It was found that females tended to use more English loanwords than males; additionally, educated and young participants tended to use more loanwords than uneducated and old participants.

2.5 The Effect of Input Modality on the Production of Novel Sounds

This section reviews prior studies on the impact of input modality on L2 production. Although there is relatively little discussion of the effects of input modality in relation to loanword phonology, orthographic effects are well-attested in the L2 literature. Previous studies have shown that orthographic representations of L2 sounds affect phonological awareness, speech production and speech perception (for an overview, see Hayes-Harb & Barrios, 2021). In the following L2 studies, researchers have reported mixed results regarding the impact of input modality on the production of novel segments.

Nimz and Khattab (2020) showed a facilitative effect of orthographic input on L2 production. Participants were Polish learners who were recruited at a high school that puts a special focus on German as a foreign language. This study aimed to examine whether orthographic cues are helpful for Polish learners in learning to produce a German vowel length contrast. In German, a vowel preceded by <h> is long and a vowel followed by double letters is short (e.g. *Höhle* /høːlə/ 'cave' vs. *Hölle* /hœlə/ 'hell'). Overall, the findings showed that these Polish learners tended to show a greater difference between short and long vowels when German orthographic cues were available.

However, orthographic input can also hinder L2 production, especially when the L2 has less transparent orthography than the L1. Vokic (2011) examined the production of English flap by native speakers of Spanish living in the United States. Participants were requested to read English words containing the flap sound /ɾ/ in carrier sentences. In English, the flap can be spelled as <t>, <d>, <tt> or <dd>. In Spanish, the flap sound is always spelled as <r> while /t/ and /d/ are spelled as <t> and <d>, respectively. Vokic (2011) found that Spanish speakers tended to produce the flap sound according to the letter-sound correspondences of Spanish; that is, they tended to produce a target flap sound as [t] or [d] when it was orthographically presented as <t> and <d>. This finding is congruent with the study of Bassetti (2017) which examined the production of English words containing double consonant letters (e.g., *floppy* and *pepper*) by Italian native speakers. Of particular interest was whether the presence of double consonant

letters would lead native speakers of Italian to produce a length contrast, because geminates are spelled as double letters in Italian. Italian participants were asked to read a list of English words in carrier sentences. Analysing the duration of consonants in participants' production of target words indicated that consonants were produced longer when they were spelled with two letters. The author concluded that Italian speakers' production of English words was affected by the transfer of Italian letter-sound correspondences.

The L2 production accuracy of non-native sounds can be affected by exposure to both orthographic and auditory inputs as opposed to auditory-only inputs. Davidson (2010) examined the production of non-native clusters by native speakers of English and Catalan. The participants performed a word repetition task in which some non-words were represented auditorily and some non-words were represented auditorily and orthographically. Speakers from the two language backgrounds tended to produce target non-native clusters more accurately when they were exposed to both orthographic and auditory inputs.

Bassetti and Atkinson (2015) examined the effect of orthography on Italian speakers' production of English words containing silent letters (e.g., *debt* and *climb*). The native Italian speakers had more than 10 years of English language instruction. They performed word reading and word repetition tasks. In the reading task, the words were written. In the repetition task, the words were written and produced by an English native speaker. Bassetti and Atkinson (2015) found that the silent letters were more likely to be produced by experienced instructed learners in the reading task than in the repetition task which involved auditory inputs.

The studies discussed above provide evidence regarding the general effect of the input modality in which a lexical item is initially encountered on the production of L2 sounds. Before concluding this chapter, let us briefly point out that the effect of input modality is influenced by individual differences in experience with the L2. Words are represented on different levels in the mental lexicon. The levels of representation convey semantic, phonological, orthographic information (Foucart & Frenck-Mestre; 2013). According to the Lexical Quality Hypothesis (Perfetti, 2007), the different levels of representations may vary in their degree of completeness; the same word may have a high-quality orthographic representation and a low-quality phonological representation, or vice versa. The quality of phonological and orthographic representations depends on how well a speaker knows how to spell and produce a word. A speaker may know how to produce a word but cannot spell it. Therefore, the quality of

representations may lead to variability when producing the same L2 sound in a word presented or encountered in different modalities.

According to the Lexical Entrenchment Hypothesis (Diependaele et al, 2013), individuals with a higher degree of L2 exposure are expected to have higher quality representations. Within this hypothesis, the processing of low-frequency L2 words improves with increasing language exposure. Having a high level of language exposure results in faster activation and less interference from similar lexical representations, leading to small differences in processing of low- and high- frequency words. Speakers with high levels of L2 exposure recognize low frequency words more accurately and faster than speakers with low levels of L2 exposure because of their high-quality lexical representations.

Two theoretical models of cross-language speech perception and acquisition, namely the Speech Learning Model (SLM; Flege, 1995) and the Perceptual Assimilation Model for second language learners (PAM-L2; Best & Tyler, 2007), relate learning of L2 phonetic categories with L2 experience. Both SLM and PAM-L2 assume that learning a new phonetic category depends on its degree of similarity to the closet L1 phonetic category and on the amount of experience with the L2. That is, non-native contrasts become more discernible with more experience of the L2. PAM-L2 states that L2 learners and naïve monolinguals differ in their L2 perception. According to PAM-L2, L2 learners will start out like naïve monolinguals, assimilating L2 categories to their closest L1 categories. That is, initially, L2 learners would assimilate two categories to one L1 category. However, new phonetic categories can be established as learners gain more knowledge in the L2 phonological and orthographic systems. The latest version of SLM (SLM-r; Flege & Bohn, 2021) states that the quality and quantity of L2 input are what matter most in learning new phonetic categories rather than the age of acquisition. Individuals should have enough exposure to an adequate quantity of high-quality inputs during L2 learning to form new L2 phonetic categories.

Numerous studies have found evidence that the amount of L2 exposure leads to differences in L2 perception and production. For example, Kwon (2017) showed that early bilinguals were more likely to perceive English final stops in novel words borrowed into Korean than late bilinguals. Mok et al. (2018) also showed that the negative impact of orthography on the production of Mandarin tones was more obvious for Cantonese learners with lower proficiency.

2.6 Summary

This chapter has defined the key terms of adaptation and importation used in this thesis, along with several examples of these phenomena in various languages. It has also discussed the status of English in Saudi Arabia and provided a review of the relevant literature on loanword phonology and L2 production.

The studies reviewed in this chapter provided important insights into different factors that may account for variable production of novel segments in loanwords. Overall, the previous literature suggests that the production of novel forms can be affected by a range of factors: the L1 phonology (e.g., Paradis & LaCharité, 2005), perception (e.g, Dupoux et al., 1999), orthography (e.g, Vendelin & Peperkamp, 2006) and social factors (e.g, Lev Ari et al., 2014).

In the present study, we focus specifically on the perception and production of target /v/ and /tʃ/ which were found to be variably produced in English loanwords into Arabic (e.g., Abu Guba, 2016; Aloufi, 2016). Studies on English loanwords into Arabic has largely examined the impact of Arabic phonology on the production of English segments (e.g., Al-Athwary, 2017; Abu Guba, 2016; Aloufi, 2016; Jarrah, 2013; Saaida, 2015); scant attention has been paid to other factors. There have been few studies that have assessed the role of perception (Alzaaq, 2017), orthography (Abu Guba, 2016; Hamdi, 2017) or social factors (Saaida, 2015) on variable adaptations of English loanwords into Arabic. Most previous studies were based on real words collected from the Internet, newspapers and magazines. The English loanwords into Arabic compiled by other researchers (e.g., Al-Athwary, 2017; Abu Guba, 2016; Aloufi, 2016) in previous studies will be checked and used in the production task in Chapter 5.

This study will examine this process from different angles contrasting with most studies on Arabic loanword phonology which focused mainly on the impact of native phonology. The study will examine the impact of input modality on the production of the two target sounds in interaction with a range of other factors: context (word position), level of English exposure and gender. The production task for non-words in Chapter 4, partially replicates the design of Vendelin and Peperkamp's (2006) experiment. As discussed in 2.31, Vendelin and Peperkamp (2006) tested French speakers' production in two conditions: aural-only and aural-written. In the former, they listened to non-words pronounced by a native speaker of English. In the latter, they listened to the non-words while viewing the written forms on a screen. However, in this study,

we will test Arabic speakers' production in three conditions: aural-only, written-only and aural-written.

In the next chapter, we will examine the discrimination accuracy in perception of the two target contrasts /v-f/ and /ʃ-tʃ/. The aim is to later determine whether there is any evidence for a link between individuals' perception and production so that individuals who produce more imported sounds [v] and [tʃ] also discriminate the target contrasts better than those who produce fewer imported sounds.

3 Discrimination Performance on the /v-f/ and /tʃ-ʃ/ Contrasts

3.1 Introduction

In this chapter, we examine the discrimination accuracy of the two target contrasts, /v-f/ and /tʃ-ʃ/, using the Perceptual Assimilation Model (PAM) framework. The question addressed in this chapter is as follows: *Do Saudi Arabic speakers find it difficult to perceive target /v/ and /tʃ/ as different from their L1 counterparts /f/ and /ʃ/?* The reason why /f/ and /ʃ/ were selected to contrast with English /v/ and /tʃ/ is because these two Arabic sounds are often used by Arabic speakers to replace /v/ and /tʃ/ (e.g., Abu Guba, 2016; Aloufi, 2016). An online oddity task[1] was conducted to determine the degree of difficulty in discriminating the two target contrasts. The perception task requires the participants to listen to three non-words and decide which non-word is different, with the option to say that the three non-words are the same if they do not hear an odd one. Following Nagle (2021), the 'all the same' option was included because participants may select correct answers by chance even though they do not perceive any of the non-words as different.

The remainder of this chapter spells out in detail the participants (Section 3.2), stimuli (Section 3.3), procedure (Section 3.4), predictions (Section 3.5), data analysis (Section 3.6), results (Section 3.7), discussion (Section 3.8), and summary (Section 3.9).

3.2 Participants

The participants were 67 Saudi speakers, including both males (31) and females (36). It was difficult to maintain equal numbers of males and females, given that this task was conducted online. They all live in Arar, a city located in the north of Saudi Arabia (see Figure 3.1). The participants were undergraduate students and therefore they were within the age range of 18 to 24. In Saudi Arabia, students should not exceed more than five years after obtaining the high school certificate or its equivalent to apply to an undergraduate degree program. These Saudi speakers use a similar consonant inventory to that of Najdi (Ingham, 1994) and Hejazi (Omar & Nydell, 1975) which are used by speakers in large cities in Saudi Arabia, such as Riyadh, Jeddah, Almaidnah, and Makkah. The participants were students at the Northern Border University with

[1] The task was piloted with a small set of participants using the same stimuli and procedure except with a minor change. As will be indicated in Section 3.3, same trials (AAA and BBB) were included in the pilot study but not in the main study.

different levels of English exposure. All the participants were expected to have some knowledge of English because it is typically introduced through primary and secondary school education and the media. However, the participants were recruited in three groups based on their expected level of English exposure (i.e. based on their academic program). The first group consisted of 24 students from the department of English, the second group consisted of 22 students from the department of Computer Science, and the third group comprised 21 students from the department of Arabic. Students in the department of English typically have the most frequent exposure to English in classes while students in the department of Arabic have the least exposure to English in classes. Students in the department of Computer Science fall in between the two other groups. All participants spent at least two years in their academic programs. None of the participants reported hearing or speaking impairments.

The University of Northern Borders was contacted to seek permission to undertake this study among their students. Considerable time and effort were expended to recruit an appropriate number of participants. Participants did not receive any payment or reward for their participation; they were invited to voluntarily participate in this study by the head of each department. Some participants accepted the invitation to participate in the study while others refused. Participants' contact information was recorded based on their agreement. Participants, who accepted to be involved in the study, agreed to give their phone numbers so that the researcher could contact them via the *WhatsApp* application to explain instructions and send them links to the present perception task, production tasks in Chapters 4 and 5 and attitudes survey in Chapter 6.

Table 3.1: Participants' details

	Females	Males	Total
Frequent Exposure to English (Students from the department of English)	12	12	24
Less Exposure to English (Students from the department of Computer Science)	12	10	22
Rare Exposure to English (Students from the department of Arabic)	12	9	21
Total	36	31	67

Figure 3.1: Location of participants' place of residence (Wikipedia, 2021)

3.3 Stimuli

Four contrasts were included, two target contrasts and two distractors. The target stimuli are contrasted by voicing (/v-f/) and manner of articulation (/ʧ-ʃ/). Selecting Arabic consonants that contrast with English consonants is based on the findings of previous studies of English loanwords into Arabic (e.g., Abu Guba, 2016; Aloufi, 2016). Target /v/ and /ʧ/ are often replaced by /f/ and /ʃ/, respectively. The distractor stimuli contrast by voicing (/t-d/) and place of articulation (/m-n/). The distractor contrasts were chosen because they exist in both Arabic and English. The target and filler items were represented by six different trials (AAB, ABA, ABB, BAA, BBA, BAB). Same trials (AAA and BBB) were excluded for two reasons. The first reason

for the exclusion is that same trials did not impose any difficulty for participants in an earlier pilot study. The second reason is that the task was already very long even without the same trials. This perception task took approximately 45 to 60 minutes to be completed. However, as mentioned in 3.1, the 'all the same option' option was included in the task because participants may select correct answers by chance even though they do not perceive any differences between the non-words.

To avoid a possible effect of lexical frequency, the stimuli in this task are CVCVC non-words, with the primary stress placed on the first syllable. The CVCVC syllable pattern is used for two reasons. The first reason is that, as it exists in Arabic (e.g., ['mælɪk] 'owner'), this template does not impose difficulty for Arabic speakers. The second reason is to determine whether the position of the consonant within a word influences participants' performance. The target contrasts, which occur in three positions — initial, intervocalic and final — are embedded in these different positions because evidence suggests that the level of difficulty in perceiving and producing a novel sound can be affected by its position within a word (e.g., Huang & Lin, 2016; Kubozono, Itô & Mester, 2008). The selected vowels are the near-low front vowel /æ/ in the first syllable and the near-high front vowel /ɪ/ in the second syllable. These two vowels were also chosen because they exist in both Arabic and English.

Table 3.2 shows the 48 target items used in the study. The target items containing /v/ and /tʃ/ were selected based on the naturalness ratings of native speakers of English. This norming study was carried out in an attempt to select non-words that sound as English-like as possible. Non-words in the norming study were presented orthographically with their phonetic transcriptions. Participants were asked to judge how natural each non-word on a scale ranging from 0 to 10 (see Appendix A for more details of the naturalness rating). Of 78 possible candidates, 24 target non-words were selected. The selected non-words received a median score of 5 or higher. Any word that was identified by a native speaker as an actual English word was excluded. Out of many possible candidates, the final list of filler items was selected by a native speaker of English (see Appendix B); unnatural items (i.e., non-words that do not sound like plausible English words) were excluded.

Recordings of all target and filler items were provided by a 55-year-old female native speaker of British English who is originally from the West Midlands. All stimuli were produced with a standard British accent. Recordings were made using a high-quality solid-state digital

recorder *(*Marantz PMD-610) and Shure SM10A microphone*,* with the default sampling frequency 44100Hz in WAV format. The native speaker recorded three versions of each target word. In the ABA trial, for example, the token in the first A is different from the token in the second A. The female native speaker was requested to produce target sounds in the same way she would produce them in her native language (e.g., the native speaker was not asked to produce final /v/ with full voicing). All the target and filler items were normalised at the same intensity level (75 dB).

Table 3.2: Target items in the perception task

/v-f/			/tʃ-ʃ/		
Initial	Intervocalic	Final	Initial	Intervocalic	Final
/væpɪt/ /fæpɪt/	/ɹævɪd/ /ɹæfɪd/	/bælɪv//bælɪf/	/tʃæɹɪs/ /ʃæɹɪs/	/ɹætʃɪm/ /ɹæʃɪm/	/mænɪtʃ/ /mænɪʃ/
/væɹɪʃ/ /fæɹɪʃ/	/sævɪʃ/ /sæfɪʃ/	/bæɹɪv/ /bæɹɪf/	/tʃæsɪt/ /ʃæsɪt/	/ɹætʃɪn//ɹæʃɪn//	/tænɪtʃ/ /tænɪʃ/
/vænɪt/ /fænɪt/	/nævɪʃ/ /næfɪʃ/	/bætɪv//bætɪf/	/tʃænɪt/ /ʃænɪt/	/ɹætʃɪl/ /ɹæʃɪl/	/lætɪtʃ/ /lætɪʃ/
/vælɪt/ /fælɪt/	/bævɪn//bæfɪn/	/gælɪv//gælɪf/	/tʃælɪt/ /ʃælɪt/	/bætʃɪn//bæʃɪn/	/pænɪtʃ/ /pænɪʃ/

3.4 Procedure

An oddity task was used to elicit participants' perception accuracy of the two contrasts. The oddity task was chosen because the chance level in this task is lower than AX or AXB tasks (Nagle, 2021). The oddity task required participants to decide which stimulus is different, with the option to say that the three auditory stimuli are the same if they do not hear an odd stimulus. Responses were made by mouse clicking on one of the four response fields (1, 2, 3 and x as an option for 'all the same'). The order of trials was randomised for each participant.

The task was conducted online due to the strict COVID-19 ban on in-person data collection. *Gorilla* (https://gorilla.sc) was used to design this task. Gorilla was chosen for two reasons. The first reason is that it is safe (fully compliant with the General Data Protection Regulation). The second reason is that it is easy to use and does not require downloading any software. The task link was sent to participants using the *WhatsApp* application. To ensure anonymity, the participants were first requested to use anonymised codes that were given to them by the researcher. They could not start the task unless they entered their code and signed the online consent form. After that, they were asked to read the instructions carefully (see Appendix

D) and then to start the task. It was explained to participants that they should listen to a sound file containing three words in each trial and then they need to select which word is odd by mouse clicking (1, 2 or 3), with the option (x) to say that the three words are the same if they do not hear an odd word. Participants were instructed to perform the task on a computer or laptop in a quiet place using a microphone and headphones within the expected time (45 minutes).

At the outset, participants were asked to put on headphones and to adjust the volume to the desired loudness level. Four practice trials, using filler items, were given first to ensure that participants knew how to do the task. 5-minute breaks between blocks of test trials were allowed because the task takes a long time to complete. On each trial, participants listened to an audio file that contains a series of three auditory stimuli. Each audio file was played automatically one after the other once the page had loaded. Participants were allowed to play the audio file again multiple times in case they were interrupted (e.g., phone calls or loss of Internet connection). The audio file was played again more than one time in 9.66% of the trials for target /v/ (466 out of 4824) and in 6.26% of the trials for target /tʃ/ (302 out of 4824).

There were two target contrasts (/v-f/and /tʃ-ʃ/) × 12 non-words (i.e. four words in the three different positions) × six trials per contrast (AAB, ABA, ABB, BAA, BBA, BAB) = 144 trials. The time between the presentation of each auditory stimulus, interstimulus interval (ISI), was 500 msec. An ISI between 500 and 1,000 ms is expected to be sufficient to produce excellent discrimination performance (Gerrits & Schouten, 2004). An example of each trial is provided in Table 3.3 and a screenshot of the task as seen by participants in Figure 3.2.

Table 3.3: An example of oddity task trials

AAB	/væpɪt/	/væpɪt/	/fæpɪt/
ABA	/væpɪt/	/fæpɪt/	/væpɪt/
BAB	/fæpɪt/	/væpɪt/	/fæpɪt/
ABB	/væpɪt/	/fæpɪt/	/fæpɪt/
BAA	/fæpɪt/	/væpɪt/	/væpɪt/
BBA	/fæpɪt/	/fæpɪt/	/væpɪt/

Figure 3.2: A screenshot of the oddity task [2]

Some participants did not follow all instructions since the study was not conducted in person. Twenty-two participants took longer than was expected to complete the task (i.e., they spent more than one hour to complete the task) and nineteen participants resumed where they left off in the task after one day or even days. Appendix F contains a histogram showing the distribution of time taken by the participants to complete the task. There were no restrictions on the type of device for participating. Some participants had to use other devices (e.g., phones and tablets) as they normally would use to access their online classes because they had no access to computers or laptops. It was difficult to ensure that all participants used headphones to listen to the stimuli, so participants were not asked if they used headphones when listening to the stimuli. All participants who completed the task were included in the study. The study did not apply strict inclusion criteria because it was already difficult to recruit and incentivize participants to complete the full sequence of online tasks, especially given that the participation was completely voluntary.

3.5 Predictions

The Perceptual Assimilation Model (PAM) was chosen to predict results in this task. The model was developed by Best (1995) to address non-native speech perception. PAM proposes

[2] Translation: Identify the word that is different. 1= the first word 2= the second word 3= the third word x= all the words are the same.
The task was presented in Arabic because the purpose of this study is to examine the realization of English loanwords in Arabic.

that a listener's ability to discriminate between non-native phonemes can be predicted depending on the degree of similarity between non-native sounds and native categories. This model was chosen because it accounts for variation in the perception of different types of non-native contrasts. PAM does not state an explicit association between speech perception and production. However, it posits that listeners are able to perceive information about the articulatory gestures of non-native phonemes which suggests a link between speech perception and production. Before delving into the expected patterns of discrimination for the two target contrasts /v-f/ and /ʃ-tʃ/, it is important to explain the six assimilation types in PAM of non-native contrasts.

Discrimination is expected to be excellent if there is *two category (TC) assimilation*, in which each non-native sound is assimilated to a different native category. Discrimination accuracy should be excellent because listeners can use their L1 phonological knowledge to make a distinction between the two sounds. In contrast, discrimination is expected to be poor if there is *single category (SC) assimilation*, in which two non-native sounds are assimilated equally to a single native category. Discrimination accuracy should be poor because listeners are unable to detect differences between the two sounds. One non-native sound may be considered a better exemplar of the category than the other, which leads to *category goodness (CG) assimilation*, in which the discrimination is greater than in SC but still less than in TC. Discrimination accuracy should be good because one sound is perceived as a better version of one L1 category than the other.

In situations where one or both of the non-native sounds fail to be assimilated to any native category, the model allows for three possible assimilation types. Two non-native sounds maybe *uncategorized (UU)* when both sounds are assimilated as poor exemplars of two or more native categories. Discrimination is expected to range from poor to excellent depending on the acoustic proximity of the non-native sounds to each other. Discrimination accuracy should be poor if the two sounds are similar. However, when one of the two non-native sounds is categorized *(categorized-uncategorized or UC),* discrimination should be good because the other sound is not assimilated to any native category. Finally, both non-native sounds may be perceived as *non-assimilable (NA)* when they have articulatory features that are quite distinct from any native speech sound. Discrimination is expected to range from good to excellent depending on their acoustic properties. Examples of each assimilation type are listed below.

1. TC: In Danish perception of English plosives (Horslunda et al., 2015), English plosives /t/, /d/, /k/, and /g/ are assimilated by Danish speakers to their counterparts in Danish.
2. SC: In English perception of Zulu bilabial stops (Best, et al., 2001), voiced plosive and implosive bilabial stops in Zulu are equally assimilated as /b/ by American English speakers.
3. CG: In English perception of Zulu velar stops (Best, et al., 2001), the ejective and voiceless velar stops /k'/ and /k/ are both assimilated by American English speakers to their English counterpart /k/, but with a preference to the latter.
4. UC: In Japanese perception of Australian English vowels (Bundgraad-Nielsen, et al., 2001), /u:/ is assimilated to its Japanese counterpart but /ɜ/ is uncategorised.
5. UU: In Japanese perception of Australian English vowels (Bundgraad-Nielsen, et al., 2001), /əʉ/ and /o:/ are both uncategorised because they have no counterparts in Japanese.
6. NA: In English perception of Zulu clicks, the clicks are not perceived as speech sounds by American English speakers (Best et al., 1988).

Best and Tyler (2007) extended the PAM, which is for naïve listeners, to account for the perception pattern of L2- learning listeners rather than naïve listeners (PAM-L2). Unlike L2-learning listeners, naïve listeners are not familiar with the acoustic details of L2 phonological categories. L2- learning listeners' experience with L2 phonology and orthography will likely influence their perception patterns (i.e., how the L2 sound is assimilated to the L1 phonological category). For example, Best and Tyler (2007) pointed to English learners of French who tend to assimilate the French /ʁ/ and English /ɹ/ to the same phonological category despite the clear acoustic differences between the two sounds. The reason for this is that the two sounds are similar in terms of phonology (i.e., occur in the same syllable positions) and orthography (represented by the same letter <r>). Evans and Alshangiti (2018) examined the perception of English vowels and consonants by 26 Saudi learners with two levels of English proficiency: high and low. It was found that Saudi learners had difficulties with English affricates, and high front, high back, and central vowels. However, learners with high proficiency in English outperform those with low English proficiency, suggesting that their L2 knowledge helped them to perceive these sounds.

The assimilation types in the original PAM also apply to L2 perceptual learning. For example, L2- learning listeners are expected to have no difficulty with an L2 contrast if the two

sounds are assimilated to two phonological categories in the L1 (TC assimilation). The discrimination is expected to be more difficult if the two sounds are assimilated to one L1 phonological category, in which one sound is perceived as a better exemplar of the of the L1 phonological category than the other (CG assimilation). For the CG assimilation type, L2-learning listeners will likely develop a new phonological category for the deviant member with more exposure to the L2. The discrimination is expected to be poor if the two sounds are equally assimilated to the same L1 phonological category (SC assimilation).

Referring to the PAM and PAM-L2, the following patterns of discrimination are expected for the two target contrasts in the present task:

a) /v-f/ are expected to belong to the same Arabic phonological category /f/ (SG assimilation). /f/ exists in all Arabic dialects but /v/ in none. That is, the /v-f/ contrast includes one familiar sound /f/ and one unfamiliar sound /v/. The unfamiliar sound shares some phonetic features with the familiar one. The discrimination accuracy for /v-f/ is expected to be poor.

b) /ʃ-tʃ/ are expected to belong to one Arabic phonological category *(*CG assimilation) in which /ʃ /is a better exemplar of the category than /tʃ/. The sound /ʃ/ exists in all Arabic dialects, but /tʃ/ only exists in certain Arabic dialects spoken in Gulf countries (e.g., Kuwait and Bahrain). Although Saudi speakers will be familiar with both consonants, /ʃ / is the ideal because it exists in their L1 inventory and /tʃ/ which exists only in other Arabic dialects is the deviant. The discrimination accuracy for /ʃ-tʃ/ is expected to be moderate to very good.

3.6 Data Analysis

Participants' responses were binary coded for accuracy (correct and incorrect). If a participant successfully identified the odd sound in each trial, their perception accuracy was labelled as 'correct'. If not, their perception response was labelled as 'incorrect'. An example is shown in Table 3.4.

Table 3.4: Illustration of how perception accuracy is identified based on the given answer

Trial	Stimuli	Context	Answer	Listeners' accuracy
ABA	/væpɪt/ /**fæpɪt/** /væpɪt/)	Initial	2	Correct
AAB	/bælɪv/ /bælɪv/ **/bælɪf/**	Final	1	Incorrect

The analysis was performed with R statistical software (R core team, 2019). First, descriptive statistics were used to summarise and visualise the binary outcomes for each contrast and independent variable using the *tidyverse* package (Wickham, 2016). Second, results for each contrast were explored in mixed-effects logistic regression (*glmer*) using the *lme4* package (Bates et al. 2015). This statistical method was used because it predicts the binary outcomes of the dependent variable (correct and incorrect responses) and accounts for random effects and repeated observations (Winter, 2013, 2019). The predictor variables were word position (initial, intervocalic and final), English exposure group (high, medium and low), and gender (male and female).

The inclusion of interactions was assessed via model comparison using the *anova* function. The simple model, without interactions, was adopted if the interactions were not significant and there were no significant differences between the models (p >. 05). The *all_fit* function from the *afex* package (Singmann et al. 2018) was used to indicate the optimizer that would converge the models successfully. All categorical variables were converted into numbers using dummy coding. For accuracy, 0 corresponded to 'incorrect' and 1 corresponded to 'correct'. That is, positive values suggested a higher accuracy for the target contrasts. The pairwise predictions of constructed models were estimated using the *emmeans* package (Lenth, 2021). Model structures used in the analysis are given in a footnote in the next section.

3.7 Results

This section presents the descriptive and inferential results of participants' discrimination accuracy concerning the two contrasts.

3.7.1 Descriptive Analysis

The total number of responses for each contrast was 4,824 (67 participants × 12 words × 6 trials). Table 3.5 shows the proportion of correct and incorrect responses for each contrast in this task.

Table 3.5: Participants' discrimination accuracy of the target contrasts

	/tʃ-ʃ/		/v-f/	
	Count	%	Count	%
Correct responses	3,528	73.13%	2,198	45.56 %
Incorrect responses	1,296	26.87%	2,626	54.44%
Total	4,824		4,824	

Figures 3.3 and 3.4 display the proportions of participants' discrimination accuracy for the /v-f/ and /tʃ-ʃ/ contrasts. Tables 3.6-3.8 present the raw results split by the key independent variables. The raw results suggest that there is impact of English exposure on the discrimination accuracy of the two contrasts. The best discrimination results for both contrasts were achieved by the group of participants with the high level of exposure to English followed by the participants with medium and low levels of English exposure. Figure 3.3 shows that the lowest discrimination accuracy of the /v-f/ contrast is in word-final position while Figure 3.4 shows that the discrimination accuracy of the /tʃ-ʃ/ contrast seems to be consistent across the three-word positions. As can be seen in the figures, there was a slight difference between the female and male participants in their discrimination performance on the two contrasts; the male participants had a somewhat lower degree of discrimination accuracy than the female participants.

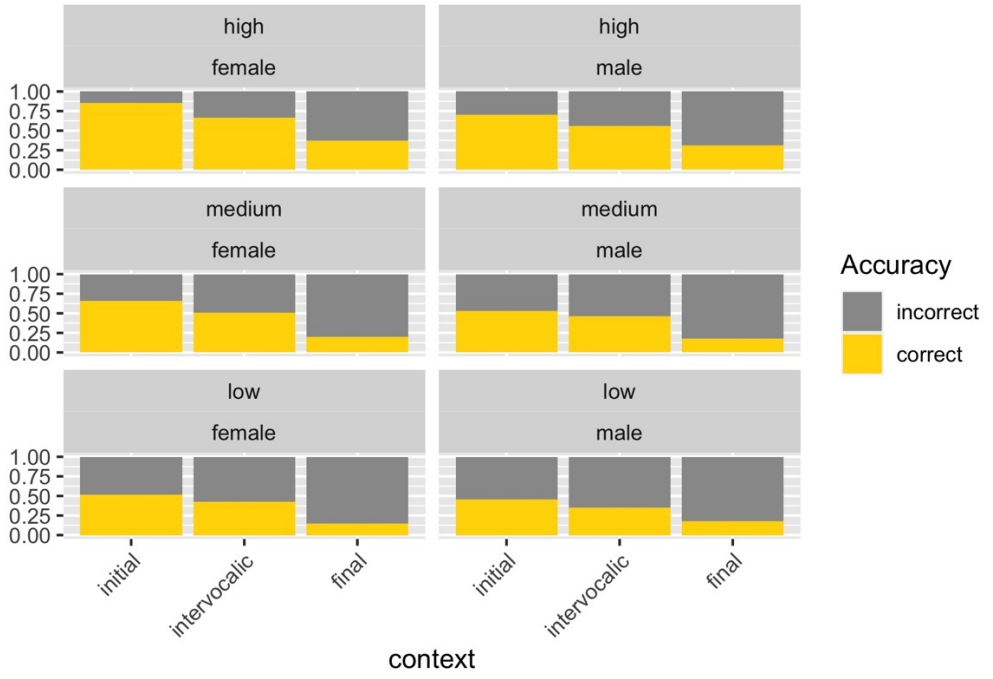


Figure 3.3: Participants' discrimination accuracy of the /v-f/ contrast by word position, language exposure group and gender

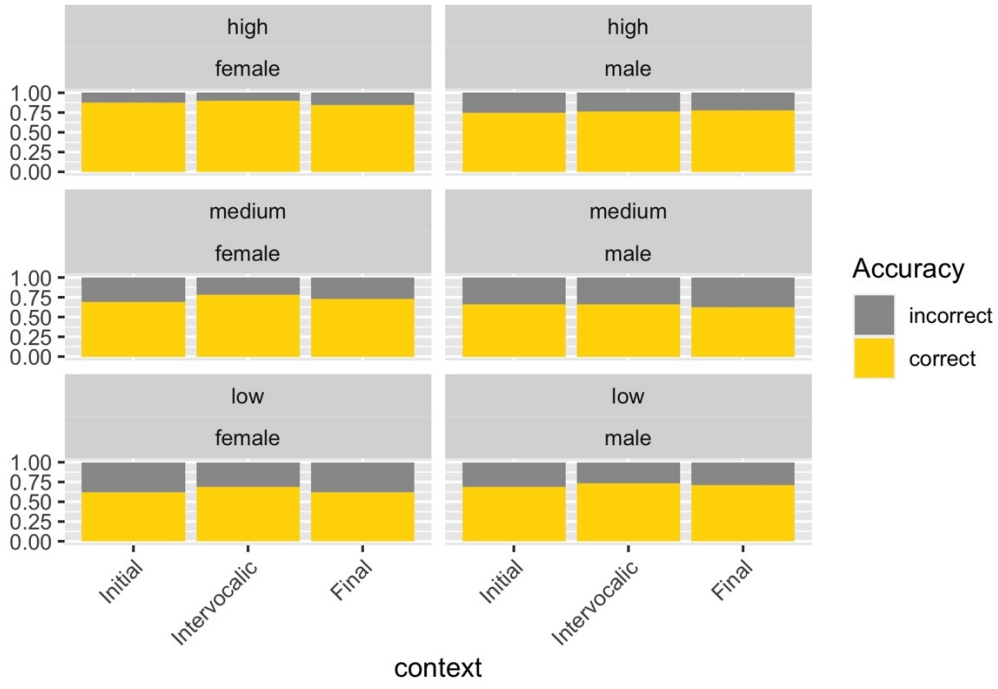


Figure 3.4: Participants' discrimination accuracy of the /tʃ-ʃ/ contrast by word position, language exposure group and gender

Table 3.6: Discrimination accuracy proportions of the /tʃ-ʃ/ and /v-f/ contrasts by language exposure group

	/tʃ-ʃ/				/v-f/			
	Correct responses		Incorrect responses		Correct responses		Incorrect responses	
	Count	%	Count	%	Count	%	Count	%
High group	1,412	81.71%	316	18.29%	999	57.81%	729	42.19%
Medium group	1,099	69.38%	485	30.62%	672	42.42%	912	57.58%
Low group	1,017	67.26%	495	32.74%	527	34.85%	985	65.15%

Table 3.7: Discrimination accuracy proportions of the /tʃ-ʃ/ and /v-f/ contrasts by word position

	/tʃ-ʃ/				/v-f/			
	Correct responses		Incorrect responses		Correct responses		Incorrect responses	
	Count	%	Count	%	Count	%	Count	%
Initial-word position	1149	71.46%	459	28.54%	1012	62.94%	596	37.06%
Intervocalic-word position	1219	75.81%	389	24.19%	809	50.31%	799	49.69%
Final-word position	1160	72.14%	448	27.86%	377	23.45%	1231	76.55%

Table 3.8: Discrimination accuracy proportions of the /tʃ-ʃ/ and /v-f/ contrasts by gender

	/tʃ-ʃ/				/v-f/			
	Correct responses		Incorrect responses		Correct responses		Incorrect responses	
	Count	%	Count	%	Count	%	Count	%
Females	1942	74.92%	650	25.08%	1251	48.26%	1341	51.74%
Males	1586	71.06%	646	28.94%	947	42.43%	1285	57.57%

3.7.2 Logistic Regression Analysis

Having discussed the descriptive results, we turn to the inferential statistical analyses. The results were tested using mixed-effects logistic regression that models the discrimination accuracy variable (correct and incorrect) as a function of different predictor variables (word

position, level of English exposure and gender). All the predictor variables were dummy coded. Two models for the two contrasts were run separately to explore the significant predictors affecting the discrimination accuracy of each contrast. Then, a combined model was run to show the significant predictors affecting the discrimination accuracy across the two contrasts. The benefit of the combined model is to show whether the discrimination accuracy of one contrast is more affected by one predictor than the other.

The model[3] for the /v-f/ contrast includes word position (initial, intervocalic and final), exposure group (high, medium and low), and gender (males and females) as fixed effects. As random effects, there were random intercepts for trial, and word, as well as a random slope by participant for the effect of word position. The reason for including a slope of word position by participant is because plotting the perception accuracy for each participant in the three groups showed that the effect of word position was not the same for all participants. Some variability was observed between participants within groups.

The coefficients of the adopted model for the /v-f/ contrast are shown in Table 3.9 and visualised in Figure 3.5. The estimate for intercept is the estimate for word-final position, low exposure group (the participants with the lowest level of English exposure), and males. The intercept is negative and significant, meaning that the /v-f/ contrast was less likely to be discriminated by the participants in the low exposure group, males, and in word-final position. The model reveals a main effect of word position: the coefficients associated with word-initial position ($\beta = 2.3776$, SE = 0.7342, $z = 3.239$, $p = 0.00$ **) and word-intervocalic position ($\beta = 1.4673$, SE = 0.7344, $z = 1.998$, $p = 0.04$ *) are each significantly different from those associated with word-final position. As shown in Figure 3.5, discrimination accuracy was significantly higher in word-initial and intervocalic positions than in word-final position. The model also confirms a significant effect of level of English exposure. The coefficients associated with the high exposure group ($\beta = 1.3708$, SE = 0.3391, $z = 4.042$ $p = 5.30e\text{-}05$ ***) were significantly different from those associated with the low exposure group. These results suggest that the /v-f/ contrast was more likely to be discriminated by the participants in the high exposure group. As shown in Figure 3.5 the participants in the medium exposure group also had a somewhat higher

[3] glmer (accuracy~ context + group + gender + (1+context | participant)+ (1 | trial) + (1 | word),data = data_v, family = "binomial")

degree of discrimination accuracy than the participants in the low exposure group (though not reaching statistical significance). Finally, the model shows that the difference between the female and male participants is not significant, indicating that they were similar in their discrimination accuracy for the /v-f/ contrast.

Table 3.9: Mixed effects logistic regression model for the /v-f/ contrast

Fixed effects	Estimate	SE	z value	Pr(>|z|)
Intercept	-2.3910	0.5993	-3.989	6.62e-05 ***
Word-initial position	2.3776	0.7342	3.239	0.0012 **
Word-intervocalic position	1.4673	0.7344	1.998	0.0457 *
High group	1.3708	0.3391	4.042	5.30e-05 ***
Medium group	0.4308	0.3374	1.277	0.2016
Females	0.3009	0.2743	1.097	0.2725

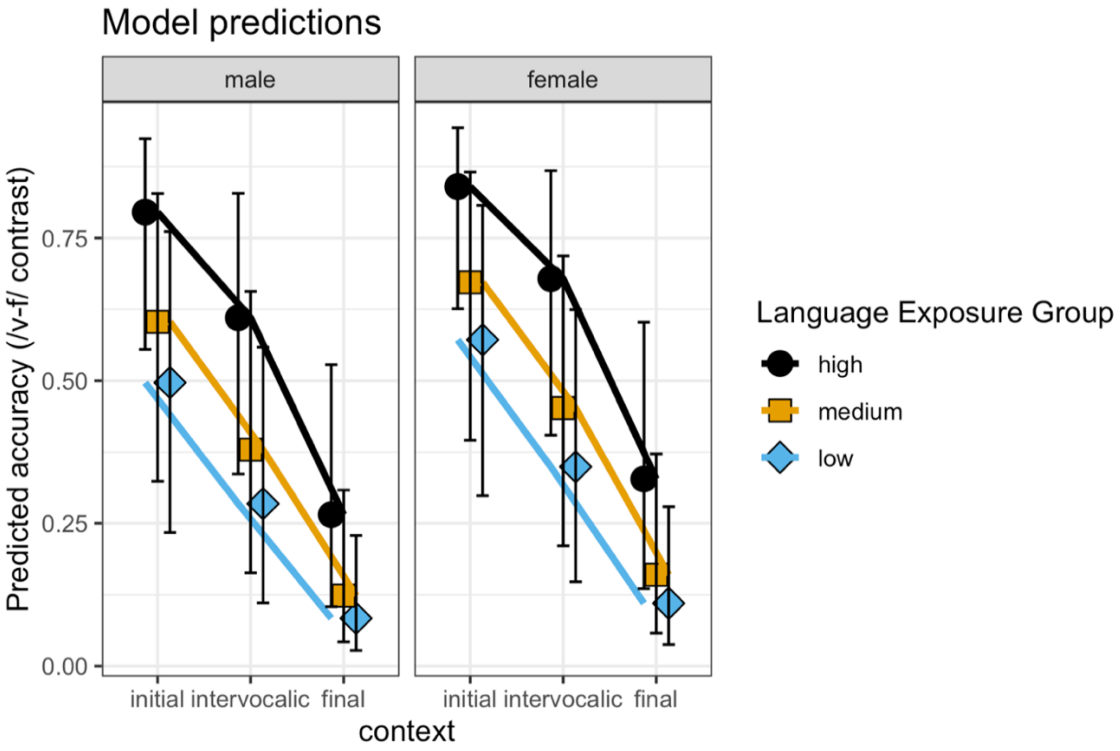


Figure 3.5: Predicted accuracy of the /v-f/ contrast by word position, language exposure group and gender

Turning to the /tʃ-ʃ/ contrast, the results of the mixed effects model[4] are presented in Table 3.10 and visualised in Figure 3.6. The estimate for the intercept is the estimate for word-final position, low exposure group, and males. The fitted model for the /tʃ-ʃ/ contrast includes word position (initial, intervocalic and final), language exposure group (high, medium and low), and gender (males and females) as fixed effects. There are random intercepts for trial and word and a slope for word position by participant. The effect of word position is not significant. Additionally, gender does not approach statistical significance. Participants in the medium and low exposure groups performed similarly. However, the difference between the high exposure group and the low exposure group is significant (β = 1.13763, SE = 0.44447, z = 2.560, p = 0.01*). As shown in Figure 3.6, there is also a significant difference between the high exposure group and the medium exposure group, suggesting that the /tʃ-ʃ/ contrast was more likely to be discriminated by the participants with the high level of English exposure than the participants with low and medium levels of English exposure.

Table 3.10: Mixed effects logistic regression model for the /tʃ-ʃ/ contrast

Fixed effects	Estimate	SE	z value	Pr(>|z|)
Intercept	0.90362	0.44978	2.009	0.0445 *
Word-initial position	-0.09266	0.30854	-0.300	0.7639
Word-intervocalic position	0.27539	0.30690	0.897	0.3696
High group	1.13763	0.44447	2.560	0.0105 *
Medium group	0.09309	0.44654	0.208	0.8349
Females	0.16738	0.37577	0.445	0.6560

[4] glmer(accuracy~ context + group + gender + (1+context | participant)+ (1 | trial) + (1 | word),data = data_ch, family = "binomial")

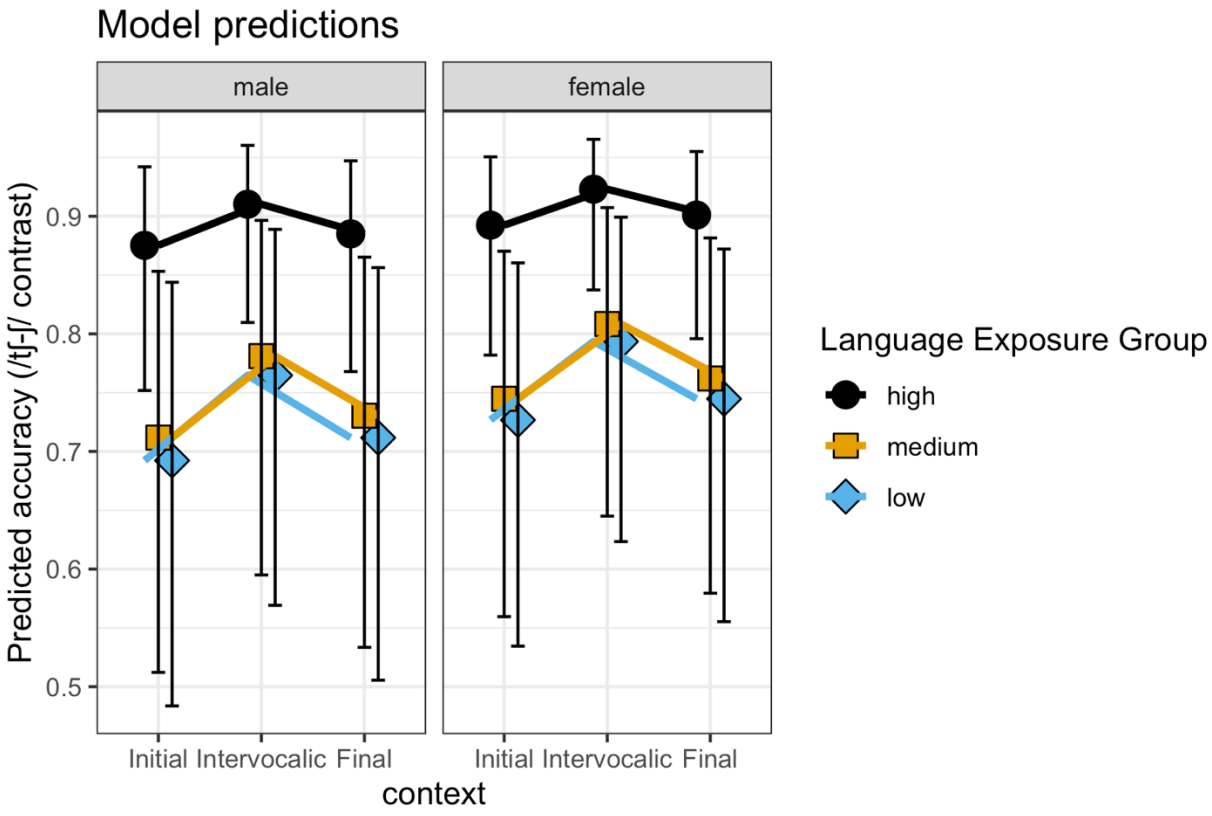


Figure 3.6: Predicted accuracy of the /tʃ-ʃ/ contrast by word position, language exposure group, and gender

Table 3.11 and Figure 3.7 present the results of the combined model[5] constructed for the two contrasts. As fixed effects, group, gender, and a two-way interaction between word position and contrast were entered into the model. The interaction was included because word position, as shown in individual models, did not have the same effect on the two contrasts. As random effects, the model included two random slopes for the effect of contrast and word position by participant, and two intercepts for trial, and word. As explained in the individual model for the /v-f/ contrast, the reason for including a slope of word position by participant was because the effect of word position was not the same for all participants within each group. The reason for including a slope of a contrast by participants is because the perceptual sensitivity for the two

[5] glmer(accuracy~ group + contrast*context + gender + (1+contrast| participant)+ (1+context| participant)+ (1 | trial) + (1 | word),data = data, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

contrasts may vary across participants. The estimate for intercept in this model is the estimate for word-final position, low exposure group, males and the /tʃ-ʃ/ contrast. The /v-f/ contrast differs significantly from the /tʃ-ʃ/ contrast (β = -2.90145, SE = 0.55187, z = -5.257, p = 1.46e-07 ***). Figure 3.7 shows that participants were more accurate at discriminating the /tʃ-ʃ/ contrast than the /v-f/ contrast. The model confirms the effect of language exposure: the high exposure group differs significantly from the low exposure group (β = 1.32259, SE = 0.33567, z = 3.940, p = 8.14e-05***), meaning that the participants in the high exposure group had a higher discrimination accuracy for both contrasts than the participants in the low exposure group. In terms of gender, the difference between males and females was not significant. The model reveals a significant interaction between the /v-f/ contrast and word-initial position (β = 2.24417, SE = 0.76206, z = 2.945, p = 0.00**), indicating that the /v-f/ contrast was more likely to be discriminated in word-initial position than in word-final position.

Table 3.11: Mixed effects logistic regression model for the /v-f/ and /tʃ-ʃ/ contrasts

Fixed Effects	Estimate	SE	z value	Pr(>|z|)
Intercept	0.60056	0.49528	1.213	0.22529
High group	1.32259	0.33567	3.940	8.14e-05 ***
Medium group	0.37439	0.33712	1.111	0.26676
Contrast /v-f/	-2.90145	0.55187	-5.257	1.46e-07 ***
Females	0.33205	0.27169	1.222	0.22166
Word-initial position	0.04195	0.54535	0.077	0.93868
Word-intervocalic position	0.31898	0.54531	0.585	0.55858
Contrast /v-f/: Initial position	2.24417	0.76206	2.945	0.00323 **
Contrast /v-f/: Intervocalic position	1.08845	0.76296	1.427	0.15369

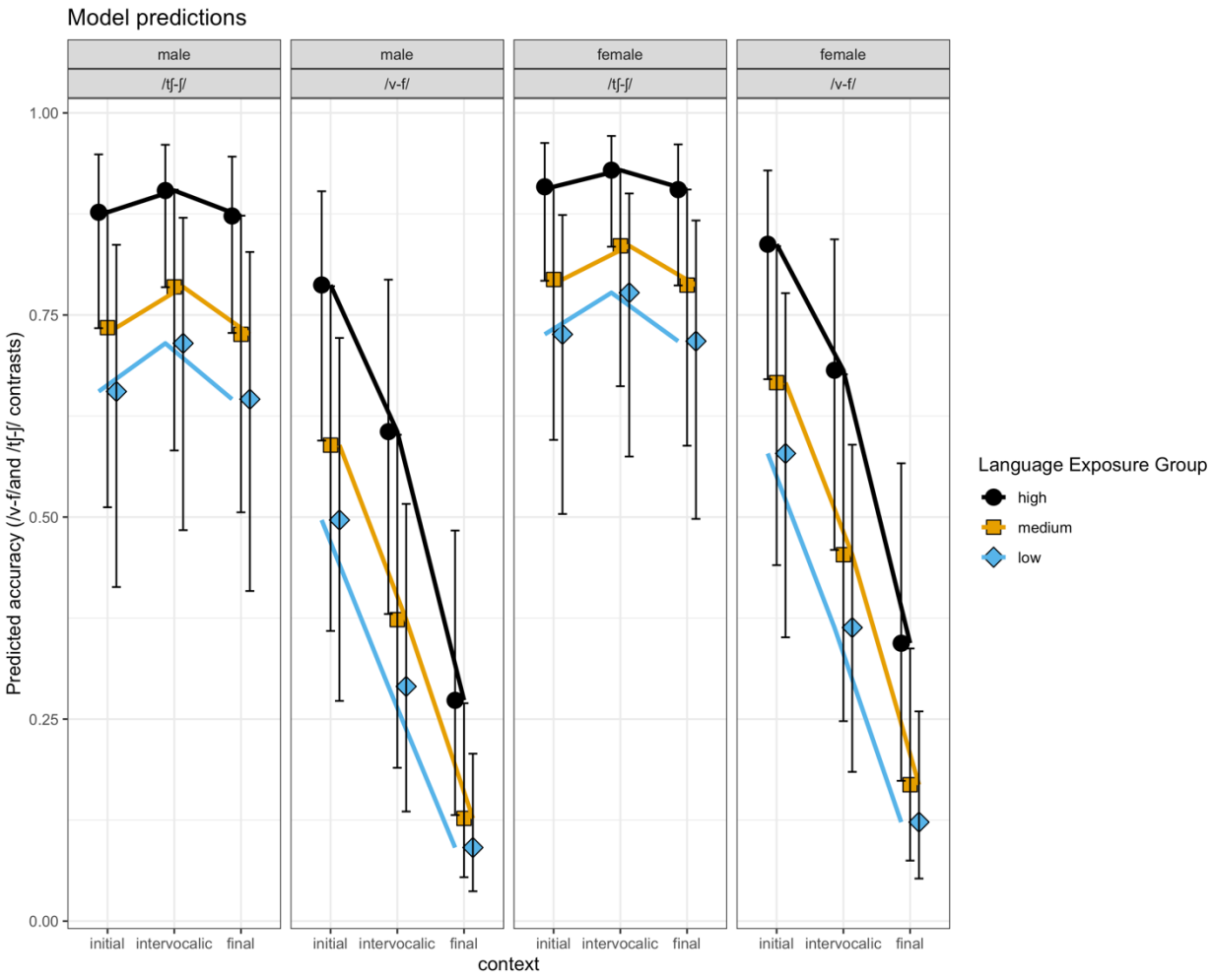


Figure 3.7: Predicted accuracy of the /v-f/ and /tʃ-ʃ/ contrasts by contrast, word position, language exposure group, and gender

3.7.3 Interim Summary

This subsection has presented the discrimination accuracy of the /v-f/ and /tʃ-ʃ/ contrasts across Saudi Arabic listeners, by level of English exposure, word position and gender. The results showed a main effect of English exposure: both contrasts were more likely to be discriminated by the participants in the high exposure group than the participants in the medium and low exposure groups. The results also confirm the effect of word position on the discrimination accuracy of the /v-f/ contrast: discrimination accuracy was lower in word-final position than in word-initial and intervocalic positions. Gender did not affect the discrimination accuracy of either of the two contrasts: there were no significant differences between the male and female participants. The discussion of these results will be taken up in detail in the next

section. Table 3.12 present significant factors affecting the discrimination accuracy of each target contrast.

Table 3.12: Factors influencing the perception accuracy of target /v-f/ and /tʃ-ʃ/ contrasts

	Exposure Group reference = low		Word position reference = final		Gender reference = male
	high	medium	initial	intervocalic	female
/v-f/	***		**	*	
/tʃ-ʃ/	*				

Asterisks refer to levels of significance: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001
Reference refers to the reference level (intercept)

3.8 Discussion

This section aims to discuss potential explanations for the main findings, in relation to PAM/PAM-L2 predictions and previous literature.

Both English /v/ and /f/ were expected to be equally assimilated to the same Arabic phonological category /f/ (single category assimilation - SC). English /tʃ/ and /ʃ/ were also expected to be assimilated to the same native category; however, /ʃ/ is perceived as a better exemplar of the category than /tʃ/ (category goodness assimilation - CG) because participants achieved a high (73.13%) but not perfect level of discrimination accuracy. The reason for this is that /tʃ/ exists in some Arabic dialects.

Overall, the findings of this task demonstrate that the discrimination accuracy of the /v-f/ contrast was lower than the /tʃ-ʃ/ contrast. These results are in agreement with PAM (Best,1995) and SLM (Flege, 1995) which posit that accurate perception of L2 sounds depends on the perceived relationship between phonetic categories in L1 and L2.

The lack of /v/ in Arabic dialects is a possible cause for participants' low discrimination performance on the /v-f/ contrast. This outcome is parallel to the poor discrimination accuracy of English speakers for the /h-ħ/ Arabic contrast, as reported in Shehata (2018). The contrast /h-ħ/ includes one familiar sound /h/ and one unfamiliar sound /ħ/, and the two sounds are acoustically similar. English speakers, regardless of their proficiency in Arabic, had more difficulty in discriminating the /h-ħ/ contrast than the /ħ-ʕ/ contrast which include two unfamiliar sounds.

The discrimination accuracy of the /ʧ-ʃ/ contrast was good but not excellent. This is to be expected since Saudi speakers are familiar with /ʧ/ because it exists as a phoneme category in other Arabic dialects. This finding is consistent with the moderate discrimination accuracy of English speakers for the Zulu velar stops (/k^h/-/k'/), as reported in Best, et al. (2001). The discrimination accuracy for the velar stops was 89.4%, which was significantly lower than lateral fricatives (95%), but still significantly higher than bilabial stops (65.9%).

Interestingly, word position significantly affected the discrimination of the /v-f/ contrast but not of the /ʧ-ʃ/ contrast. The participants, regardless of their level of English exposure, had more difficulty with the /v-f/ contrast in word-final position than in other positions. This finding add further weight to universal markedness claims whereby voiced sounds are universally more difficult to perceive and produce in word-final position (Eckman, 1981).

A possible interpretation of this outcome may also be due to the nature of the stimuli used in the task. In word-final position, /v/ was produced by the English native speaker with partial voicing. The partial voicing of /v/ makes it even more similar to /f/ in Arabic. English voiced fricatives commonly lose their voicing partially or completely in word-final position (e.g., Bayley & Holland, 2014; Ogden, 2009). However, Ogden (2009) demonstrates that voicing is merely one of the several phonetic features that the /v-f/ contrast involves; the friction duration is longer in /f/ than in /v/, and there is less friction noise for /v/ than for /f/. The discrimination of the /v-f/ contrast where voicing, intensity and duration differences occur seems easier than the discrimination where only intensity and duration provide the cues to make a distinction between the two sounds.

Figures 3.8 and 3.9 show both waveforms and spectrograms[6] for the non-words 'bativ' and 'batif' as produced by the English native speaker. For 'bativ', final /v/ was produced with partial voicing. However, other differences were still found: /f/ was produced more loudly than /v/ (i.e., it looks darker on the spectrogram). The average intensity was higher in /f/ (56.59) than in /v/ (51.35). Finally, the friction duration in /f/ (242 ms) was longer than in /v/ (142 ms).

[6] As can be seen in the waveforms and spectrograms, there are also large differences in the duration of the preceding vowel.

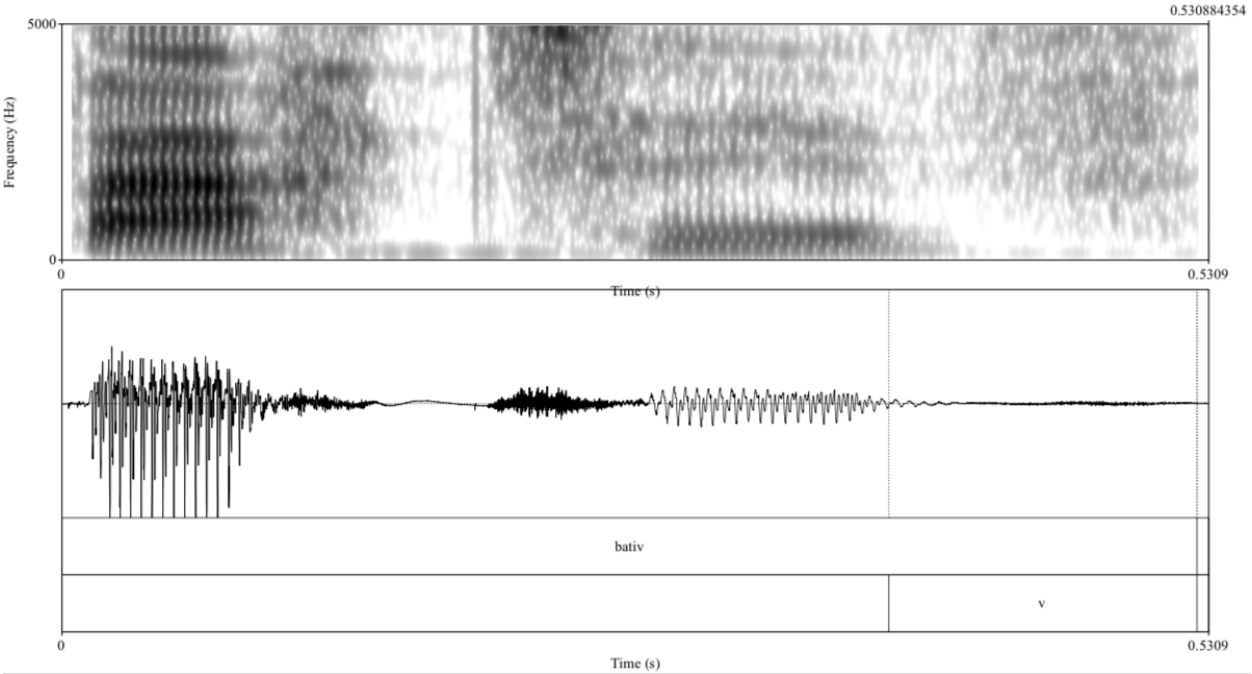


Figure 3.8: 'bativ' as produced by the English native speaker

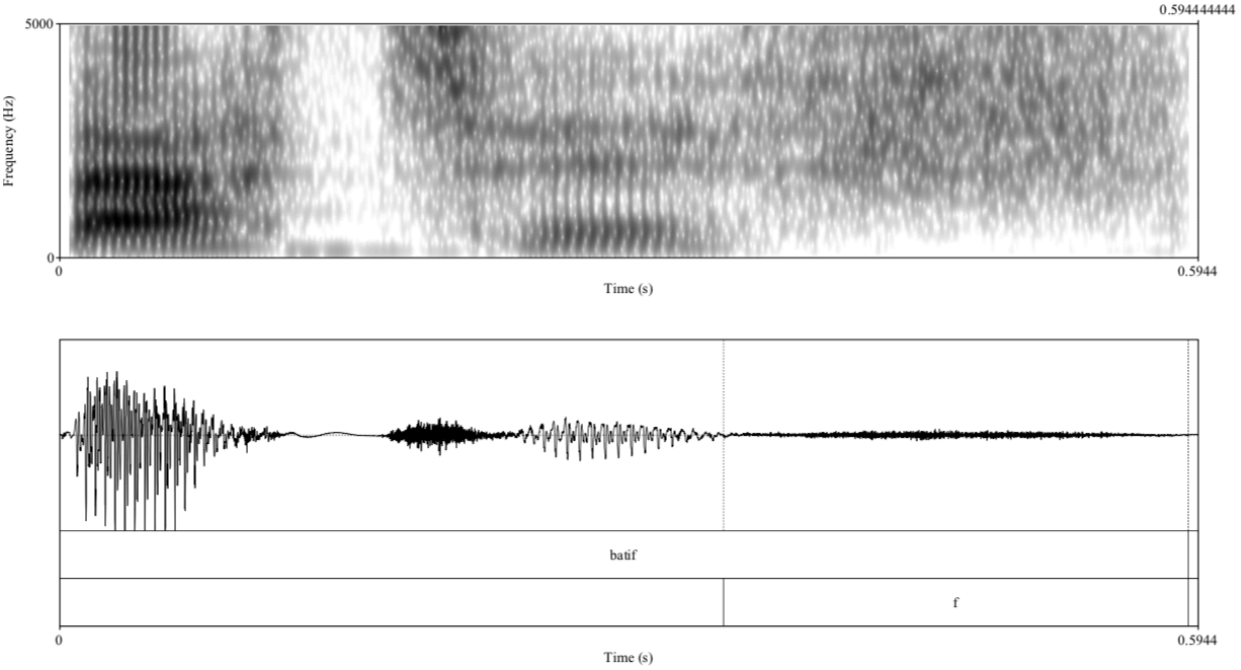


Figure 3.9: 'batif' as produced by the English native speaker

Level of English exposure was a significant predictor of variation in this task; the discrimination accuracy for the two contrasts varied depending on participants' levels of exposure to English. The participants in the high exposure group were more likely to

discriminate both contrasts than the participants in the medium and low exposure groups. This finding is consistent with the PAM-L2 (Best & Tyler, 2007) which assumes that a non-native contrast becomes more discernible with more exposure to the L2. The finding also agrees with SLM-r (Flege & Bohn, 2021) proposing that formation of a new phonetic category depends on the degree of perceived dissimilarity to the closet L1 phonetic category, and the quantity and quality of L2 inputs. Further, the finding is in line with previous literature on loanword phonology showing that relative exposure to the donor language can be a source of variability in the perception of L2 sounds (e.g., Kwon, 2019).

Participants with high levels of English exposure were perhaps more likely to differentiate between /v/ and /f/ because of their frequent exposure to English, in which the two sounds are contrastive. On the other hand, participants with the lower level of English exposure were perhaps less likely to pay attention to the phonetic details of the two sounds because of their frequent exposure to Arabic, in which this voicing contrast does not exist. According to SLM-r (Flege & Bohn, 2021), learners become gradually able to detect differences between L1 sounds and L2 sounds, as they gain more experience in the L2, which will result in forming new phonetic categories for certain L2 sounds.

The significant impact of level of English exposure on the discrimination accuracy of the /ʧ-ʃ/ contrast was unexpected given that all groups have high exposure to Arabic dialects that have this sound. However, this finding is predictable from a previous study (Evans & Alshangiti, 2018) showing that /ʧ/ was poorly identified by low-proficiency Saudi learners.

3.9 Summary

This section provides a summary of the aims, procedure and key findings of the perception task. The main aim of this task was to examine the discrimination accuracy of the /v-f/ and /ʧ-ʃ/ contrasts. The participants were 31 males and 36 females who were recruited based on their expected level of exposure to English (high, medium, and low). An oddity task was conducted online, using *Gorilla* (https://gorilla.sc), to determine the degree of difficulty in discriminating the target contrasts. Participants were asked to decide which stimulus was different, with the option to say that the three auditory stimuli are the same if they do not hear an odd stimulus. Overall, the results revealed the degree of difficulty in perceiving the /v-f/ and /ʧ-ʃ/ contrasts. The /v-f/ contrast was more difficult to discriminate than the /ʧ-ʃ/ contrast, in line with PAM's discriminability ranking for the two contrasts (CG /ʧ-ʃ/ > SC /v-f/). Participants'

discrimination accuracy for the two contrasts varied in terms of their levels of exposure to English. The participants in the high exposure group were more likely to discriminate both contrasts than the participants in the medium and low exposure groups. The discrimination accuracy of both contrasts was not affected by gender: there were no significant differences between the female and male participants. Word position significantly influenced the discrimination accuracy of /v-f/ but not /tʃ-ʃ/: participants were less likely to discriminate the /v-f/ contrast in word-final position than in other positions.

The results of this perception task will be used in the next chapter to generate predictions in the non-word production task. The non-word production task focuses on the relative contribution of acoustic cues and orthographic representation to variable production of target /tʃ/ and /v/, alongside other factors that could influence the production of the target sounds (word position, gender, and level of English exposure).

4 Production of /v/ and /tʃ/ in Target Non-words

4.1 Introduction

Having discussed the perception accuracy of target /v/ and /tʃ/ in the previous chapter, this chapter and the next one will explore their production in non-words and real words. The findings of the last perception task revealed that listeners' perception was more accurate for the /tʃ-ʃ/ contrast than the /v-f/ contrast. Overall, the perception accuracy of the /v-f/ contrast was poor while the perception accuracy of the /tʃ-ʃ/ contrast was good. There was a main effect of English exposure on the perception accuracy of the two contrasts: participants in the high exposure group had the highest perception accuracy. There was also a main effect of word position on the perception accuracy of the /v-f/ contrast: the lowest perception accuracy was in word-final position.

As explained in Chapter 1, the production task[7] reported in this chapter seeks to examine the possible impact of input modality (aural, written and aural-written) and a range of other factors (level of English exposure, word position and gender) on the production of the target sounds. The question addressed in this chapter is as follows: *To what extent do input modality, level of English exposure, word position, and gender account for variability in the production of target /v/ and /tʃ/ in non-words?*

The design of Vendelin and Peperkamp's (2006) experiment was adopted in this study. Vendelin and Peperkamp's (2006) study was carried out to examine French speakers' production of eight English vowels. Participants were tested in aural-only and aural-written conditions. In the former, they listened to non-words pronounced by a native speaker of English. In the latter, they listened to non-words while viewing the written forms on a screen. However, the present production task consists of three conditions: aural-only (auditory input), written-only (orthographic input), and aural-written (auditory-orthographic input). Non-words were used in all three conditions. The same participants in the perception task performed the present production task. To reiterate, this production task preceded the perception task reported in Chapter 3 to avoid drawing participants' attention to the contrasts of interests and prevent them becoming familiar with auditory inputs because the same non-words were used in the two tasks. The

[7] The task was piloted with a small set of participants using the same stimuli and procedure.

perception task was presented first in this thesis because its results were first analysed and used as predictions for analysis of the current production task (see Section 4.5).

The remainder of this chapter discusses participants, (Section 4.2) stimuli (Section 4.3), procedure (Section 4.4), predictions (Section 4.5), data analysis (Section 4.6), number of observations (Section 4.7), the classification process of the target segments (Section 4.8), the acoustic measurements of the target segments (Section 4.9), main results (Section 4.10), discussion (Section 4.11) and summary (Section 4.12).

4.2 Participants

The same participants in the perception task performed the production task in this chapter. As a reminder, the participants were 67 Saudi students at the University of Northern Border, split by gender (31 males and 36 females) and expected level of English exposure (high, medium and low) (for more details, see 3.2).

4.3 Stimuli

The stimuli consisted of 48 items: 24 target items and 24 filler items. The target items were the same CVCVC non-words used in the perception task. The reason for using non-words is that speakers' production of real words may depend on stored representations rather than the input prompt (Eisenbeiss, 2010). Target non-words do not meet criteria outlined by Poplack (2018) for established loanwords (see 2.1.1) because the interest of this task is to examine how production patterns may vary depending on how borrowers are first exposed to a source word. (or, in other words, how much of variability is due to English orthography and/or perception).

To briefly reiterate, the selected vowels are the near-low front vowel /æ/ in the first syllable and the near-high front vowel /ɪ/ in the second syllable. The target items were selected based on naturalness ratings of English native speakers (see Appendix A for more details of the naturalness rating). The motivation for the choice of this syllable pattern and the selected vowels is detailed in Chapter 3 (see 3.3). The filler items were also non-words, but without /v/ or /tʃ/, so that the target sounds would not stand out. Out of 60 possible candidates, the final list of filler items was selected based on a native English speaker's judgment (see Appendix B for the list of filler items). The native speaker who selected the filler items speaks the standard British accent and has phonetic training.

In aural-only and aural-written conditions, the same female native speaker who produced the stimuli in the previous perception task also rerecorded the stimuli in the present task. To

reiterate, the female native speaker is originally from the West Midlands and she was 55 years old at the time of recording. All target and filler items were recorded with a standard British accent. Marantz PMD-610 recorder and Shure SM10A microphone were used to record target and filler items in WAV format with the default sampling frequency 44100Hz. The target and filler items normalised to the same intensity level (75 dB).

In written-only and aural-written conditions, non-words were written in English orthography. Target /v/ was spelled with the diagraph <v> and target /tʃ/ was spelled with the diagraph <ch>. /æ/ was spelled with the diagraph <a> and /ɪ/ with the diagraph <i>.

Table 4.1: Target items in the non-word production task

/v/	/væpɪt/ <vapit>	/ɹævɪd/ <ravid>	/bælɪv/ <baliv>
	/væɪɪʃ/ <varish>	/sævɪʃ/ <savish>	/bæɹɪv/ <bariv>
	/vænɪt/ <vanit>	/nævɪʃ/ <navish>	/bætɪv/ <bativ>
	/vælɪt/ <valit>	/bævɪn/ <bavin>	/gælɪv/ <galiv>
/tʃ/	/tʃæɹɪs/ <charis>	/ɹætʃɪm/ <rachim>	/mænɪtʃ/ <manich>
	/tʃæsɪt/ <chasit>	/ɹætʃɪn/ <rachin>	/tænɪtʃ/ <tanich>
	/tʃænɪt/ <chanit>	/ɹætʃɪl/ <rachil>	/lætɪtʃ/ <latich>
	/tʃælɪt/ <chalit>	/bætʃɪn/ <bachin>	/pænɪtʃ/ <panich>

4.4 Procedure

The present task tested Saudi Arabic speakers in three conditions: aural-only, written-only and aural-written. This design was used to measure variability in participants' production of target /v/ and /tʃ/ when they are exposed to different input sources. Participants were asked to produce each target and filler item twice. Due to the strict COVID-19 ban on in-person data collection at the time, Gorilla (https://gorilla.sc) was used to present stimuli and record responses. The reasons for selecting Gorilla to design this task are explained in the last perception task. The procedures for each condition are presented below:

a) Aural-only condition (Listen-Say): Participants were asked to listen to each non-word and then produce it twice in an Arabic frame sentence presented on the screen. The audio file was played automatically when the page finished loading.

b) Written-only condition (Read-Say): Participants were asked to read each non-word and then produce it twice in an Arabic frame sentence. Orthographic representations of non-words appeared immediately on the screen after the page finished loading.

c) Aural-written condition (Listen-Read-Say): Participants were requested to listen to each non-word with its accompanying orthographic representation and then produce it twice in an Arabic frame sentence. The auditory and orthographic representations of non-words appeared simultaneously on the screen. Like aural-only condition, the audio file was played automatically when the page finished loading.

Given that this task was conducted online, it was considered that participants might be interrupted (e.g., phone call or loss of Internet connection) while listening to auditory inputs in aural and aural written conditions. Therefore, it was decided to give participants the opportunity to play sound files again multiple times. In aural-only and aural-written conditions, the audio file was played again in 10.03% of the trials for target /v/ (79 out 787) and in 11.43% of the trials for target /tʃ/ (90 out 787). Figure 4.1 shows a screenshot of the aural-written condition as seen by participants.

Figure 4.1: Screenshot of aural-written condition

The order of aural-only, written only, and aural-written conditions was counterbalanced. Table 4.2 illustrates how the three conditions were counterbalanced.

Table 4.2: Illustration of counterbalancing the order of non-word conditions

Participant	Conditions		
FH01	aural-only	written-only	aural-written
FH02	aural-only	aural-written	written-only
FH03	written-only	aural-only	aural-written
FH04	written-only	aural-written	aural-only
FH05	aural-written	aural-only	written-only
FH06	aural-written	written-only	aural-only

In each condition, participants were requested to produce six target words twice: three non-words containing /v/ and three non-words containing /tʃ/ in the three different positions (initial/ intervocalic/ final). That is, 36 tokens were expected to be collected from each participant (2 target sounds × 3 non-words × 2 repetitions × 3 conditions). As shown in Table 4.3, a subset of non-words was given randomly to every participant, with the rule not to see or hear the same set of non-words in each condition. For example, a non-word 'vapit' was produced by a participant in aural-only condition and by another participant in written-only condition.

Table 4.3: Illustration of how target non-words were randomly distributed across conditions

Participant	aural-only	written-only	aural-written
FH01	vapit/ ravid/ baliv	varish/ savish/ bariv	vanit/ navish/ bativ
FH02	varish/ savish/ bariv	vanit/ navish/ bativ	valit/ bavin/ galiv
FH03	vanit/ navish/ bativ	valit/ bavin/ galiv	vapit/ ravid/ baliv
FH04	valit/ bavin/ galiv	vapit/ ravid/ baliv	varish/ savish/ bariv

Arabic frame sentences were randomised for each non-word in the three conditions. Participants did not hear or read target non-words in the frame sentences. In written-only and aural-written conditions, the frame sentences appeared below target non-words. The participants were told that their task would be to insert novel company names in Arabic sentences. The frame sentences were of similar structure and length, each with the same beginning but with different

endings. The target non-words were recorded after the word 'company' [ʃærɪkət] <sharikat>[8] at the beginning of each sentence. The reason for giving the word 'company' was to make participants produce target non-words as if they were real English names because some participants may have a good level of English. The frame sentences are presented in Table 4.4.

Table 4.4: Frame sentences used for target non-words

1	[ʃærɪkət _____ mɪn ʔkbər ʃærɪkaat təsˤni:ʕ ʔlmæwæd ʔlɣiθæʔjəh.] '_____is one of the large food manufacturing companies.'
2	[ʃærɪkət _____ mɪn ʔkbər ʃærɪkæt mu:stəhdˤræt ʔtəjmi:l.] '_____is one of the large cosmetics companies.'
3	[ʃærɪkət _____ mɪn ʔkbər ʃærɪkæt fi: sˤɪnæʕət ʔlɣəsəlæt.] '_____is one of the large washing machine companies.'
4	[ʃærɪkət _____ mɪn ʔkbər ʃærɪkæt ʔtɪlɪfɪzju:nat.] '_____is one of the large TV companies.'
5	[ʃærɪkət _____ mɪn ʔkbər ʃærɪkaat fi: sˤɪnæʕət ʔlʔləktru:nəjæt.] '_____is one of the large electronics manufacturing companies'

Participants accessed the task through a web link sent to them via the *WhatsApp* application. To ensure anonymity, participants were requested to use the same anonymised codes that were given to them by the researcher in both perception and production tasks. For each condition, instructions were given in Arabic (Appendix D). It was explained to participants that they will hear and/or read international brand names and they should produce these brand names aloud within given sentences. As in the previous perception task, participants were instructed to finish this task along with the production task in Chapter 5 within the expected time (45-60 minutes) in a quiet room. All participants were also instructed to perform the task using a computer or laptop and any type of headphones or earphones that they had access to. Participants were not allowed to start this task and the production task reported in Chapter 5 without entering their codes, reading the information sheet, and then signing the consent form. Two questions

[8] The romanised transliteration in this example was adopted from the Intonational Variation in Arabic Corpus (Hellmuth & Almbark, 2017).

appeared asking about their gender and academic program. The reason for asking these two questions was to confirm that participants were given the correct codes (i.e., Based on their gender and expected level of English exposure).

At the outset, participants were asked to create a short recording, over their computers with Gorilla, and then play it back to test whether their microphones were working or not. Three practice trials with distractor items were given prior to the testing of each condition to ensure that participants understood the procedure. The task was self-paced. Participants were given as much time as they needed to complete the task; they were able to control the start of a new trial by pressing the 'next' button. Since the study was not conducted under the researcher's supervision, seven participants in fact completed the task over a number of days and thirteen participants spent more than one hour to finish the task in one day. Appendix F contains a histogram illustrating the distribution of time taken by participants to complete the task. There were no restrictions placed on the type of device for participating. Some participants had to use other devices (e.g., phones and tablets). It was difficult to ensure that all participants used headphones to listen to the stimuli, so participants were not asked if they used headphones when listening to the stimuli in aural and aural-written conditions. The study did not apply strict inclusion criteria because recruiting participants who were willing to complete a full sequence of online tasks was difficult. Participants were only excluded if they did not complete the task.

4.5 Predictions

This section explains predicted answers to four specific questions derived from the main question. To reiterate, the main question asked in this task is: *to what extent do input modality, level of English exposure, word position, and gender account for variability in the production of target /tʃ/ and /v/ in non-words?*

The first specific question is: *is the variable production of target /v/ and /tʃ/ affected by input modality?* I hypothesise that the likelihood of adaptation vs. importation may be affected by input modality. Before moving on to the predictions, I will briefly review some of the findings from previous research regarding the role of perception and orthography on the production of novel forms.

While the link between L2 perception and production is still a matter of debate, it is often believed that the two are related (PAM-L2; Best & Tyler, 2007; SLM; Flege, 1995). Some studies on loanword adaptation have shown a positive link between perception and production

(e.g., Boersma & Hamann, 2009), while others have not (e.g., Martin et al., 2022). Regarding orthography, some L2 studies found that orthographic information may help in facilitating non-native production (Nimz and Khattab, 2020); however, other studies indicated that orthography may have negative impact on non-native production, especially on speakers whose L1 has a transparent writing system. If the L2 writing system is less transparent than that of the L1, speakers may produce L2 forms based on their L1 knowledge of letter-sound correspondences. They may result in them producing a sound other than what is in written input (e.g., Vokic, 2011; Bassetti, 2017). Some studies showed that orthographic information could improve the production accuracy of non-native forms when combined with auditory information (e.g., Davidson, 2010) while other studies have not (e.g., Bassetti and Atkinson, 2015; Vendelin and Peperkamp, 2006), suggesting that the effect of input modality can vary for different non-native sounds in different languages.

The studies discussed above have informed our predictions about how input modality may affect the production of target /v/ and /ʧ/. In our case, we expect that Arabic speakers' production of target /v/ and /ʧ/ may be affected by input modality and this effect may vary for the two sounds. Recall that the results of the perception task showed that the discrimination accuracy of the /v-f/ contrast was poor while the discrimination accuracy of the /ʧ-ʃ/ contrast was good. These results may suggest that auditory information may be less helpful to make a distinction between /v-f/ than /ʧ-ʃ/. In contrast, orthographic information may be less helpful in making a distinction between /ʧ-ʃ/ than /v-f/ because Arabic speakers may produce the target sounds based on their L1 knowledge of letter-to-sound correspondences. Arabic orthography is transparent, meaning one letter represents only one sound. In English script, /v/ and /f/ are often represented by two different graphemes <v> and <f>. However, both /ʧ/ and /ʃ/ can be represented by the same grapheme <ch>.

Based on the hypothesis and evidence from the previous perception task, the following predictions can be put forward:

a) If Saudi Arabic speakers are exposed to written and auditory inputs, /v/ is more likely to be imported as [v] in response to written inputs than auditory inputs because /v/ does not involve salient acoustic features for Arabic speakers.

b) If Saudi Arabic speakers are exposed to written and auditory inputs, /tʃ/ is more likely to be imported as [tʃ] in response to auditory inputs than written inputs because the grapheme <ch> represents /tʃ/, /ʃ/ and /k/ in English.

The second specific question is: *is the variable production of /v/ and /tʃ/ affected by level of English exposure?* We hypothesise that the likelihood of adaptation vs. importation may be affected by level of English exposure. The previous perception task showed that participants with high levels of English exposure were more likely to discriminate the target contrasts than participants with low levels of English exposure. Additionally, in the realm of loanword phonology, evidence suggests that the higher one's level of exposure to the source language is, the more likely one is to import rather than adapt novel sounds in loanwords (e.g, Kang, 2017; Kwon 2019; Poplack, 2018). Based on the hypothesis, the following prediction can be put forward:

c) Saudi Arabic speakers with high levels of exposure to English are more likely to produce the imported sounds [v] and [tʃ] than speakers with medium and low levels of exposure to English.

The third specific question is: *is the variable production of /v/ and /tʃ/ affected by word position?* We hypothesise that the likelihood of adaptation vs. importation could be affected by word-final position. The findings of the perception task showed that word position affected the perception accuracy of the /v-f/ contrast, with lowest perception accuracy in word-final position. Previous literature suggests that the substitution rate for non-native sounds can be higher in some phonetic contexts than others due to universal markedness (e.g., Eckman, 1977, 1984). In word-final position, cross-linguistically, voiced obstruents are more marked than their voiceless counterparts and affricates are more marked than fricatives (Eckman & Iverson 1994). Therefore, we expect that the substitution of target /v/ and /tʃ/ for the Arabic sounds /f/ and /ʃ/ might be higher in word-final position than in other positions. Based on the hypothesis, the following prediction can be made:

d) For target /v/ and /tʃ/ elicited in word-final position, Saudi Arabic speakers are more likely to produce the adapted sounds [f] and [ʃ] than the imported sounds [v] and [tʃ] because /f/ and /ʃ/ are less marked in word-final position.

The fourth specific question is: *is the variable production of /v/ and /tʃ/ affected by gender?* We hypothesised that the likelihood of adaptation vs. importation could be affected by

gender. Edwards (2008) claims that gender does not have explanatory power concerning accuracy in L2 production. However, social factors, including gender, that come into play in the L1 may arise in producing L2 forms (e.g., Adamson and Regan, 1991). Previous sociolinguistic studies in various Arabic dialects showed that women tend to adopt prestigious novel forms more than men (e.g, Almuhannadi, 1991; Assiri, 2014; Omari & Van Herk, 2016). If this scenario also holds in Saudi Arabic, we expect that females will be more likely to produce the imported sounds [v] and [tʃ] than males because English is a prestigious language in Saudi Arabia. Based on the hypothesis, the following prediction can be put forward:

e) Females are more likely to produce the imported sounds [v] and [tʃ] than males because English is a prestigious language in Saudi Arabia.

4.6 Data Analysis

This section is divided into two subsections. The first subsection discusses the acoustic analysis and the second subsection discusses the statistical analyses.

4.6.1 Acoustic Analysis

The acoustic analysis in this chapter serves to corroborate the results of manual labelling of target sounds. Praat (Boersma & Weenik, 2016) was used in two ways. The first was to classify the target sounds as will be explained, in detail, in section 4.7. The tokens were segmented into two tiers using a Praat script to assist with labelling; the first tier contained the target words, and the second tier contained the target sounds. The second was to extract acoustic measurements of the target segments using a Praat script.

Several measurements were made and compared across the three word positions. As cues for the distinction between /v/ and /f/, friction duration, intensity, and the centre of gravity (COG) were measured. Literature on acoustic features of English fricatives demonstrates that voiced fricatives have shorter friction duration, lower intensities, and lower COG values than voiceless fricatives (e.g., Ogden, 2009; Jongman et al., 2000). For the duration, the onset and offset of the friction noise were used as benchmarks for determining the beginning and end of a segment. Intensity was measured by calculating the average intensity throughout the segment. As cues for the distinction between /tʃ/ and /ʃ/, friction duration, and the amplitude rise time were measured (Hayward, 2013). According to Hayward (2013), fricatives have longer friction durations and rise time than affricates. The duration was measured as the time between the onset to offset of the friction noise. Rise time was measured as the time between the beginning of the

friction and its maximum amplitude (Hayward, 2013). The stop portion of /tʃ/ was not measured because its acoustic annotation, as will be explained in the next section, was somewhat difficult in word-initial position. It is also worth noting that the burst of /tʃ/ was not measured as a part of the friction. According to Hayward (2013), the burst release creates a barrier between the stop silence and the friction portion. Therefore, the burst was not included in the friction portion.

4.6.2 Statistical Analyses

R software (R Core Team, 2019) was used to perform two types of statistical analysis: descriptive and inferential. For the acoustic analysis, linear mixed effects models were used for each parameter with a parallel structure: acoustic parameter ~ segment + (1 |speaker) + (1 |word). The model structure includes the target sounds as fixed effects and random intercepts for speaker and word. This statistical method was used because the dependent variable, the acoustic parameter, is continuous. The *LmerTest* package (Kuznetsova et al., 2017) was used to calculate p-values in the linear mixed effects models.

Main results were summarised and plotted using the *tidyverse* package (Wickham, 2016). Following the descriptive analysis, inferential statistics with mixed-effects models were used with the *lme4* package (Bates et al. 2015) to consider random effects that account for repeated measures of speakers or items (Winter, 2013, 2019). Mixed effects logistic regression models were run for the target sounds to estimate the probability of the binary outcome (i.e., imported sounds [v] and [tʃ] vs. adapted sounds [f] and [tʃ]) based on different predictors. The target sounds were examined with the same predictor variables: input modality, level of English exposure, word position, and gender.

In cases of non-convergence, the *all_fit* function from the *afex* package (Singmann et al. 2018) was used to test a variety of optimizers for convergence. However, if convergence was still not achieved, the solution was to separate the intercept and slope using zero character as in (1| speaker) (0 + context | speaker) (Winter, 2019). If convergence was again not achieved, backward elimination was run to simplify the model. The model was compared by including and excluding each random slope. That is, convergence was tested with and without each random slope. The inclusion of interactions was assessed via model comparison using the *anova* function. The simple model, without interactions, was adopted if the interactions were not significant and there were no significant differences between the models ($p >. 05$). All categorical variables were converted into numbers using dummy coding. For the target sounds, 0

corresponded to [v] and [tʃ], and 1 corresponded to [f] and [ʃ]. That is, positive values suggested a higher probability for the adapted sounds than the imported sounds. The pairwise predictions of constructed models were estimated using the *emmeans* package (Lenth, 2021). R codes used in the task are given in a footnote in the next section.

4.7 Number of Observations

In total, 2,412 tokens (9 non-words × two sounds × two times × 67 participants) were expected to be collected. However, 2,354 tokens were collected (1,182 tokens for /v/ and 1,172 tokens for /tʃ/) because the participants did not produce all the target non-words (i.e., some participants moved forward to the next page without recording the target non-word). Only 1,130 tokens for /v/ and 1,103 tokens for /tʃ/ were included. For the sake of this study, 52 tokens for /v/ (4.40%) and 69 tokens for /tʃ/ (5.89%) were excluded because they underwent other consonantal substitutions or syllabic modifications, such as epenthesis and deletion.

4.8 Classification of Imported and Adapted Sounds

The terms *importation* and *adaptation* were clarified and explained in detail in Chapter 2. The current study focuses only on [f] and [ʃ] as adapted sounds for /v/ and /tʃ/ to discuss the likelihood of importing and adapting novel sounds. In this task, a production of [v] or [tʃ] is treated as a case of being imported from English whereas a production of [f] or [ʃ] is a case of being adapted to the native phonology. [9]

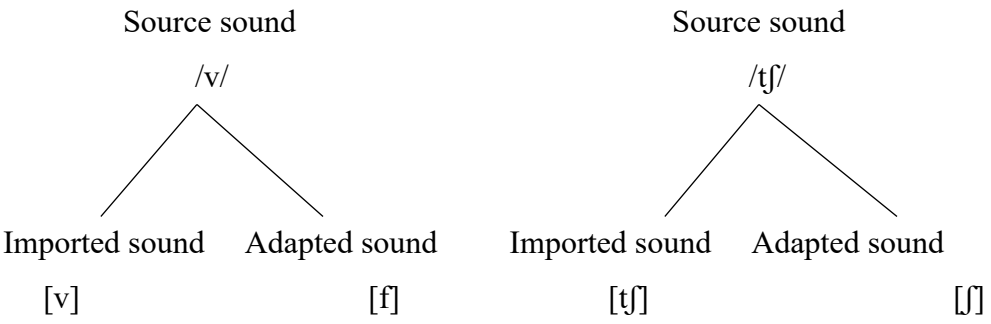


Figure 4.2: *Imported sounds vs. adapted sounds*

The researcher's classification process was based on acoustic representation and auditory judgment. The target sounds were annotated manually. Classifying the realizations of /v/ and /tʃ/

[9]We should note, to avoid any confusion, that square brackets were always used in this thesis to refer to speakers' production (i.e., either imported sounds or adapted sounds).

into imported sounds and adapted sounds in both non-words and real words was mainly based on observed acoustic features. Reliance on acoustic features wherever possible was used to increase the objectivity of the classification process. As explained in detail in 4.81 and 4.82, pure auditory analysis (impressionistic judgement) was necessary in cases where it was not possible to classify the sounds acoustically. For example, the distinction of /v/ and /f/ is acoustically not clear in word-final position because final /v/ can be fully devoiced. In this case, target /v/ and /f/ were identified impressionistically.

To validate the impressionistic labelling in target non-words in this chapter and real words in Chapter 5, an interrater reliability test was conducted using Cohen's Kappa. Following Mackey and Gass (2015), 25% of impressionistically identified tokens were selected randomly to be labelled by a second Arabic trained analyst. The agreement rate was 100% (Cohen's kappa = 1) for /v/ tokens (n=83) which is perfect and 84.6% for /tʃ/ tokens (n=52) (Cohen's kappa = 0.66) which is considered as very good (Roever and Phakiti, 2017). The sounds that were identified acoustically and impressionistically were labelled using the same categories (i.e., either imported sound or adapted sound) to yield one dependent variable with two levels for input to mixed effects logistic regression. The next subsections describe how and when acoustic representations and auditory judgment were used to classify tokens as imported or adapted.

4.8.1 Classification of /v/ into [v] and [f]

English is traditionally described as having a phonological contrast between voiced and voiceless fricatives. The major phonetic cue to differentiate English voiced and voiceless fricatives is the presence or absence of vocal fold vibration characterised by a voice bar at the bottom of the spectrogram. In a waveform, voiced fricatives are typically characterised by the overlap of two components: periodicity (voicing) and noise (friction). In contrast, voiceless fricatives are characterised only by the friction component (Ogden, 2009*)*. Besides voicing, English voiced fricatives also have lower intensity and shorter duration than voiceless fricatives (Hayward, 2013; Ogden, 2009). The contrast is thus realized through a range of phonetic properties shown in Table 4.5.

Table 4.5: The phonetic cues for voiced and voiceless fricatives (Ogden, 2009)

Symbols	Phonetic Properties
/ v/	Voicing overlaps with friction Low intensity The duration of friction is short
/f/	Only friction High intensity The duration of friction is long

Following previous literature, /v/ and /f/ were acoustically classified with reference to both waveforms and spectrograms in Praat (Boersma & Weenik, 2016). Voicing was used as the main cue to make an acoustic judgement. Fully voiced segments overlapping with friction were always acoustically annotated as [v]. Fully voiceless segments with only friction noise were always acoustically annotated as [f].

Figures 4.3 and 4.4 show two waveforms for the non-word ‘vanit’ as produced by two male speakers with high and low levels of exposure to English. Figure 4.3 shows the imported sound [v] while Figure 4.4 shows the adapted sound [f]. Voicing can be seen throughout the sound in Figure 4.3, but the sound in Figure 4.4 has only friction noise.

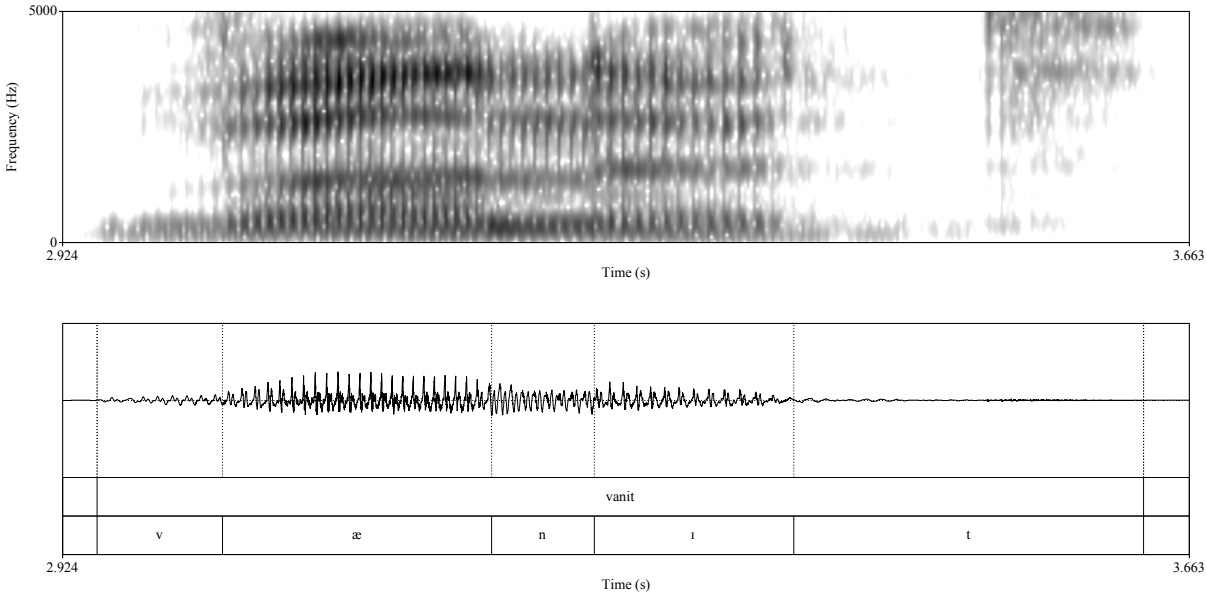


Figure 4.3: Sample waveform and spectrogram of the imported sound [v] (MH08_O_vanit2)

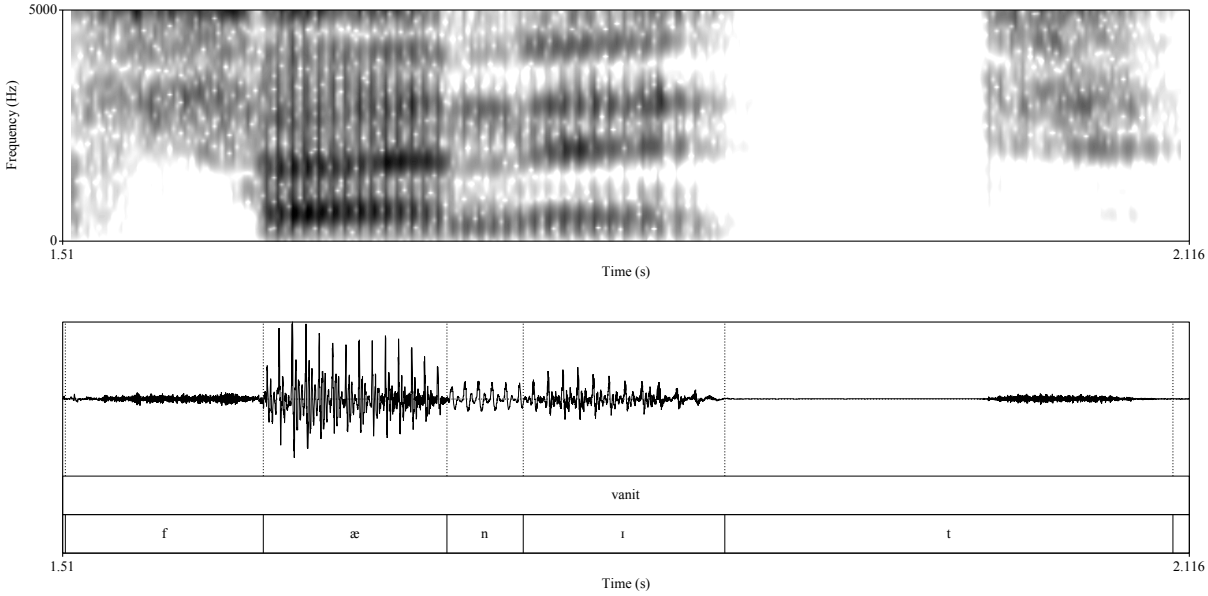


Figure 4.4: Sample waveform and spectrogram of the adapted sound [f] (ML02_O_vanit1)

At first glance, classifying /v/ into [v] or [f] might appear straightforward. However, the classification process was not always simple. Partial devoicing is common in utterance-final position, as the following silence may lead to an early offset of vocal fold vibration. In utterance-initial position, the preceding silence may lead to late onset of vocal fold vibration (Laver, 2003).

Similarly, voiced segments display partial devoicing in syllable-initial and syllable-final positions. For example, in English, the early portion of a voiced fricative is likely to be voiceless in syllable-initial position. Correspondingly, the latter portion of a voiced fricative is likely to be voiceless in syllable-final position. In intervocalic position, the middle part of an English voiced fricative can also be quiet or voiceless. That is, voicing is retained at the end of the fricative construction only in syllable-initial and intervocalic positions (Cruttenden & Gimson, 2001; Ogden, 2009). One explanation for this partial devoicing in different syllable positions is that friction and voicing are difficult to produce at the same time (Ohala, 1983). Voiceless fricatives are also prone to partial voicing intervocalically. That is, some voicing may come into a voiceless fricative from a neighbouring segment. The process of voicing can also be triggered at word boundaries. In discussing phonation of English voiceless stops and fricatives in connected speech, Davidson (2018) indicated that partial voicing may spill over into a fricative from a preceding segment.

Due to these issues, acoustic criteria were tailored to different contexts where partial voicing exists. In word-initial position, segments that have partial voicing which overlaps with the latter portion of friction are voiced (Cruttenden, 2001). However, in some cases, it was still justified to consider a segment as [v] if the latter portion was voiceless as in Figure 4.5. The waveform shows the non-word ‘vapit’ as produced by a female speaker with a high level of exposure to English. The early portion of friction in initial [v] is voiced but the latter portion is voiceless. In this case, we assume that voicing did not spread into the fricative from the preceding segment because it was a voiceless [t] (the final segment in the preceding word, [ʃælɪkət], in the frame sentence). Therefore, it was acoustically annotated as the imported sound [v].

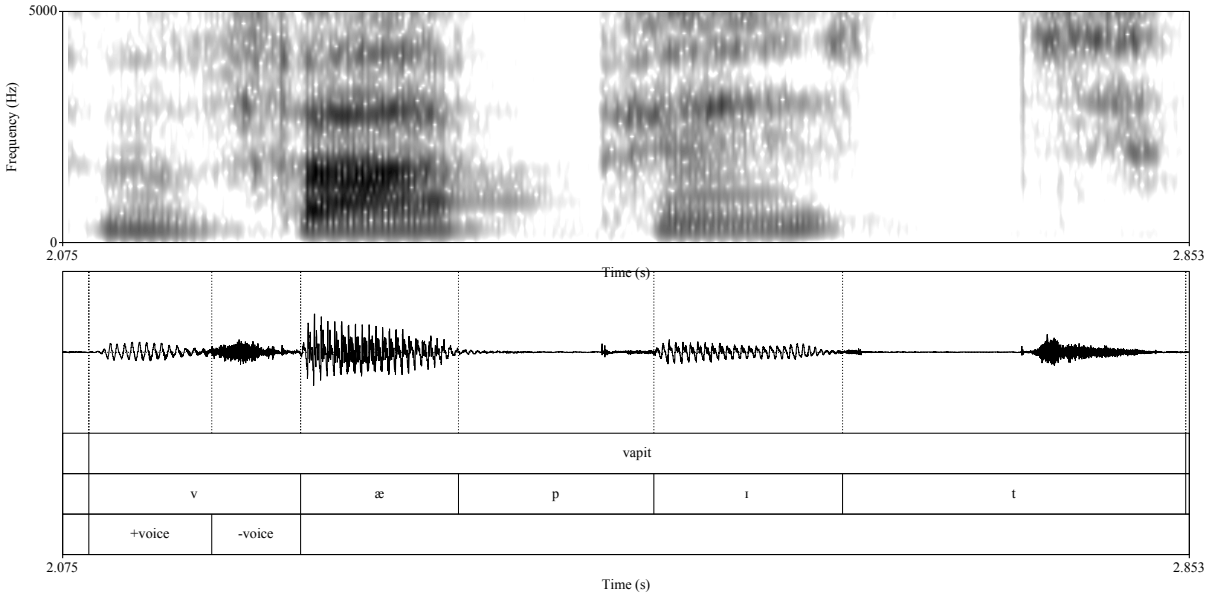


Figure 4.5: Sample waveform and spectrogram of partially devoiced [v] (FH03_W_vapit2)

In intervocalic position, partially voiced segments in which voicing overlaps with friction at the beginning and the end of the segment were acoustically annotated as [v] (Ogden, 2009). However, partially voiced segments were acoustically annotated as [f] if voicing was not retained at the end of the period of friction. Figures 4.6 and 4.7 show two cases of partial voicing in intervocalic position by female speakers with high and low levels of exposure to English. Figure 4.6 shows voicing at the beginning and the end of friction. In contrast, Figure 4.7 shows voicing only present at the beginning of friction.

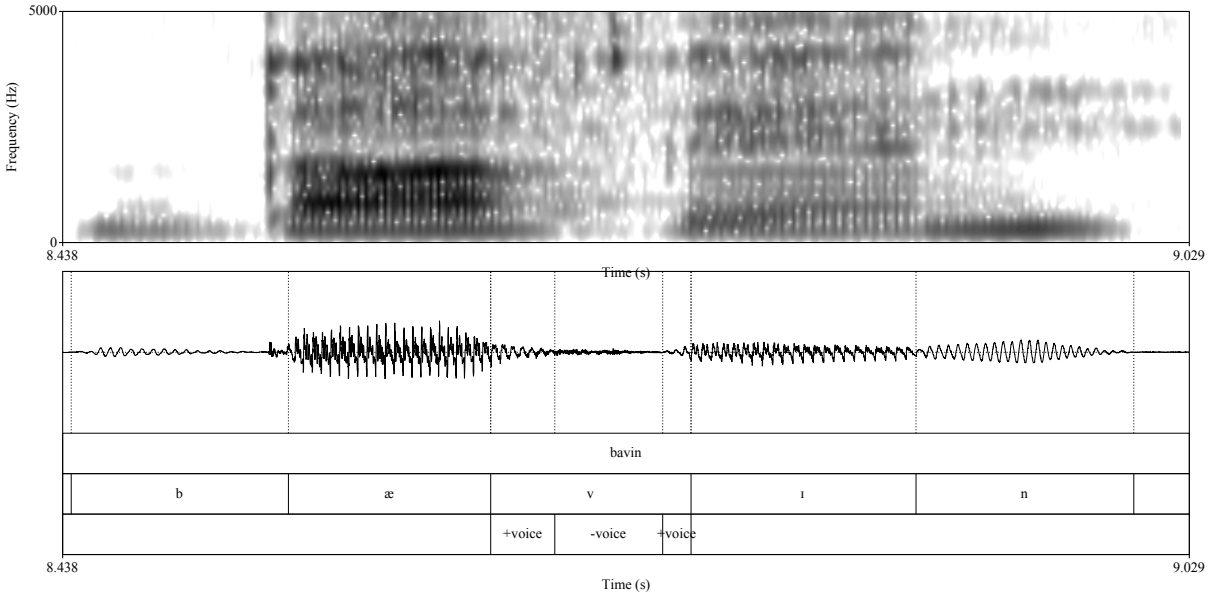


Figure 4.6: Sample waveform and spectrogram of partially devoiced [v] (FH03_OW_bavin1)

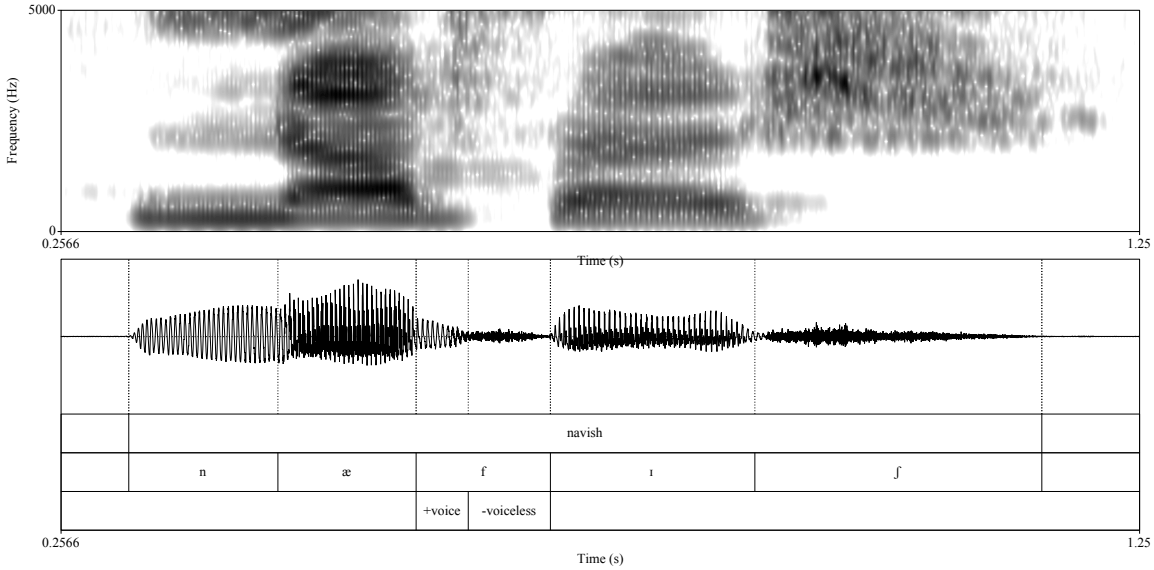


Figure 4.7: Sample waveform and spectrogram of partially voiced [f] (FL08_OW_navish1)

In word-final position, segments with any amount of voicing overlapping with friction were acoustically annotated as [v]. [v] and [f] were always analysed impressionistically if the sound was fully voiceless because the voiced fricative can be completely devoiced in this position (e.g., Bayley & Holland, 2014; Ogden, 2009). Figure 4.8 shows a waveform of the non-

word ‘baliv’ as produced by a female speaker with low exposure to English. The presence of voicing in final [v] can be seen at the beginning of the friction (lasting 68 ms).

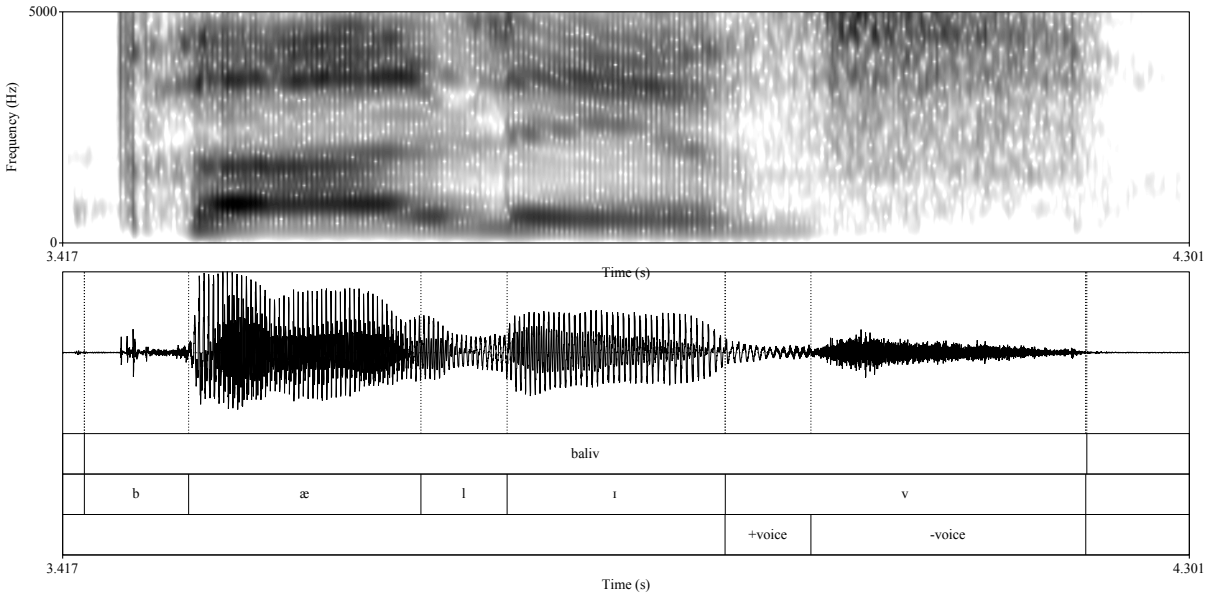


Figure 4.8: Sample waveform and spectrogram of partially devoiced [v] (FL11_W_bativ2)

To sum up, the classification process for identification of [v] and [f] in all production data is summarized in Table 4.6.

Table 4.6: Classification criteria for [v] and [f]

Annotation	Criteria
Acoustically annotated as [v]	• The segment had full voicing overlapping with friction (Ogden, 2009) • The segment had partial voicing in the later portion of friction in word- initial and intervocalic positions. An exception was made for segments that had no voicing in the early portion of friction in word-initial position if preceded by a voiceless segment (Cruttenden, 2001; Ogden, 2009). • The segment had partial voicing in the early portion of friction in word-final position (Cruttenden, 2001)
Acoustically annotated as [f]	• The segment had only friction without voicing (Ogden, 2009) • The segment had partial voicing in the early portion of friction in word-initial and intervocalic positions (Cruttenden, 2001; Ogden, 2009).
Impressionistically annotated as either [v] or [f]	• The segment had only friction without voicing in word-final position (Cruttenden, 2001; Ogden, 2009).

Table 4.7 shows the acoustic and impressionistic annotations for [v] and [f] in the different word positions.

Table 4.7: Total number of annotations for [v] and [f] in target non-words

Annotation	Initial	Intervocalic	Final	Total
Acoustically annotated as [v]	240	218	164	622 (52.62%)
Acoustically annotated as [f]	136	167	None	303 (25.63%)
Impressionistically annotated as [f]	None	None	205	205 (17.34%)
Others (other sounds)				52 (4.40%)

4.8.2 Classification of /tʃ/ into [tʃ] and [ʃ]

English affricates are a combination of two sounds: a plosive followed by a fricative. The major phonetic cue to differentiate fricatives and affricates is the presence or absence of a plosive portion. That is, /tʃ/ can be characterised by three components: silence, burst release and friction (Hayward, 2013). The burst can be seen as a spike in a waveform and a vertical line in a spectrogram. Acoustically, the friction portion in /tʃ/ is different from that in /ʃ/. The phonetic distinctions between the two sounds are explained in Table 4.8.

Table 4.8: The phonetic cues for fricatives and affricates (Hayward, 2013)

Symbols	Phonetic Properties
/ʃ/	Only friction The friction noise increases gradually The duration of friction is long
/tʃ/	Closure followed by a burst and friction The friction noise increases rapidly The duration of friction is short

Following previous literature, /tʃ/ and /ʃ/ were acoustically classified with reference to both waveforms and spectrograms in Praat (Boersma & Weenik, 2016). Segments that have a period of silence followed by friction with a burst between the two were acoustically annotated as [tʃ]. Segments that have only friction without a period of silence or burst were acoustically annotated as [ʃ]. Figures 4.9 and 4.10 show waveforms for the non-word ‘panich’ as produced by two different female speakers with low level of exposure to English. Figure 4.9 shows an example of the imported segment [tʃ] while Figure 4.10 shows the adapted segment [ʃ].

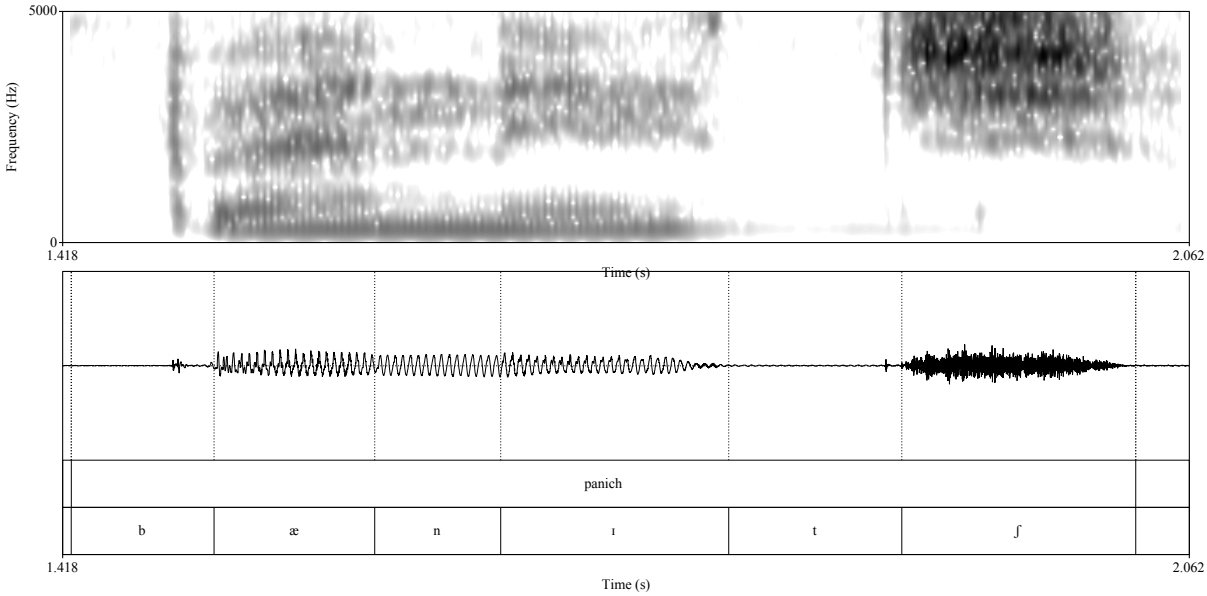


Figure 4.9: Sample waveform and spectrogram of the imported segment [tʃ] (FL03_OW_panich1)

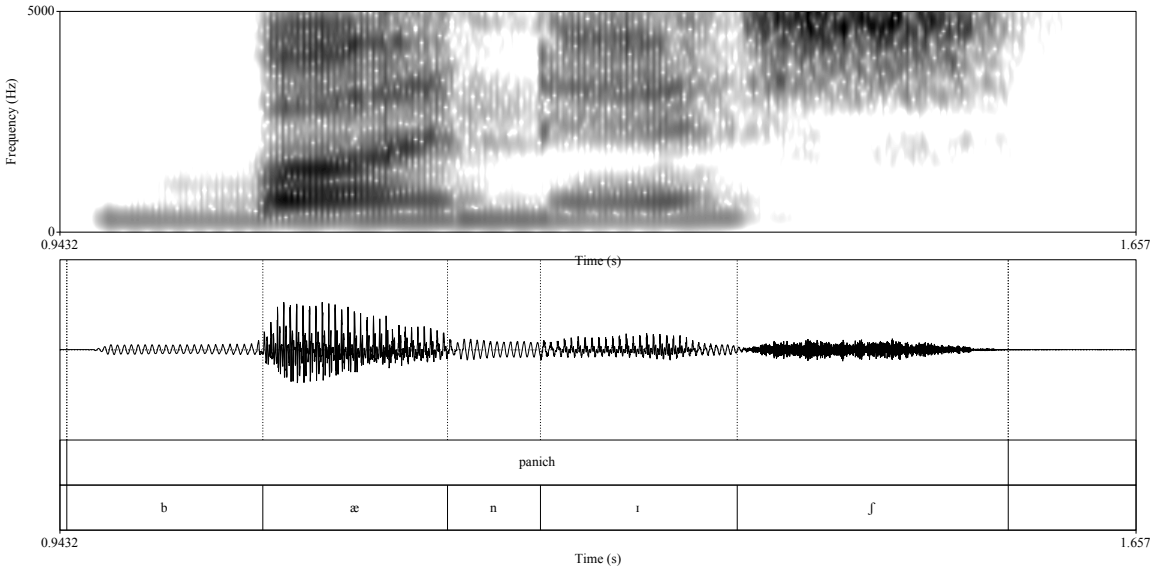


Figure 4.10: Sample waveform and spectrogram of the imported segment [ʃ] (FL06_OW_panich1)

It was difficult to acoustically annotate tokens as either [tʃ] or [ʃ] in word-initial position due to a design flaw in the frame sentence. The word ‘company’ [ʃæɹɪkət] that precedes the

target non-words ends with /t/. It would have been better to provide participants with a word that ends with a vowel or glide. To reiterate, the reason for using this word was to make participants produce non-words as if they were real English names for companies because some participants may have a good level of English. Two clear bursts were expected if the segment were /tʃ/: one burst for the final stop in [ʃæɪkət] and one burst for the stop portion of /tʃ/ in the next non-word. In fact, there was only one burst in many tokens. Acoustically, it was therefore difficult to identify whether the stop closure and burst belong to /tʃ/ or to the preceding word. In this study, therefore, auditory analysis had to be used in such cases. The friction portion of [tʃ] and [ʃ] was analysed impressionistically in word-initial position. However, acoustic analysis was possible in cases where tokens were produced as isolated words (i.e., when participants did not produce the target words in the carrier sentences) or inserted a vowel after the preceding word (i.e., when participants inserted Arabic case markers). The number of acoustically identified tokens for target [tʃ] is reported in Table 4.10. The classification process for [tʃ] and [ʃ] in all production data is summarized in Table 4.9:

Table 4.9: Classification criteria for [tʃ] and [ʃ]

Annotation	Criteria
Acoustically annotated as [tʃ]	• The segment had a period of silence followed by friction with a burst between the two (Hayward, 2013).
Acoustically annotated as [ʃ]	• The segment had friction without a period of silence or burst (Hayward, 2013).
Impressionistically annotated as either [tʃ] or [ʃ]	• The word-initial segment, produced in a word within a carrier sentence, had a period of silence followed by friction with a burst between the two. However, in word-initial position, an exception was made for segments which were preceded by another clear burst release for the final stop in the preceding word. That is, if there were two bursts, the segment was acoustically annotated as [tʃ].

Table 4.10 shows the acoustic and impressionistic annotations for [tʃ] and [ʃ] in the different word positions.

Table 4.10: Total number of annotations for [tʃ] and [ʃ] in target non-words

Annotation	Initial	Intervocalic	Final	% of Annotations
Acoustically annotated as [tʃ]	139	326	297	762 (65.02%)
Acoustically annotated as [ʃ]	33	50	61	144 (12.29 %)
Impressionistically annotated as [tʃ]	172	None	None	172 (14.68 %)
Impressionistically annotated as [ʃ]	25	None	None	25 (2.13%)
Others				69 (5.89%)

4.9 Acoustic Measurements of Imported and Adapted Sounds in Target Non-words

As explained in the previous section, classifying the realizations of /v/ and /tʃ/ into imported sounds [v] and [tʃ] and adapted sounds [f] and [ʃ] was based on acoustic representation and auditory judgment. The primary purpose of this section is to corroborate the results of the classification process. In this section, we examine whether the acoustic properties of the imported and adapted sounds match what we expect for those sounds. Imported [v] is expected to have a shorter duration, lower COG and greater average intensity than that of adapted [f] while imported [tʃ] is expected to have a shorter friction duration and rise time than that of adapted [ʃ]. The measurements reported here were taken from the sounds that were either identified acoustically or impressionistically.

4.9.1 Acoustic Measurements of [v] and [f]

Three measurements were taken for the distinction between [v] and [f]: friction duration, intensity, and COG (Ogden, 2009). Differences between [v] and [f] in each parameter are depicted in Figures 4.11, 4.12 and 4.13. [v] had a shorter duration, lower COG, and greater average intensity than that of [f]. Regarding intensity, it is worth noting that the findings should be interpreted with caution because the production tasks were conducted online in an uncontrolled setting and intensity may vary across speakers since they used different recording devices. The intensity of [v] was higher than [f] which is not in agreement with Hayward (2013)

and Ogden (2009) who describe English voiced fricatives as having a lower intensity than voiceless fricatives. However, the finding here is congruent with results reported by Jongman et al. (2000), in which voiced fricatives had higher intensity than their voiceless counterparts. The mean intensity in /f/ and /v/ were 55.7dB and 63.2 dB, respectively.

Hayward (2013) and Ogden (2009) describe generally English voiced fricatives in different syllables. However, Jongman et al. (2000) focused only CVC words. It is also perhaps relevant that the work of Hayward (2013) and Ogden (2009) was based on British English, but the work of Jongman et al. (2000) was based on American English. The target variety for the Saudi speakers in this study probably is the American English. The primary purpose of the acoustic analysis here, however, is to support the classification process. Further examination of whether the properties of [f] and [v] found in this thesis match those of English fricatives is left for future study.

Tables 4.11 and 4.12 illustrate the results of a comparison between [v] and [f] in target non-words across the three different positions. The mean and standard deviation for each parameter was calculated.

Table 4.11: Duration, COG, and intensity for [v] in target non-words across the three word positions

	[v]					
	Initial		Intervocalic		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	114.79	41.73	85.08	20.25	116.50	50.08
COG values in Hz.	2233.81	1336.169	1667.195	837.78	2570.636	1602.25
Intensity values in dB.	63.13	8.59	62.90	8.06	51.73	7.73

Table 4.12: Duration, COG, and intensity for [f] in target non-words across the three word positions

	[f]					
	Initial		Intervocalic		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	122.22	58.47	105.27	23.80	204.68	89.46
COG values in Hz.	2676.863	1560.997	2219.392	1446.12	2924.658	1818.097
Intensity values in dB.	48.63	9.41	54.79	9.26	48.46	9.04

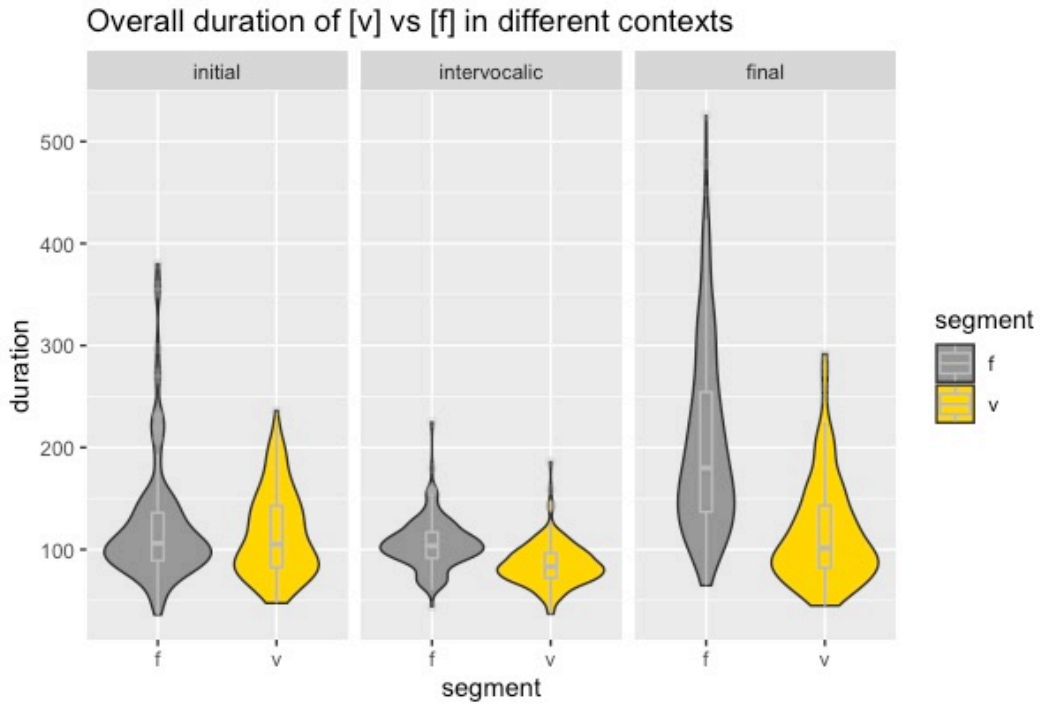


Figure 4.11: Friction duration for [v] and [f] in target non-words across the three word positions

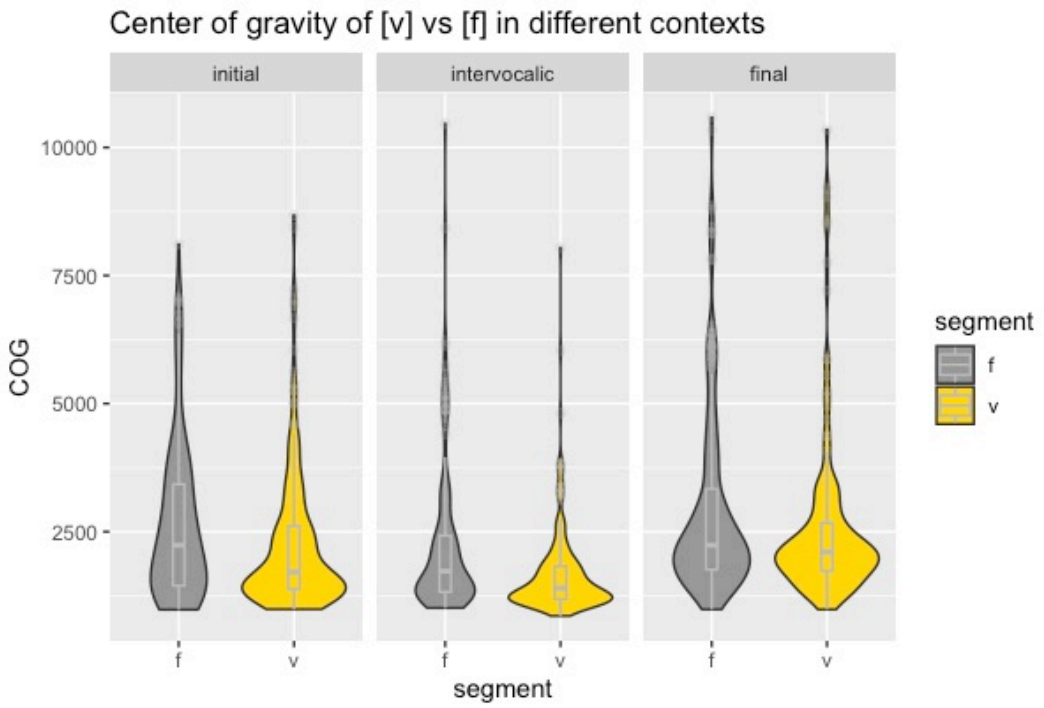


Figure 4.12: Center of gravity of [v] and [f] in target non-words across the three word positions

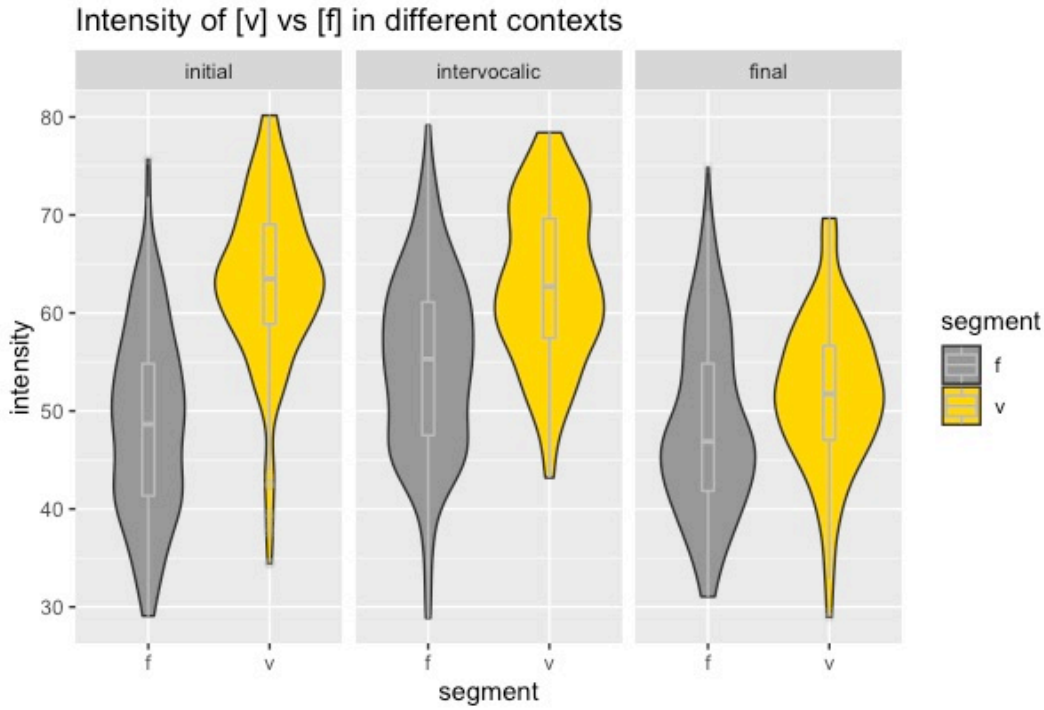


Figure 4.13: Mean intensity for [v] and [f] in target non-words across the three word positions

As an example, consider the non-word ‘vanit’ as produced by a male speaker in the high English exposure group. As shown in Figures 4.14 and 4.15, the speaker produced the initial sound in the same target non-word differently across repetitions. He produced it as [f] in the first token and as [v] in the second token. The average intensity levels of [f] (72.23 dB) and [v] (73.46 dB) were almost identical. However, other differences were still found. The friction duration was longer in [f] (189.38 ms) than in [v] (128.95ms). The centre of gravity was higher in [f] (5626.06 Hz) than in [v] (1593.08 Hz).

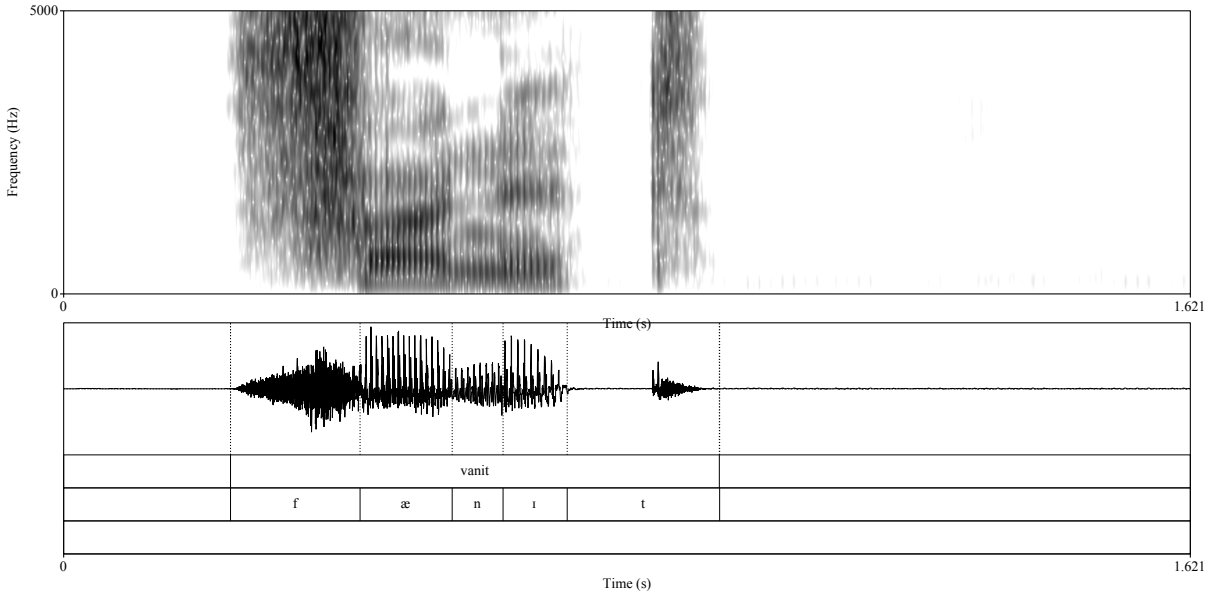


Figure 4.14: Sample waveform and spectrogram of the adapted segment [f] (MH07_O_vanit1)

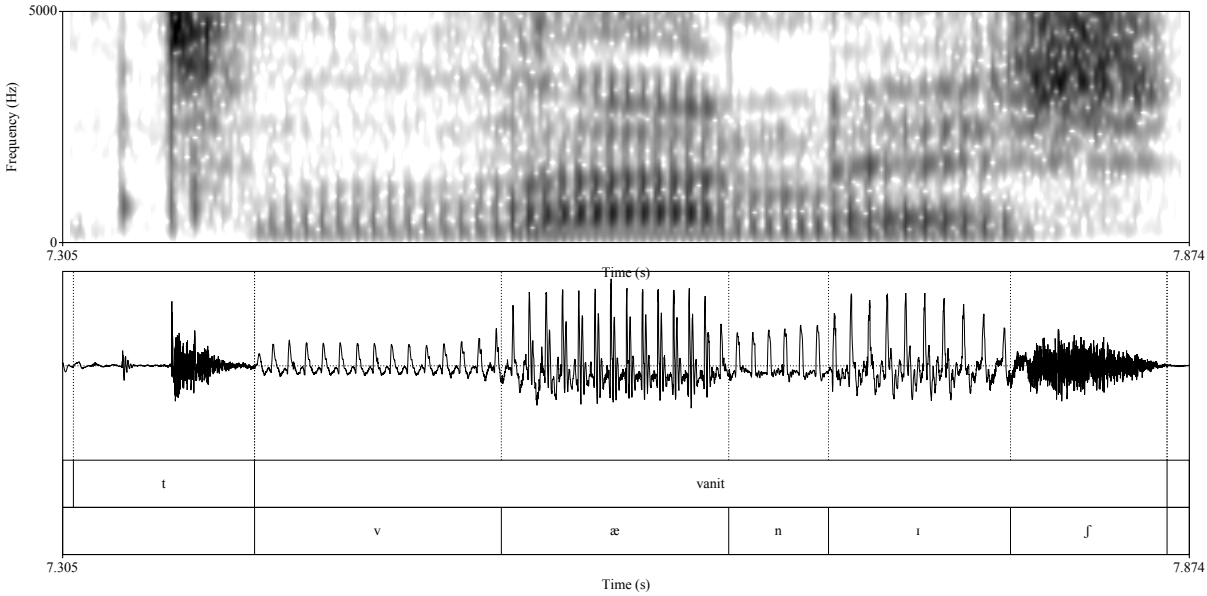


Figure 4.15: Sample waveform and spectrogram of the imported sound [v] (MH07_O_vanit2)

In a final step, linear mixed-effects models were constructed to examine whether [v] and [f] show significant differences in duration, intensity, and COG in each word position. [v] and [f] were included as fixed effects. As random effects, there were random intercepts for 'speaker', and 'word'. Differences between [v] and [f] were found to be significant in each parameter. Tables 4.13, 4.14 and 4.15 show general models constructed to examine the differences between the two segments in each parameter.

Table 4.13: Linear mixed effects model for the distinction between [v] and [f] in duration in target non-words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	127.000	9.161	15.385	13.86	4.2e-10 ***
Segment1	18.446	1.703	1120.373	10.84	< 2e-16 ***

Table 4.14: Linear mixed effects model for the distinction between [v] and [f] in COG in target non-words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	2359.91	148.40	27.35	15.903	2.40e-15 ***
Segment1	174.46	43.40	1126.65	4.019	6.22e-05 ***

Table 4.15: Linear mixed effects model for the distinction between [v] and [f] in intensity in target non-words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	55.4222	1.3972	25.9891	39.67	<2e-16 ***
Segment1	-3.6834	0.2114	1086.8622	-17.42	<2e-16 ***

4.9.2 Acoustic Measurements of [tʃ] and [ʃ]

Two measurements were taken for the distinction between [tʃ] and [ʃ]: friction duration and the amplitude rise time (Hayward, 2013). Friction duration and the amplitude rise time for [tʃ] and [ʃ] in the three word positions are reported in Tables 4.16 and 4.17 and visualized in Figures 4.16 and 4.17. [ʃ] had a longer friction duration than that of [tʃ]. Additionally, the friction of [ʃ] had a longer rise time (i.e., a gradual friction noise) than that of [tʃ]. These results suggest that the two sounds are like English in terms of friction duration and rise time

Table 4.16: The friction duration and amplitude rise time for [tʃ] in target non-words across the three word positions

	[tʃ]					
	Initial		Intervocalic		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	72.44	23.63	91.12	23.83	140.3	66.78
Mean rise time values in ms.	40.93	22.37	39.71	23.91	45.02	32.49

Table 4.17: The friction duration and amplitude rise time for [ʃ] in target non-words across the three word positions

	[ʃ]					
	Initial		Intervocalic		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	208.66	62.21	142.66	30.92	220.95	97.03
Mean rise time values in ms.	158.43	63.96	60.87	36.50	66.16	48.66

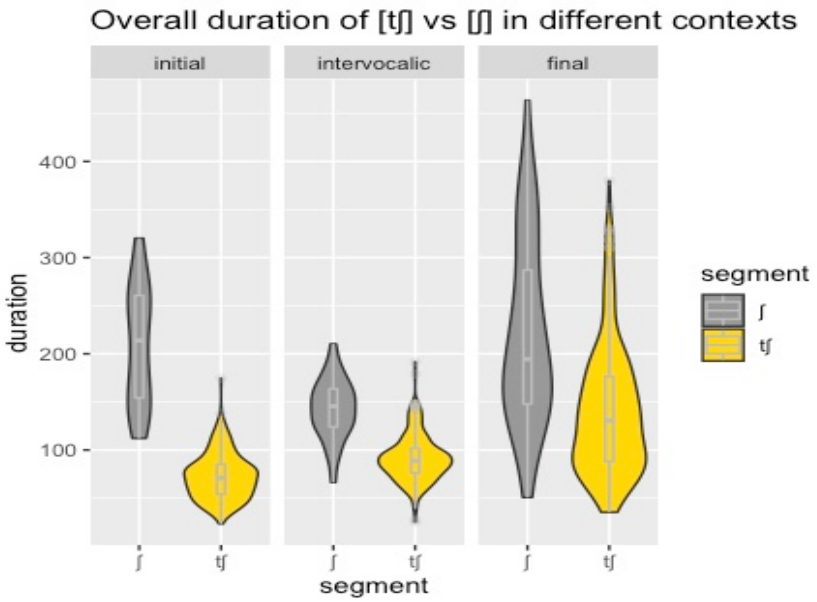


Figure 4.16: Durations of [tʃ] and [ʃ] in target non-words across the three word positions

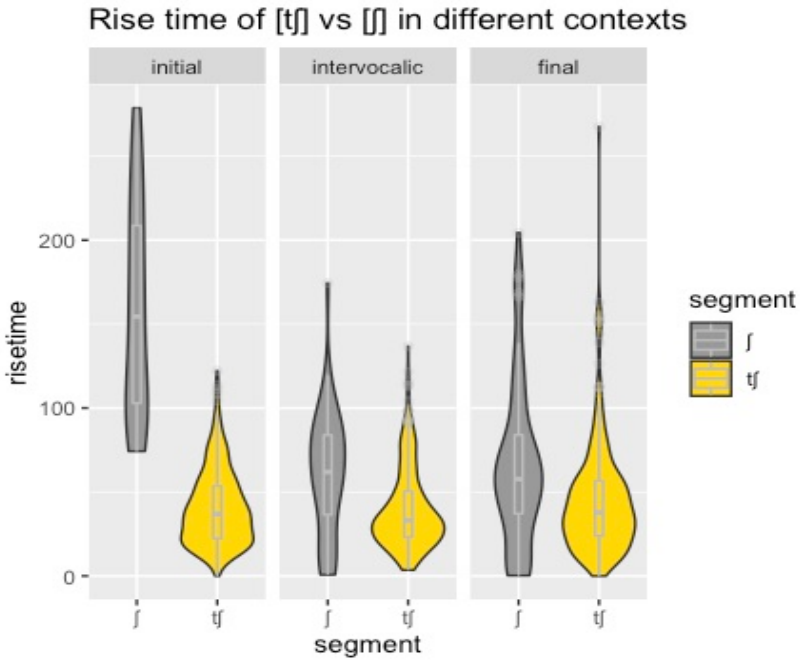


Figure 4.17: Amplitude rise time of [tʃ] and [ʃ] in target non-words across the three word positions

As an example, Figures 4.18 and 4.19 show the non-word ‘rachim’ as produced by a male speaker in the high English exposure group in two repetitions. The intervocalic sound was produced as [ʃ] in the first token and as [tʃ] in the second token. As expected, the friction duration was longer in [ʃ] (192.54 ms) than in [tʃ] (104.71 ms). The rise time was also longer in [ʃ] (52.19 ms) than in [tʃ] (23.38 ms).

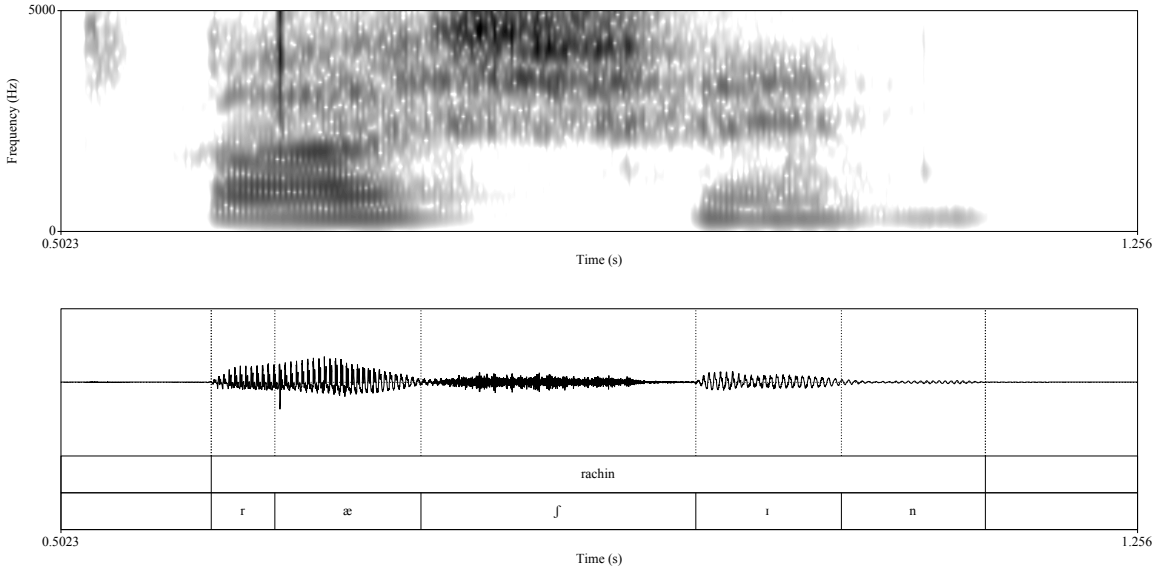


Figure 4.18: Sample waveform and spectrogram of the adapted sound [ʃ] (FM12_W_rachin1)

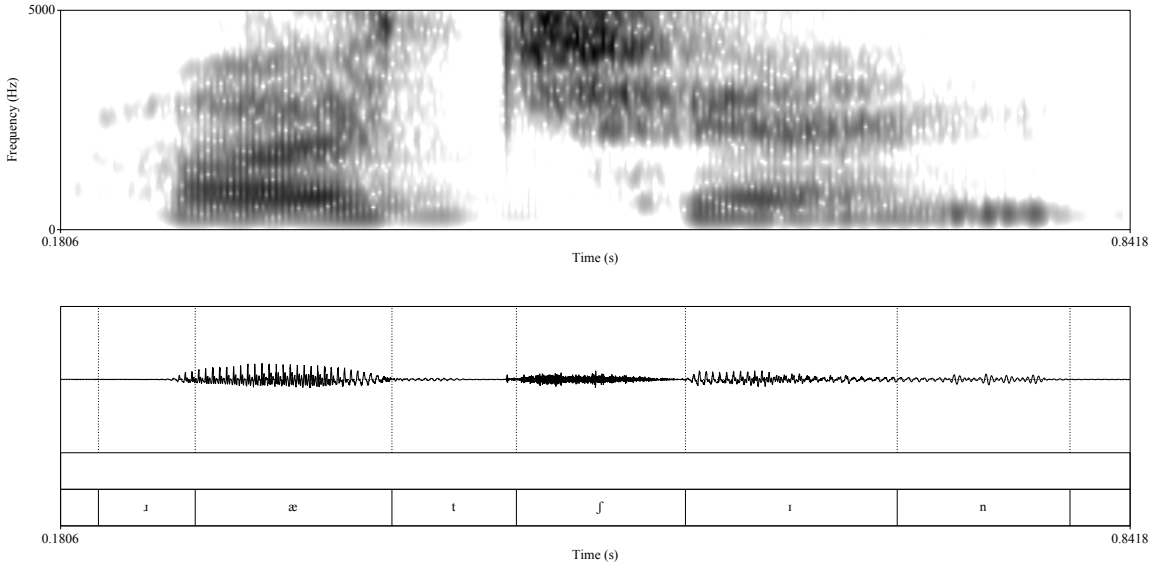


Figure 4.19: Sample waveform and spectrogram of the imported sound [tʃ] (FM12_W_rachin2)

Finally, linear mixed-effects models were constructed to examine whether [tʃ] and [ʃ] show significant differences in duration and rise time. The segments were included as fixed effects. As random effects, there were random intercepts for 'speaker', and 'word'. Differences between [tʃ] and [ʃ] in each parameter were found to be significant. Tables 4.18 and 4.19 show the findings of models constructed to examine differences between the two sounds in each parameter.

Table 4.18: Linear mixed effects model for the distinction between [tʃ] and [ʃ] in duration in target non-words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	134.610	9.441	16.924	14.26	7.36e-11 ***
Segment1	32.757	1.938	1045.904	16.90	< 2e-16 ***

Table 4.19: Linear mixed effects model for the distinction between [tʃ] and [ʃ] in amplitude rise time in target non-words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	60.814	2.818	46.141	21.58	<2e-16 ***
Segment1	18.034	1.491	1072.116	12.10	<2e-16 ***

4.10 Main Results

In this section, an analysis of /v/ and /tʃ/ realizations (i.e., imported sounds [v] and [tʃ] vs. adapted sounds [f] and [ʃ]) in target non-words is presented with both descriptive statistics and inferential statistics using mixed-effects logistic regression.

4.10.1 Descriptive Analysis

Figures 4.20 and 4.21 display the proportion of /v/ and /tʃ/ realizations in target non-words. Tables 4.20-4.23 show the raw results split by key independent variables.

The overall rate of importation for target /v/ was lower in aural-only condition (52.02%) and aural-written condition (53.89%) than in written-only condition (59.25%). These descriptive results may suggest that target /v/ was less likely to be produced as the imported sound [v] with the availability of auditory inputs. The importation rate differed across the three exposure

groups. Speakers with the highest level of exposure to English showed the highest overall level of importation (69.29%). Correspondingly, speakers with the lowest level of English exposure showed the lowest level of importation (38.4%). Speakers with the medium level of English exposure showed a somewhat lower level of importation than the high group but a higher level of importation than the low group (54.57%). As shown in Figure 4.20, the importation rate of /v/ was generally lower in word-final position compared to the other positions. The proportion of [v] realizations was lowest in word-final position (44.44%), followed by word-intervocalic position (56.62%) and then word-initial position (63.83%). Finally, in terms of gender, males showed a lower degree of importation than females (46.24% and 62.52% respectively).

The importation rate of target /tʃ/ was high, though as predicted, speakers in the high exposure group showed a higher proportion of importation (93.83%) than speakers in the medium and low exposure groups (82.96% and 75.59 % respectively). Interestingly, in contrast to the findings for /v/, speakers in all three groups showed a lower importation rate with the availability of written inputs. The proportion of [tʃ] realizations was higher in aural-only condition (96.13%) than in the other two conditions. Regarding word position, there seems to be no obvious pattern. As shown in Figure 4.21, there were small differences between the three positions, suggesting there was little if any effect of word position. The proportion of [tʃ] realizations was highest in word-intervocalic position (86.7%), followed by word-initial position (84.28%), and then word-final-position (82.96%). Additionally, small differences can be seen in the figure between males and females in the three exposure groups. However, Table 4.23 shows that males showed a lower degree of importation than females (82.78% and 86.32%, respectively).

Table 4.20: Summary of participants' production of /v/ and /tʃ/ in target non-words by condition

Non-words								
	[v]		[f]		[tʃ]		[ʃ]	
	Count	%	Count	%	count	%	Count	%
Aural condition	193	52.02%	178	47.98%	348	96.13%	14	3.87%
Aural-written condition	208	53.89%	178	46.11%	346	89.64%	40	10.36%
Written condition	221	59.25%	152	40.75%	240	67.61%	115	32.39%

Table 4.21: Summary of participants' production of /v/ and /tʃ/ in target non-words by level of English exposure

	Non-words							
	[v]		[f]		[tʃ]		[ʃ]	
	Count	%	Count	%	Count	%	count	%
High group	291	69.29%	129	30.71%	380	93.83%	25	6.17%
Medium group	197	54.57%	164	45.43%	297	82.96%	61	17.04%
Low group	134	38.4%	215	61.6%	257	75.59%	83	24.41%

Table 4.22: Summary of participants' production of /v/ and /tʃ/ in target non-words by word position

	Non-words							
	[v]		[f]		[tʃ]		[ʃ]	
	Count	%	count	%	count	%	count	%
Initial position	240	63.83%	136	36.17%	311	84.28%	58	15.72%
Intervocalic position	218	56.62%	167	43.38%	326	86.7%	50	13.3%
Final position	164	44.44%	205	55.56%	297	82.96	61	17.04%

Table 4.23: Summary of participants' production of /v/ and /tʃ/ in target non-words by gender

	Non-words							
	[v]		[f]		[tʃ]		[ʃ]	
	count	%	count	%	Count	%	count	%
Females	382	62.52%	229	37.48%	511	86.32%	81	13.68%
Males	240	46.24%	279	53.76%	423	82.78%	88	17.22%

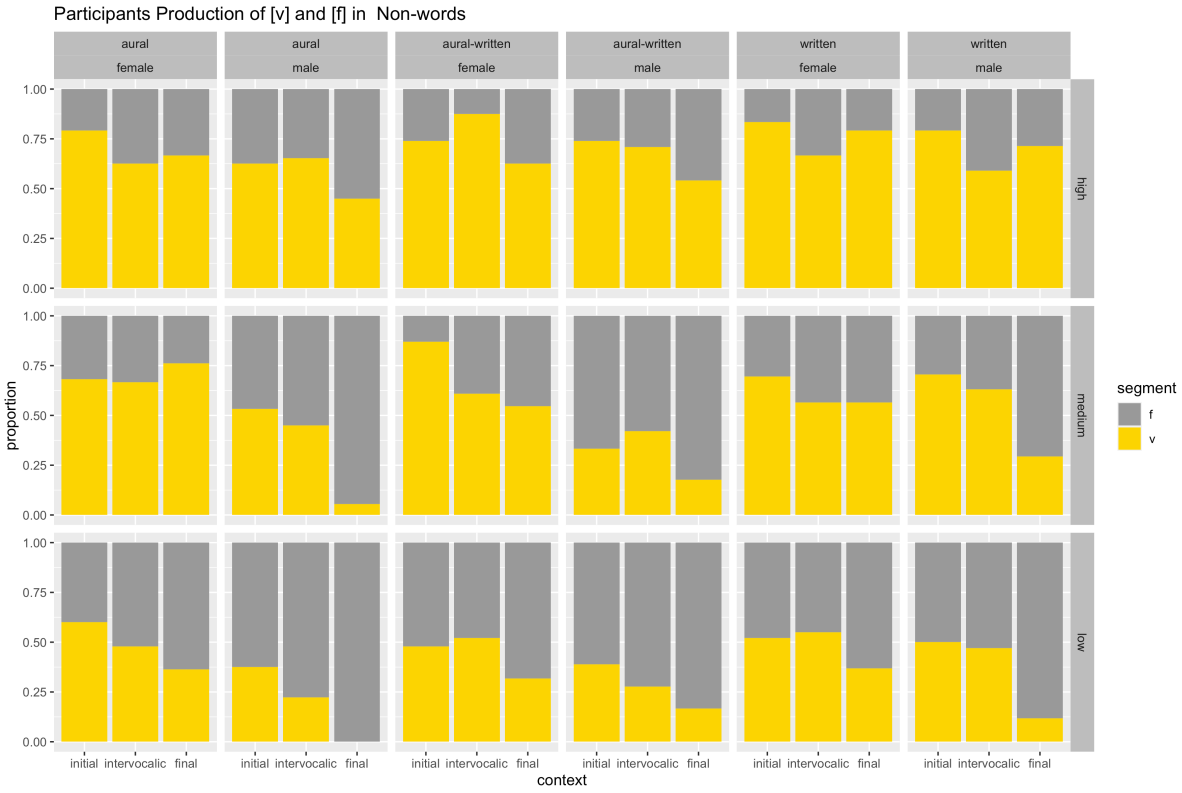


Figure 4.20: Proportion of /v/ realizations in target non-words by condition, word position, language exposure group and gender

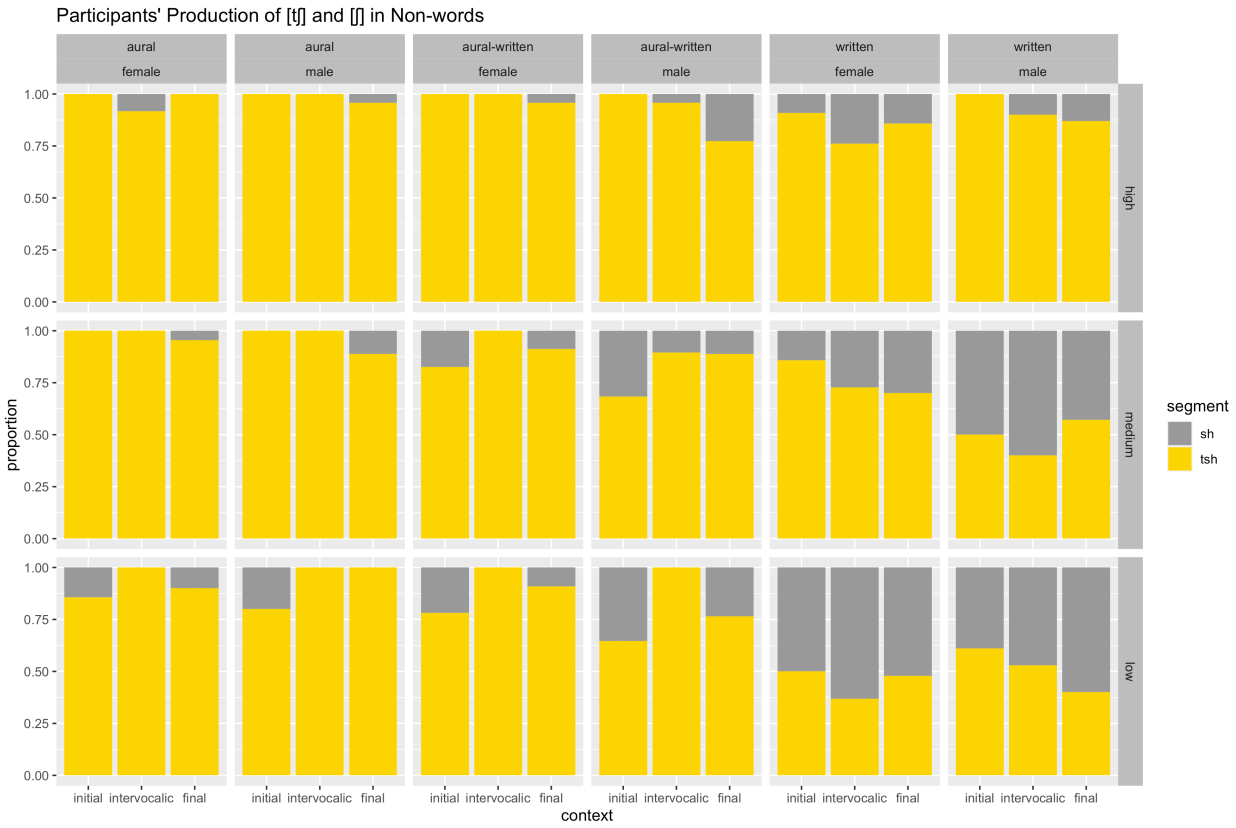


Figure 4.21: Proportion of /tʃ/ realizations in target non-words by condition, word position, language exposure group and gender

4.10.2 Logistic Regression Analysis

Turning to the inferential statistics, it is worth reiterating that the target sounds were dummy coded (the imported sounds [v] and [tʃ] = 0 and the adapted sounds [f] and [ʃ] = 1). Slopes were not included if the model did not converge. Additionally, interactions between fixed factors were not included if they were not significant and did not improve the model fit. Table 4.24 and Figure 4.22 present the results of the mixed effects model constructed for target /v/. The model[10] includes word position (initial, intervocalic and final), exposure group (high, medium and low), and a two-way interaction between condition (aural-only, aural-written and written-only) and gender (males and females) as fixed effects. There is a random intercept for word and a slope for word position by speaker. It is worth noting that the random slope for condition by

[10] glmer(segment ~ context + group + gender*condition + (1 + context| speaker)+ (1 | word),data = nonwords_v, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

speaker was excluded because the model failed to converge. Interactions between other fixed factors were not included because they were not significant and did not improve the model fit.

The estimate for intercept is the estimate for aural-only condition, word-final position, low exposure group (i.e., speakers with the lowest level of English exposure), and males. The intercept is positive and significant. This indicates that the adapted sound [f] was more likely to be produced in word-final position and in aural-only condition by speakers in the low exposure group and males. The model confirms the main effect of word position: the imported sound [v] was more likely to be produced in both word-initial position (β - 1.5524, SE =0.4960, z = -3.130, p = 0.00**) and word-intervocalic position (β -0.8802, SE = 0.4316, z = -2.040, p = 0.04*) than in word-final position. The model also reveals a main effect of level of English exposure. The coefficients associated with the high exposure group (β = -2.1889, SE = 0.4171, z = -5.247, p = 1.54e-07 ***) and the medium exposure group (β = -1.0199, SE = 0.4059, z = -2.513, p = 0.011*) are each significantly different from those associated with the reference level (low exposure group). In terms of the effect of condition, the difference between aural-only and aural-written conditions is not significant (β = -0.3332, SE = 0.3048, z = -1.093, p = 0.27439), indicating that the importation rate was similar in these two conditions. However, the difference between aural-only and written-only conditions is significant (β -1.0458, SE = 0.3225, z = -3.243, p = 0.00118 **), meaning that the imported sound [v] was more likely to be produced when written input was visible. The effect of gender is significant (-1.9756 (SE = 0.4181, z = -4.725, p = 2.30e-06 ***), meaning that females were more likely to produce the imported sound [v] than males. There is a significant interaction between gender and condition, indicating that females were more likely than males to produce the imported sound [v] in the aural-only condition (β = 1.2801, SE = 0.4087, z = 3.132, p = 0.00174 **).

Table 4.24: Summary of mixed effects logistic regression model for [v] and [f] in target non-words

Fixed Effects	Estimate	SE	z value	Pr(>|z|)
Intercept	2.7468	0.5065	5.424	5.84e-08 ***
Word-initial position	-1.5524	0.4960	-3.130	0.00175 **
Word-intervocalic position	-0.8802	0.4316	-2.040	0.04139 *
High group	-2.1889	0.4171	-5.247	1.54e-07 ***
Medium Group	-1.0199	0.4059	-2.513	0.01198 *
Aural-written condition	-0.3332	0.3048	-1.093	0.27439
Written-only condition	-1.0458	0.3225	-3.243	0.00118 **
Females	-1.9756	0.4181	-4.725	2.30e-06 ***
Gender female: Aural-written condition	0.5896	0.4010	1.470	0.14148
Gender female: Written condition	1.2801	0.4087	3.132	0.00174 **

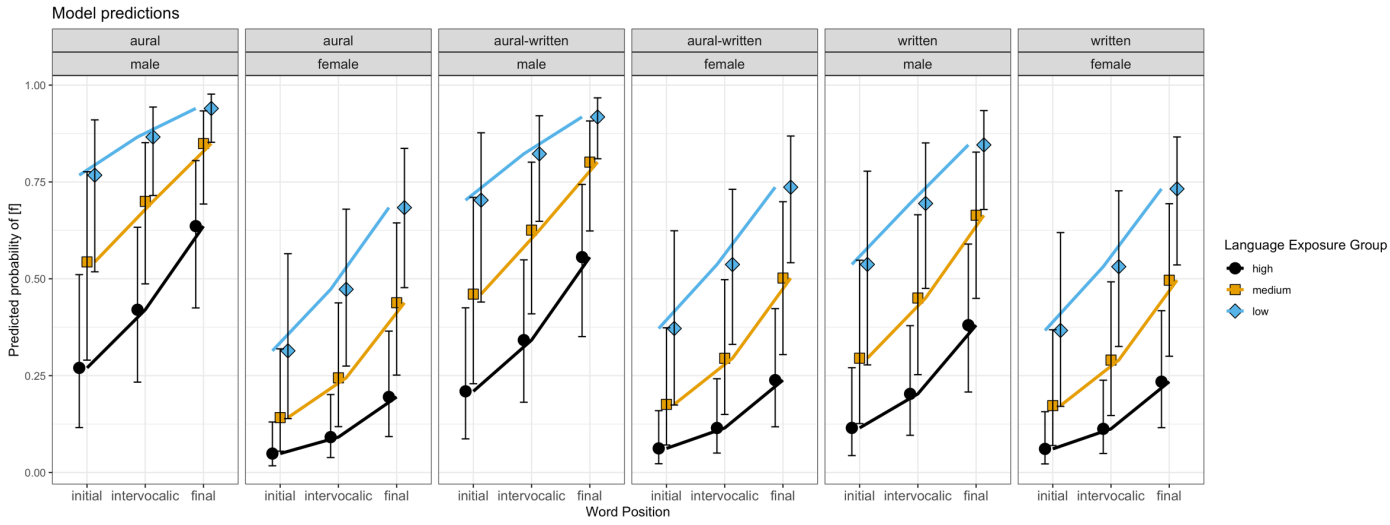


Figure 4.22: Predicted probability of [f] production in target non-words by condition, language exposure group, word position and gender

The model for target /tʃ/ [11] includes word position (initial, intervocalic and final), exposure group (high, medium and low), gender, (males and females) and condition (aural-only, aural-written, and written-only) as fixed effects. As random effects, there are intercepts for word and speaker. It is important to note that this simple model was adopted because the model with the random slopes (condition by participant and condition by item) failed to converge. The interaction between exposure group and condition was not included because it was not significant and did not improve the model fit.

The estimate for the intercept is the estimate for aural-only condition, low group, males, and word-final position. The model shows that word position was not a significant factor. However, the model reveals the predicted main effect of condition. As shown in Figure 4.23, the imported sound [tʃ] was less likely to be imported in aural-written condition (β =1.1494, SE = 0.3543, z = 3.245, p = 0.00**) and in written-only condition (β = 3.1923, SE = 0.3534, z = 9.033, p = < 2e-16 ***) than in aural-only condition. Concerning the effect of English exposure, the model shows that there is no significant difference between the low and medium exposure groups. However, the high exposure group differs significantly from the low exposure group (β = -2.4332, SE = 0.6027, z = -4.037, p = 5.4e-05 **), indicating that the imported sound [tʃ] was more likely to be produced by speakers in the high exposure group. Finally, the model reveals no significant gender differences (β = -0.5435, SE = 0.4652, z = -1.169, p = 0.24), meaning that females and males had similar rates of importation.

[11] glmer(segment ~ context + group + condition + gender + (1 | speaker) + (1 | word), data = nonwords_ch, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

Table 4.25: Summary of mixed effects logistic regression model for [tʃ] and [ʃ] in target non-words

Fixed Effects	Estimate	SE	z value	Pr(>|z|)
Intercept	-2.7528	0.5682	-4.845	1.27e-06 ***
Word-initial position	-0.1318	0.2666	-0.494	0.62120
Word-intervocalic position	-0.4800	0.2850	-1.684	0.09215
High group	-2.4332	0.6027	-4.037	5.41e-05 ***
Medium Group	-0.7766	0.5503	-1.411	0.15822
Aural-written condition	1.1494	0.3543	3.245	0.00118 **
Written condition	3.1923	0.3534	9.033	< 2e-16 ***
Females	-0.5435	0.4652	-1.169	0.24260

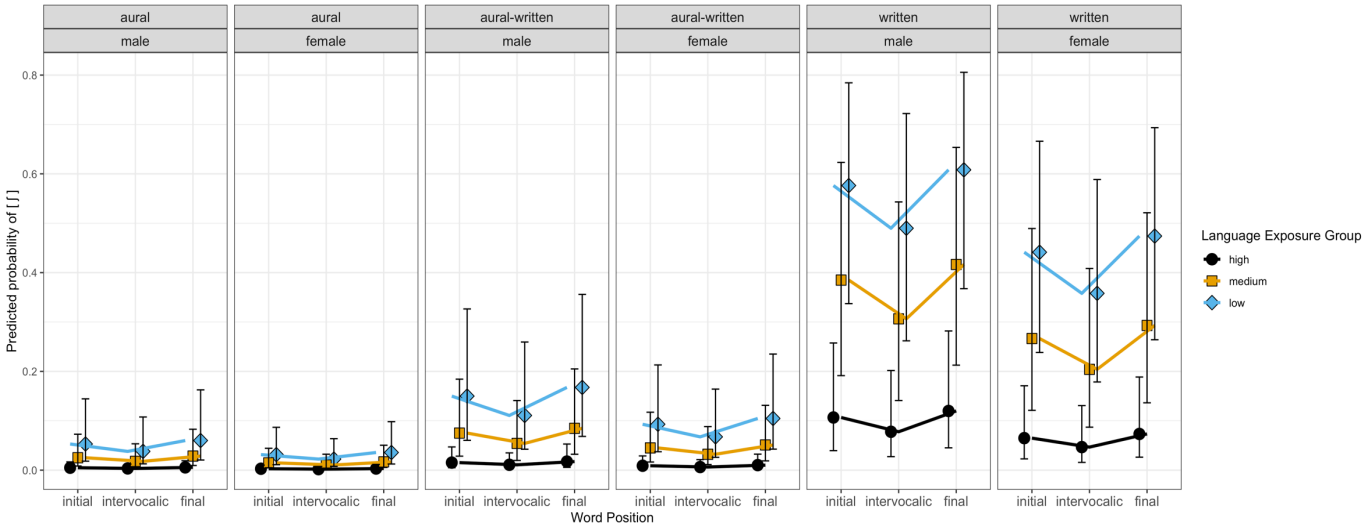


Figure 4.23: Predicted probability of [ʃ] production in target non-words by condition, language exposure group, word position and gender

The discussion of these results will be taken up in detail in Section 4.11, in what follows is a summary of all factors affecting the likelihood of producing the adapted sounds [f] and [ʃ] in target non-words.

4.10.3 Interim Summary

This subsection has presented the results of /v/ and /tʃ/ production in target non-words. The results were obtained through descriptive statistics and inferential statistics using mixed-effects logistic regression. Table 4.26 shows only the statistically significant results obtained from the mixed effects logistic regression models for the two target sounds.

The significant factors affecting the importation rate of /v/ in target non-words were the following: word position, level of English exposure, condition, and gender. Target /v/ was more likely to be imported as [v] in word-initial and intervocalic positions than in word-final position. Speakers with high and medium levels of English exposure were more likely to produce the imported sound [v] than speakers with low levels of English exposure. Speakers' production was affected by condition: the imported sound [v] was more likely to be produced in written-only condition than in the other two conditions. Finally, females were more likely to produce the imported sound [v] than males, especially when auditory input is available.

The significant factors affecting the importation rate of /tʃ/ were condition and level of English exposure. Overall, the imported sound [tʃ] was less likely to be produced in aural-written and written-only conditions than in aural-only condition. Speakers in the high exposure group were more likely to produce the imported sound [tʃ] than speakers in the other exposure groups.

Table 4.26: Factors influencing the likelihood of producing [f] and [ʃ] for target /v/ and /tʃ/ in non-words

	Group reference = low		Word position reference = final		Condition reference = aural-only		Gender reference = female	Condition *Gender
	high	medium	initial	intervocalic	aural-written	written	female	written*female
[v] reference = [f]	***	*	**	*		**	***	**
[tʃ] reference = [ʃ]	***				**	***		

Asterisks refer to levels of significance: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001
Reference refers to the reference level (intercept)

4.11 Discussion

This section aims to discuss several possible explanations for the key findings in relation to the predictions and previous literature. In what follows are attempts to offer alternative explanations for the different effects of the factors on the production of the two target sounds. The possible explanations for these findings are not mutually exclusive. That is, a combination of explanations could account for the same finding.

4.11.1 Effect of Input Type

The presence of auditory input was predicted to result in an increase in the importation rate for /tʃ/ but not for /v/ (Prediction a and Prediction b). As predicted, the findings showed that the effect of input modality varies for the different target sounds. Unlike /v/, /tʃ/ was less likely to be imported with the availability of written inputs. As shown in figures 4.22 and 4.23, the effect of the written input type was greater for /tʃ/ than /v/. The differences between the three conditions for target /v/ were small although they were significant. Interestingly, exposure to two input types (auditory + orthographic) did not result in an increase in the importation rate of either target sound. This finding for target /tʃ/ disagrees with previous studies demonstrating that orthographic input along with auditory input can improve the production accuracy of non-native sounds (e.g., Davidson, 2010), but is in line with other studies showing that the orthographic input accompanying the auditory input is not helpful (e.g., Bassetti and Atkinson, 2015, Vendelin and Peperkamp, 2006).

Overall, these patterns may relate to the results of the perception task reported in Chapter 3. Listeners' perception was less accurate for the /v-f/ contrast than the /tʃ-ʃ/ contrast. The discrimination accuracy of the /v-f/ contrast was overall weak; thus, it is unsurprising that speakers' importation rate of target /v/ was low with the availability of auditory inputs. The interaction between condition and gender was significant for written-only condition. As shown in Figure 4.22, females were more likely than males to produce [v] in aural-only and aural-written conditions. The differences between males and females disappeared in written-only condition, indicating that males were misled by auditory input in aural-only and aural-written conditions.

The discrimination accuracy of the /tʃ-ʃ/ contrast was good; correspondingly, auditory input was more helpful than written input, especially for speakers who are less familiar with English orthography. Note that /tʃ/ was spelled as <ch> and /v/ was spelled as <v> in the stimuli. The grapheme <ch> seems to be unhelpful for Arabic learners of English. In English, /tʃ/, /ʃ/ and

even /k/ can be represented by the same grapheme (e.g., chair, chat, machine and chef, brochure, school, and character). Arabic orthography is transparent, one letter represents only one sound. This spelling-sound inconsistency may explain the low importation rate of /tʃ/ in written-only condition. Arabic speakers are used to a script that is quite transparent, with consistent sound-to grapheme correspondence. The difference in the transparency between the two writing systems can make the grapheme <ch> confusing for Arabic speakers because it represents different sounds in English.

4.11.2 Effect of Language Exposure

Level of exposure to English was predicted to influence the importation rate of both /v/ and /tʃ/. As predicted, speakers in the high exposure group were more likely to produce the imported sounds [v] and [tʃ] than speakers in the medium and low exposure groups (Prediction c). The influence of language exposure is most evident in how the input type contributes to the production of /v/ and /tʃ/. The written-only condition affected the importation rate of /v/; this effect was greater with more exposure to English. Speakers in the high exposure group were more likely to utilize orthographic information to produce the imported sound [v]. The presence of written input had little, if any, impact on the production of /v/ for speakers in the low exposure group.

In contrast, the absence of auditory input significantly decreased the importation rate of /tʃ/ in all three exposure groups. Speakers in the three exposure groups were less likely to produce [tʃ] in written-only condition than in aural-only condition. However, speakers in the high exposure group were found to have the highest importation rate in written-only condition as they are probably more familiar with English orthography because they are students at the department of English. As shown in Figure 4.23, there was a large degree of variability within the high and medium exposure groups, suggesting that some of the speakers struggled less with written inputs than others.

These results are in line with previous literature on loanword phonology demonstrating the effect of exposure to the source language on the production of loanwords (e.g, Kang, 2021; Poplack, 2018). According to Kang (2021), variation in the production of novel structures in loanwords is constrained by speakers' knowledge of the source language. The importation rate of /v/ and /tʃ/ was the highest for speakers with more exposure to English because they have a greater opportunity to hear and read the target sounds. However, regardless of speakers' level of

exposure to English, the importation rate of /tʃ/ was high especially with the availability of auditory input. The possible reason for this finding is that /tʃ/ exists in other Arabic dialects.

4.11.3 Effect of word position

The importation rate of /v/ and /tʃ/ was predicted to be affected in word-final position (Prediction d). The results showed that word position was a major factor influencing the importation rate of /v/ but not /tʃ/. The adapted sound [f] was more likely to be produced in word-final position than in other positions. It is worth noting that final /v/ was partially voiced in auditory input, which is a feature of English word-final fricatives. Thus, /v/ was even more similar to participants' closest native sound /f/ in word-final position. Final /v/ and /f/ do not differ much in voicing and they also have the same place of articulation (formed with upper teeth and lower lip) and manner of articulation (fricatives). In this case, the salient cues for perceiving the two sounds were solely duration and intensity, /v/ in auditory input was produced with shorter friction and lower intensity. Previous literature demonstrates that L2 sounds are more difficult to perceive and produce if they are phonetically similar to L1 sounds (e.g., Flege, 1987).

The source of this devoicing might be not their L1 which includes voicing contrast (e.g., [dəz] 'push' and [dəs] 'hide'). It is possible that markedness rather than the L1 had a greater influence on participants' production of final /v/. Markedness is used here to refer to cross-linguistic distribution; for example, voiced fricatives, cross-linguistically, are disfavoured in word-final position than voiceless fricatives (Eckman & Iverson, 1994). This distribution probably has an articulatory explanation which is the aerodynamic difficulty of maintaining friction and voicing simultaneously. For voicing to occur, the pressure below the glottis should be higher than pressure above the glottis. To generate friction, sufficient pressure is required across the locus of constriction (Ohala, 1983). In initial and intervocalic positions, voicing is likely to be maintained or retained because of following vowels or sonorant consonants.

This effect of word position can be also captured by the Interlanguage Structural Conformity Hypothesis (ISCH) (Eckman,1991) which states that L2 structure is governed by cross-linguistic tendencies rather than L1 transfer. Eckman's hypothesis was built on a study with similar results to those presented here. The study showed that Farsi speakers devoiced final English obstruents, even though Farsi has the same voicing contrast in word-final position (Eckman, 1984).

4.11.4 Effect of Gender

Gender was predicted to influence the likelihood of producing the imported sounds (Prediction e) because previous sociolinguistic studies on Arabic showed that women tend use novel forms more than men (e.g, Almuhanadi, 1991; Assiri, 2014; Omari & Van Herk, 2016.) Gender had a main effect on target /v/; females were only more likely to produce [v] than males when auditory input was available. The written-only condition made no significant gender difference in the importation of target /v/. Recall that the findings of the perception task showed that females and males were similar in their perception accuracy for the /v-f/ contrast.

A possible explanation is that females had more positive attitudes towards the English language and American culture than males. Previous studies showed that attitudes can strongly influence the importation of novel sounds in loanwords (e.g., Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2011). Attitudes towards the English language and American culture will be examined in Chapter 6 using survey data from the same participants in this production task.

4.12 Summary

In this section, we briefly summarise the aims, procedure, and key findings of the non-word production task.

The primary aim of this task was to examine the extent to which the production of the target sounds /v/ and /tʃ/ varies as a result of the following factors: the mode of input, level of English exposure, word position, and gender. The production of the two target sounds was tested in three conditions: aural-only (auditory inputs), written-only (orthographic inputs), and aural-written (auditory-orthographic inputs).

The findings of this task showed that input modality affected the likelihood of adaptation vs. importation. Speakers' production of target /v/ and /tʃ/ was different when their production cued by written inputs compared when cued by the auditory inputs. The presence of auditory input facilitated the production of [tʃ] but hindered the production of [v]. In contrast, the presence of written input facilitated the production of [v] but hindered the production of [tʃ]. Auditory information for target /v/ was less helpful for Arabic speakers because, as shown in the perception task, they had difficulty in discriminating /v/ from native /f/. Target /tʃ/ might be more susceptible to the adaptation than target /v/ in written-only condition due to inconsistency in mapping between /tʃ/ and <ch>.

The effect of language exposure was robust in our results; speakers in the high exposure group, those with more familiarity with English, tended to produce the imported sounds [v] and [tʃ] more than speakers in the medium and low exposure groups. This tendency might be explained be the fact that speakers in the high exposure group were likely to have read and heard more English words than speakers in the other two exposure groups. Word position significantly influenced the importation rate of /v/ but not /tʃ/: the lowest importation rate for /v/ was found in word-final position. A possible explanation for this result might be that voiced sounds, cross-linguistically, are more marked than their voiceless counterparts in word-final position; hence, voiced sounds are more likely to be devoiced in this position. Finally, gender had a main effect on the importation rate of /v/; however, females' importation rate was higher than males only when auditory input was available. It is difficult to explain this result, but we speculated that it might be due to differences in attitudes between females and males towards the English language and culture.

These findings reported throughout this chapter suggests that the variable adaptation of target sounds can be affected by different factors. It is essential to consider the effects of different factors (i.e., input modality, level of language exposure, word position, and gender) when discussing the process of loanword adaptation. In this chapter, we discussed factors affecting the production of target /v/ and /tʃ/ in non-words. However, it is still not clear whether these factors have the same effect on production of target /v/ and /tʃ/ in real words. In the next chapter, we will extend our understanding of how level of English exposure, word position, and gender may influence the production of the two sounds in real words.

5 Production of /v/ and /tʃ/ in Target Real Words

5.1 Introduction

Having considered the production of target /v/ and /tʃ/ in non-words, we turn to the production of the two sounds in real words. The findings of the previous production task showed that language exposure was a significant factor for target /v/: speakers in the high and medium exposure groups were more likely to produce the imported sound [v] than speakers in the low exposure groups. There was also a main effect of word position on the importation rate of target /v/: the imported sound [v] was more likely to be produced in both word-initial position and word-intervocalic position than in word-final position. Females showed a higher importation rate than males for target /v/. There was a significant interaction of gender with condition, whereby males were less likely to produce the imported sound [v] whenever auditory input is available. The performance of females and males did not differ in written-only condition. Regarding target /tʃ/ in non-words, input modality had a main effect: the imported sound [tʃ] was less likely to be produced in aural-written and written conditions than in aural-only condition. The effect of language exposure was also significant for target /tʃ/: speakers in the high exposure group were more likely to produce the imported sound [tʃ] than speakers in the other two exposure groups. The importation rate of target /tʃ/ was not affected by word position or gender.

In this chapter, we examine whether these effects of word position, gender, and level of English exposure reported in Chapter 4 extend to the production of /v/ and /tʃ/ in target real words. The production of /v/ and /tʃ/ in real words[12] was elicited using a fill-in-the-blank task in which participants were asked to read sentences and then to fill in the blanks with target real words represented in pictures. The question addressed in this chapter is as follows: *To what extent do level of English exposure, word position, and gender account for variability in the production of target /v/ and /tʃ/ in real words?*

The remainder of this chapter discusses participants, (Section 5.2) stimuli (Section 5.3), procedure (Section 5.4), predictions (Section 5.5), data analysis (Section 5.6), number of observations (Section 5.7), the classification process of the target segments (Section 5.8), the acoustic measurements of the target segments (Section 5.9), main results (Section 5.10), discussion (Section 5.11) and summary (Section 5.12).

[12] The task was piloted with a small set of participants using the same stimuli and procedure.

5.2 Participants

The same participants in the previous perception and production tasks also performed the production task reported in this chapter. To reiterate, the participants were 67 Saudi Arabic students (31 males and 36 females) at the University of Northern Border, divided by their expected level of English exposure into three exposure groups (high, medium and low) (see 3.2 for more details).

5.3 Stimuli

The stimuli consist of 29 real words: 21 words containing /v/ and 8 words containing /tʃ/. The selected real words vary in terms of the syllable structure (monosyllabic, disyllabic and trisyllabic). Fillers were not used in this production task because it was already very long even without fillers.

Table 5.1: Target real words

/v/			/tʃ/		
Video	Receiver	Glove	Cheesecake	Cappuccino	Clutch
Vanilla	Cover	Microwave	Chips	Ketchup	
Vaseline	Avocado	Live (adj)	Chocolate	Snapchat	
Vitamin	Boulevard		Chimpanzee		
Virus	Lavender				
Vase	Red velvet				
Van	Seven up				
Veto	Caravan				
Valentine	Mauve				

Initially, a set of English loanwords into Arabic were collected from previous studies (e.g., Al-Athwary, 2017; Abu Guba, 2016; Aloufi, 2016). Some of these words (e.g., video, virus, vitamin, chimpanzee) are standardised and included in contemporary Arabic lexicons (e.g., Omar, 2008) and dictionaries of Arabic for learners (e.g., Buckwalter & Parkinson, 2014) but some of these words (e.g., mauve, boulevard, cheesecake) are not standardised because the standardization process is very slow in Arabic countries (Alabdaly & Metwally, 2021). As explained more fully below, all the selected real words meet the following two criteria (Poplack, 2018):

a. Enjoying widespread use in the community.
b. Showing the same morphosyntactic behaviour as native words (e.g., شيبساتي chips –PL-1sPOSS, فديوات video.PL).

An online familiarity-rating questionnaire (see Appendix A) was designed using Qualtrics (www.qualtrics.com) to identify words that enjoy widespread use in the community (criterion a). For this reason, the target sounds did not have equal numbers of real words. The participants in the familiarity-rating questionnaire were 21 Saudi speakers of Arabic (males = 7; females = 14), from different age groups (ranging from 21 to 55, mean = 34.80; SD= 9.70). They were requested to rate how often they use, hear, or read each word using the ten-point scale. Words which received a median score of 5 or higher were selected (see Appendix A for more details of the familiarity rating questionnaire).

After meeting criterion a, the selected real words were then checked if they behave like Arabic words morphologically and syntactically. It was difficult to find any Arabic corpus that contains a large number of English loanwords, so a number of measures were taken to meet criterion b. To search for morphsyntactic integration, the researcher looked at Twitter data. Twitter was chosen because it is widely popular in Saudi Arabia; it reflects the most contemporary uses of Arabic. Figure 5.1 shows an example of a tweet containing the borrowed word 'chips' <شيبس>. In this example, 'chips' behaves like Arabic words phonologically, morphologically and syntactically. English /tʃ/ was replaced by Arabic /ʃ/ and the noun 'chips' <شبس> [ʃi:bs] was attached by the Arabic plural suffix <ات> [a:t] and the singular possessive pronoun <ي> [i:]. Additionally, both 'chips' and the following adjective 'favorite' <المفضلة> [ʔlmu:fəḍʿəlləh] show syntactic (feminine) agreement.

IPA Transcription: [wa:ħid min ʃi:bsa:ti: ʔlmu:fəd^{ʕ}əlləh ʔlgədi:məh ixtəfa: wmu: ga:drəh ʔlga:h illa bili:ma:ra:t]

Romanised transliteration[13]: <wa:Hid min shi:bsa:t: 2lmufaDallah 2lgadi:mah ixtafa: wmu: ga:drah 2lga:h illa bili:ma:ra:t>

Figure 5.1 A screenshot of a tweet in Saudi Arabic containing the borrowed word 'chips'

5.4 Procedure

The aim of the task is to elicit how Saudi Arabic speakers might produce target /v/ and /tʃ/ in real words. As stated in 1.3, all the participants started with this task, after which, they were tested in the other production task reported in Chapter 4 (where non-words were presented). Target real words were elicited via a fill-in-the-blank task. The frame sentences were not the same for each real word. The provided sentences, each of which was from 2 to 11 words in length, were written in Arabic (see Appendix C). Target real words were presented in a different random order for each participant. Pictures of familiar objects, along with some helpful hints, were used to elicit the missing words. Visualisations of target real words were used to eliminate the possible effect of orthography. Prior to the task, participants received three practice trials with distractor items to ensure that they understood the procedure.

Participants were asked to create a short recording, over their computers, and then play it back to test the functionality of their microphones. Gorilla (https://gorilla.sc) was also used in this task to present stimuli and record responses due to strict ban on in-person gathering at the time of study due to the COVID-19 pandemic. The reason why Gorilla was used to implement this task is explained in Chapter 3 (see 3.4). Participants were given as much time they needed to

[13] The romanised transliteration in this example was adopted from the Intonational Variation in Arabic Corpus (Hellmuth & Almbark, 2017).

complete the task, as they were able to control the start of each trial by pressing the 'next' button. Figure 5.1 shows two examples of this task

Figure 5.2: Screenshots of the real words task[14]

Initially, participants were asked to enter their anonymised codes, reading the instructions (Appendix D) and then signing the form. It was explained to participants that they should look at pictures, fill in the blanks with target real words and then read each sentence twice. Participants were instructed to finish this task along with the production task in Chapter 4 within the expected time (45-60 minutes) in a quiet room and to perform the task on a computer or laptop. However, three participants completed this task over a number of days and eight participants spent more than one hour to complete it. Appendix F contains a histogram that shows the distribution of time taken by the participants to complete the task. As mentioned in the previous perception and production tasks, some participants had no access to a computer or laptop and hence they used other devices (e.g., smart phones and tablets). Again, the study did not apply rigorous inclusion criteria because it was difficult at the time to recruit an appropriate number of participants with different levels of exposure to English and to incentivize them to complete a series of online tasks. Participants who completed this task along with the previous perception and production tasks were included in the study.

5.5 Predictions

This section explains predicted answers to the three specific questions derived from the main question in this chapter. The main question addressed in this chapter is: *to what extent do*

[14] Translation: ____ is my best app (hint: social media app), I do not like to eat _____ (hint: tropical fruit)

level of English exposure, word position, and gender account for variability in the production of target /v/ and /tʃ/ in real words?

The first specific question is: *is the variable production of target /v/ and /tʃ/ in real words affected by level of English exposure?* As discussed in Chapter 2, the selection of adapted sounds vs. imported sounds in loanwords can be affected by speakers' level of exposure to the source language (e.g, Kwon 2019; Poplack, 2018). Speakers with high levels of exposure to the source language may have more opportunities to hear or read imported sounds which, in turn, result in using imported sounds more often in their production. Therefore, we hypothesise that level of exposure to English may influence the likelihood of adaptation vs. importation. On the basis of the hypothesis, we can deduce the following prediction:

a) Saudi Arabic speakers with high levels of exposure to English are more likely to produce the imported sounds [v] and [tʃ] in real words than speakers with medium and low levels of exposure to English.

The second specific question is: *is the variable production of target /v/ and /tʃ/ in real words affected by word position?* As was hypothesised in the previous chapter, the likelihood of adaptation vs. importation may be affected by word position. Some studies on loanword phonology showed that the position of a novel sound within a word may affect how it is produced (e.g., Huang & Lin, 2016; Kubozono, Itô & Mester, 2008). Cross-linguistically, voiced obstruents are more marked than their voiceless counterparts and affricates are more marked than fricatives (Eckman & Iverson 1994). Therefore, the imported sounds [v] and [tʃ] may be more likely to be replaced respectively by the adapted sounds [f] and [ʃ] in word-final position. Based on the hypothesis, the following prediction can be put forward:

b) For target /v/ and /tʃ/ elicited in word-final position, Saudi Arabic speakers are more likely to produce the adapted sounds [f] and [ʃ] than the imported sounds [v] and [tʃ] in real words because [f] and [ʃ] are less marked in word-final position.

The third specific question is: *is the variable production of target /v/ and /tʃ/ affected by gender?* As mentioned in the last chapter, previous sociolinguistic studies in various Arabic dialects showed that women tend to adopt prestigious novel forms more than men (e.g, Almuhanadi, 1991; Assiri, 2014; Omari & Van Herk, 2016). In the same way, we expect that gender may influence the selection of one linguistic variant in loanwords. Namely, females are more likely to produce imported sounds than males when the source

language is prestigious. As discussed in Chapter 2, English enjoys a high status in Saudi Arabia. Thus, we hypothesise that the likelihood of adaptation vs. importation may be affected by gender. Based on this hypothesis, the following prediction can be put forward:

c) Females are more likely to produce the imported sounds [v] and [tʃ] than males in real words because English is prestigious language in Saudi Arabia.

5.6 Data Analysis

5.6.1 Acoustic Analysis

The acoustic analysis in this chapter aims to corroborate the results of the classification process. As explained in more detail in Chapter 4 (see 4.8), the cues for the distinction between [v] and [f] are friction duration, intensity, and the centre of gravity (Ogden, 2009) while the cues for the distinction between [tʃ] and [ʃ] are friction duration, and the amplitude rise time (Hayward, 2013). The acoustic measurements of the target sounds were extracted using a Praat script. In R, linear mixed effects models were run, using the *lmer* function in the *lme4* package (Bates et al. 2015), for each parameter with a parallel structure: acoustic parameter ~ segment + (1 |speaker) + (1 |word). P-values in the linear mixed effects models were calculated using the *LmerTest* package (Kuznetsova et al., 2017). If p-values are 0.05 or lower, acoustic differences between adapted sounds ([f] and [ʃ]) and imported sounds ([v] and [tʃ]) are significant.

5.6.2 Statistical Analysis of Main Results

Data for the main results was first summarised using the *tidyverse* package (Wickham, 2016) and then entered into mixed effects logistic regression using the *lme4* package (Bates et al. 2015) to estimate the probability of the binary outcome (imported sound vs. adapted sound) based on the different predictors. The predictor variables are level of English exposure (high, medium, and low), word position (initial, medial, and low), and gender (males and females). Mixed effects regression was used because it accounts for random effects and repeated tokens. In cases of non-convergence, the same steps in Chapter 4 (see 4.6) were employed. The first step was using the *all_fit* function from the *afex* package (Singmann et al. 2018) to test a variety of optimizers. If convergence was still not achieved, the second step was separating the intercept and slope using zero character as in (1| speaker) (0 + context | speaker) (Winter, 2019). If convergence was again not achieved, the final step was simplifying the model. Again, as indicated in 4.6, the inclusion of interactions was assessed via model comparison using the *anova* function. The simple model, without interactions, was adopted if the interactions were not

significant and there were no significant differences between the models (p >. 05). All categorical variables were dummy coded. For the target sounds, the imported sounds [v] and [tʃ] were coded as 0, and the adapted sounds [f] and [ʃ] were coded as 1. That is, positive values indicated a higher probability for the adapted sounds than the imported sounds. The pairwise predictions of constructed models were estimated using the *emmeans* package (Lenth, 2021). R codes used in the task are given in a footnote in the next section.

5.7 Number of observations

In total, 2,814 tokens for /v/ (21 words × two times × 67 participants) and 1,072 /tʃ/ tokens for /tʃ/ (8 words× two times × 67 participants) were expected to be collected. However, 999 tokens for /v/ and 383 tokens for /tʃ/ were collected because participants did not produce all the target real words (e.g., some participants produced other Arabic words). Only 978 tokens for /v/ and 375 tokens for /tʃ/ were included. For the sake of this study, 21 tokens for /v/ (2.10%) and 8 tokens for /tʃ/ (2.9%) were excluded because they underwent other consonantal substitutions or syllabic modifications (e.g., epenthesis and deletion).

5.8 Classification Process of Target Sounds.

This section briefly explains the classification process and reports /v/ and /tʃ/ realizations as imported sounds and adapted sounds in target real words.

The classification process was based on acoustic representation and auditory judgment. The criteria used to classify target /v/ and /tʃ/ into imported sounds and adapted sounds are explained in detail in Chapter 4 (see 4.8). Participants' productions were manually coded as one of either [v~f] or [ʃ~tʃ], with reference to the waveform and spectrogram in Praat (Boersma & Weenink, 2016). Tables 5.2 and 5.3 show the acoustic and impressionistic annotations for the target sounds in the three word positions.

Table 5.2: Total number of annotations for [v] and [f] in target real words

Annotation	Initial	Medial	Final	Total
Acoustically annotated as [v]	118	122	26	266 (26.63%)
Acoustically annotated as [f]	313	286	None	599 (59.96 %)
Impressionistically annotated as [f]	None	None	113	113 (11.31 %)
Others				21 (2.10 %)

Table 5.3: Total number of annotations for [tʃ] and [ʃ] in target real words

Annotation	Initial	Medial	Final	% of Annotations
Acoustically annotated as [tʃ]	144	59	39	242 (63.19%)
Acoustically annotated as [ʃ]	90	33	2	125 (32.64 %)
Impressionistically annotated as [ʃ]	8	None	None	8 (2.09 %)
Others				8 (2.09 %)

5.9 Acoustic Measurements of Imported and Adapted Sounds in Real Words

As in the previous chapter, the acoustic analysis here aims to corroborate the results of the classification process. This section presents the acoustic measurements of the target sounds in real words with descriptive statistics and inferential statistics.

Several measurements for the target sounds were made and compared across the three different word positions (initial, medial and final). To reiterate, the cues for the distinction between [v] and [f] are friction duration, intensity, and the centre of gravity (Ogden, 2009). The cues for the distinction between [tʃ] and [ʃ] are friction duration and amplitude rise time (Hayward, 2013).

5.9.1 Acoustic Measurements of [v] and [f]

The acoustic results of a comparison between [v] and [f] are similar to those reported in Chapter 4 (see 4.9.1). Overall, [v] had a shorter duration, lower COG, and greater average intensity than that of [f]. Friction duration, intensity, and the centre of gravity for [v] and [f] in the three words positions are reported in Tables 5.4 and 5.5 and visualized in Figures 5.2, 5.3 and 5.4. As shown in Tables 5.6, 5.7, and 5.8, linear regression models showed that the differences between the two sounds were statistically significant in each parameter.

Table 5.4: Duration, COG, and intensity for [v] in target real words across the three word positions

	[v]					
	Initial		Medial		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	84.94	25.90	72.89	16.82	102.26	46.08
COG values in Hz.	2575.257	1482.922	1948.651	1202.768	2586.93	1483.749
Intensity values in dB.	62.05	7.72	62.36	7.09	49.90	7.99

Table 5.5: Duration, COG, and intensity for [f] in target real words across the three word positions

	[f]					
	Initial		Medial		Final	
	M	SD	M	SD	M	SD
Duration values in ms.	120.62	36.16	103.09	21.18	218.37	83.49
COG values in Hz.	2919.526	1590.837	2610.771	1685.147	3163.416	2285.199
Intensity values in dB.	51.18	8.61	53.17	8.40	47.25	9.37

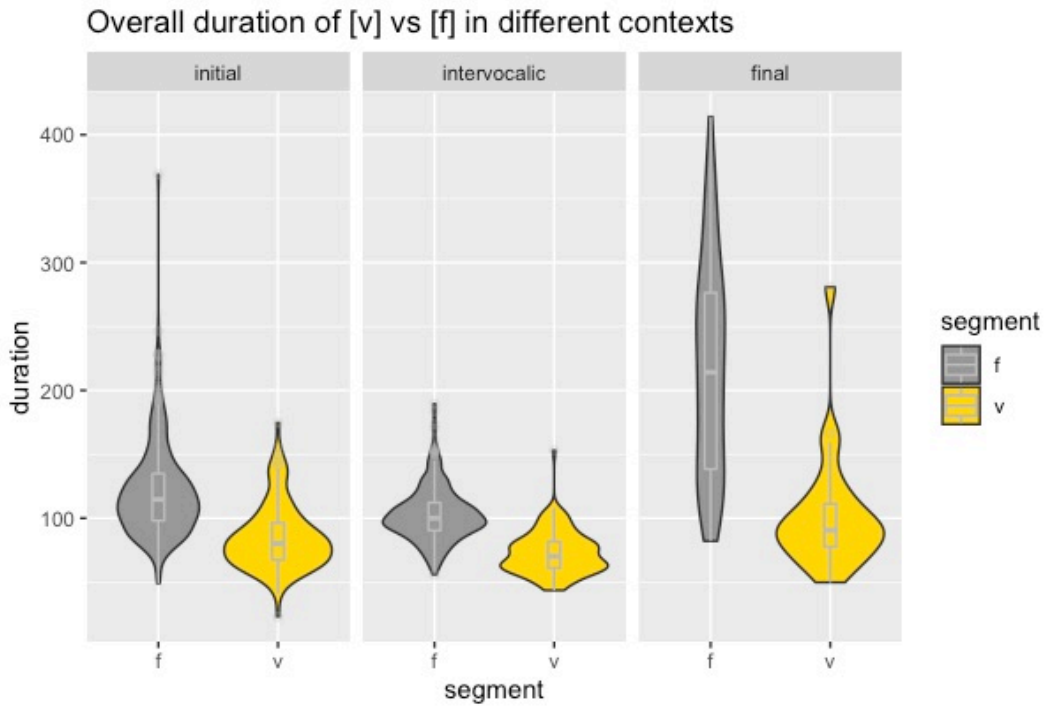


Figure 5.3: Friction duration for [v] and [f] in target real words across the three word positions.

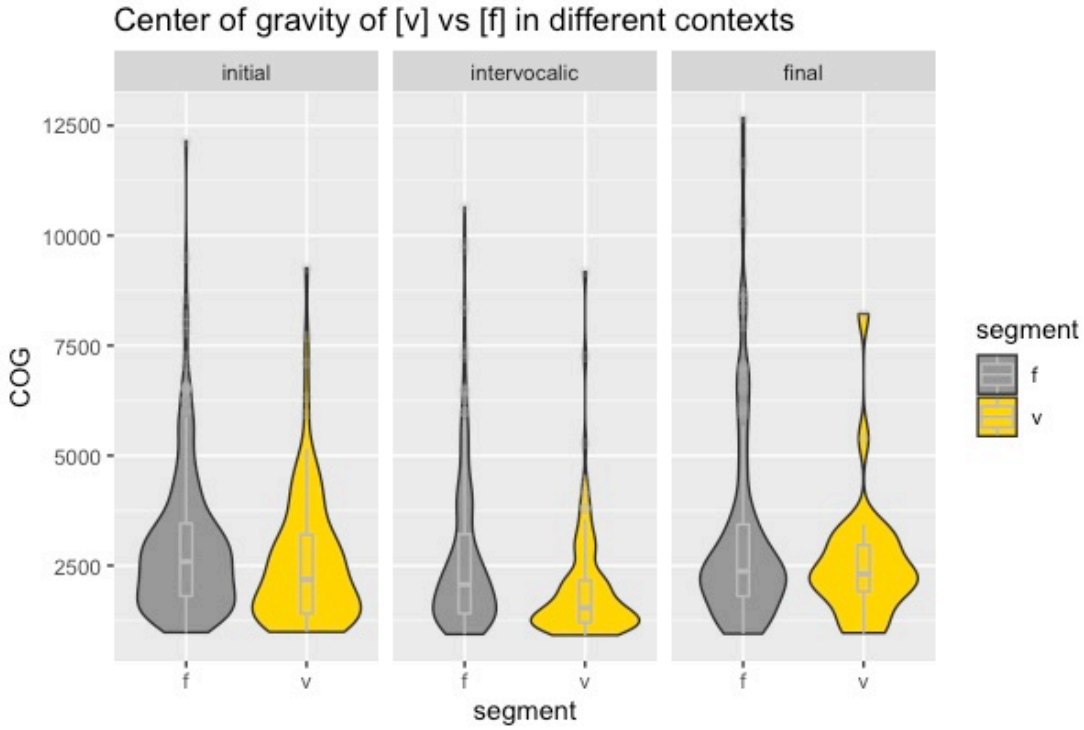


Figure 5.4: Centre of gravity of [v] and [f] in target real words across the three word positions.

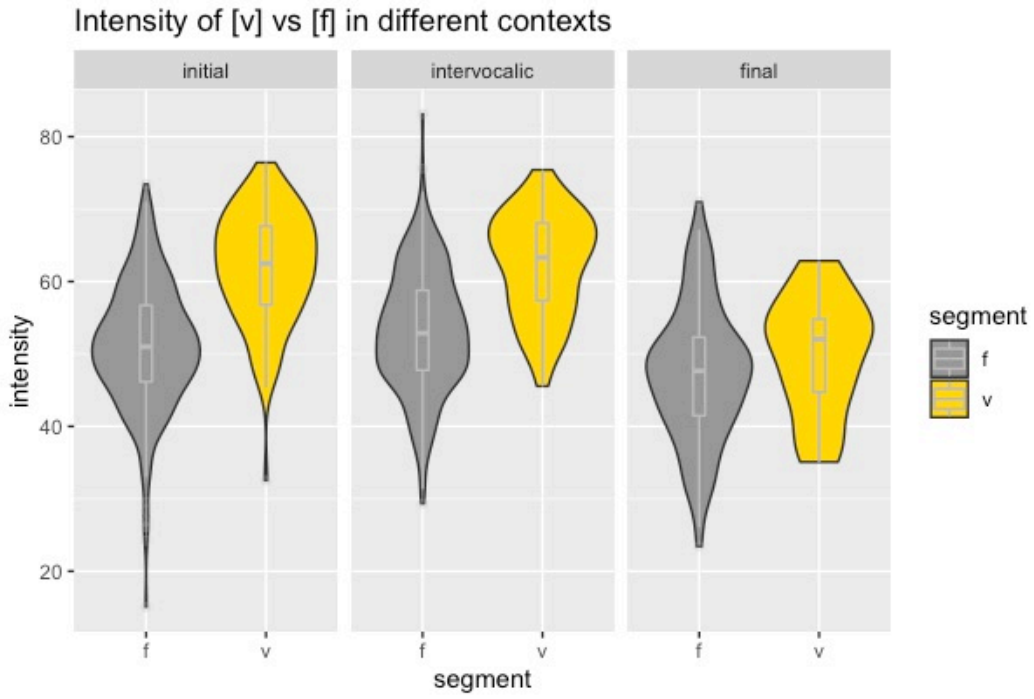


Figure 5.5: Mean intensity for [v] and [f] in target real words across the three word positions

Table 5.6: Linear mixed effects model for the distinction between [v] and [f] in duration in target real words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	109.364	8.397	21.060	13.03	1.52e-11 ***
Segment1	19.729	1.636	786.879	12.06	< 2e-16 ***

Table 5.7: Linear mixed effects model for the distinction between [v] and [f] in COG in target real words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	2655.37	160.88	73.49	16.506	< 2e-16 ***
Segment1	180.53	58.50	973.84	3.086	0.00209 **

Table 5.8: Linear mixed effects model for the distinction between [v] and [f] in intensity in target real words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	55.7154	0.9978	69.4184	55.84	<2e-16 ***
Segment1	-4.0789	0.2701	952.5658	-15.10	<2e-16 ***

5.9.2 Acoustic Measurements of [tʃ] and [ʃ]

The acoustic results for [tʃ] and [ʃ] are also similar to those reported in Chapter 4 (see 4.9.2), [tʃ] had a shorter friction duration and rise time than that of [ʃ]. Friction duration and the amplitude rise time for [tʃ] and [ʃ] in the three word positions are reported in Tables 5.9 and 5.10 and visualized in Figures 5.5 and 5.6[15]. As shown in tables 5.11, and 5.12, the mixed effects models revealed significant differences between [tʃ] and [ʃ] in each parameter.

Table 5.9: The friction duration and amplitude rise time for [tʃ] in target real words across the three word positions.

	[tʃ]					
	Initial		Medial		Final	
	M	SD	M	SD	M	SD
Duration values in ms	81.62	23.21	88.37	24.43	197.73	70.09
Mean rise time values in ms	38.89	27.38	38.00	25.65	55.23	33.84

Table 5.10: The friction duration and amplitude rise time for [ʃ] in target real words across the three word positions.

	[ʃ]					
	Initial		Medial		Final	
	M	SD	M	SD	M	SD
Duration values in ms	150.12	31.64	137.21	29.21	171.46	11.15
Mean rise time values in ms	81.51	45.51	82.21	35.17	49.68	65.07

[15] Note that there are only box plots for the final [ʃ] (without violin plots), the reason for this is that there are only two words ending with [ʃ] in word final position.

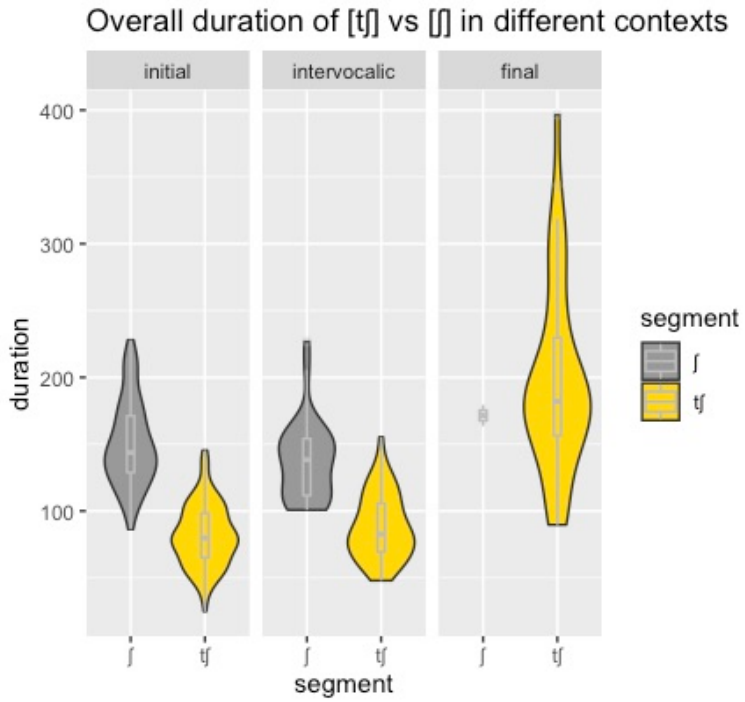


Figure 5.6: Durations of [tʃ] and [ʃ] in target real words across the three word positions

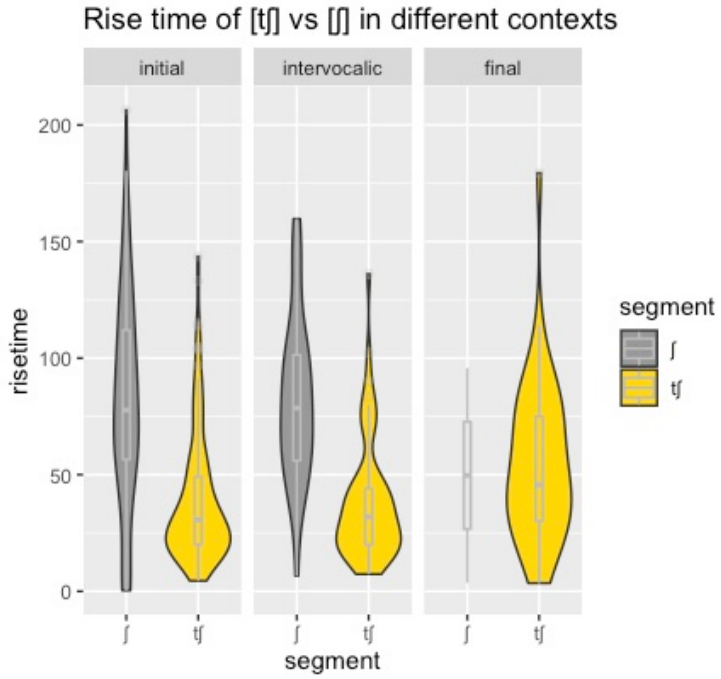


Figure 5.7: Amplitude rise time of [tʃ] and [ʃ] in target real words across the three word positions

Table 5.11: Linear mixed effects model for the distinction between [tʃ] and [ʃ] in duration in target real words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	130.284	14.330	7.256	9.092	3.2e-05 ***
Segment1	31.947	2.393	369.262	13.350	< 2e-16 ***

Table 5.12: Linear mixed effects model for the distinction between [tʃ] and [ʃ] in amplitude rise time in target real words

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	62.844	3.466	10.869	18.13	1.81e-09 ***
Segment1	21.880	2.176	84.860	10.05	4.08e-16 ***

5.10 Main Results

This section presents the results concerning /v/ and /tʃ/ realizations (i.e., imported sounds [v] and [tʃ] vs. adapted sounds [f] and [ʃ]) in target real words using both descriptive statistics and mixed-effects logistic regression.

5.10.1 Descriptive analysis

Figures 5.7 and 5.8 show the proportion of /v/ and /tʃ/ realizations in target real words. Tables 5.13-5.15 show the raw results split by key independent variables. Regarding /v/ realizations in target real words, Figure 5.7 shows differences between the three exposure groups. Speakers in the high exposure group showed the highest level of importation rate (36.41%), followed by speakers in the medium and low exposure groups (32.68% and 9.56%, respectively). In terms of gender, males showed a lower importation than females (18.88% and 32.76%, respectively). With regard to word position, the importation rate of /v/ was lower in word-final position (18.71%) than in word- initial and medial positions. Figure 5.7 may suggest an unexpected interaction between word position and group; speakers in medium exposure group showed a lower importation in word-final position than speakers in the other two exposure groups.

Moving to /tʃ/ realizations in target real words, speakers in the high exposure group showed the highest importation rate (67.79%), followed by speakers in the medium and low exposure groups (66.96% and 57.89%, respectively). Regarding gender, females had a higher

importation rate than males (68.37% and 59.38% respectively). In terms of word position, unlike target /v/, the highest importation rate was in word-final position (95.12%). However, we should keep in mind that target /tʃ/ occurred word-finally only in one word. To reiterate, this was because target real words were chosen based on a familiarity rating questionnaire; thus, the two target sounds did not have equal numbers of real words.

Table 5.13: Summary of speakers' production of /v/ and /tʃ/ in target real words by level of English exposure

	[v]		[f]		[tʃ]		[ʃ]	
	count	%	Count	%	Count	%	count	%
High group	138	36.41%	241	63.59%	101	67.79%	48	32.21%
Medium group	100	32.68%	206	67.32%	75	66.96%	37	33.04%
Low group	28	9.56%	265	90.44%	48	57.89%	66	42.11%

Table 5.14: Summary of speakers' production of /v/ and /tʃ/ in target real words by gender

	[v]		[f]		[tʃ]		[ʃ]	
	count	%	count	%	count	%	count	%
Females	192	32.76%	394	67.24%	147	68.37%	68	31.63%
Males	74	18.88%	318	81.12%	95	59.38%	65	40.62%

Table 5.15: Summary of speakers' production of /v/ and /tʃ/ in target real words by word position

	[v]		[f]		[tʃ]		[ʃ]	
	Count	%	count	%	count	%	count	%
Word-initial position	118	27.38%	313	72.62	144	59.5%	98	40.5%
Word- medial position	122	29.9%	286	70.1	59	64.13%	33	35.87%
Word-final position	26	18.71%	113	81.29%	39	95.12%	2	4.88%

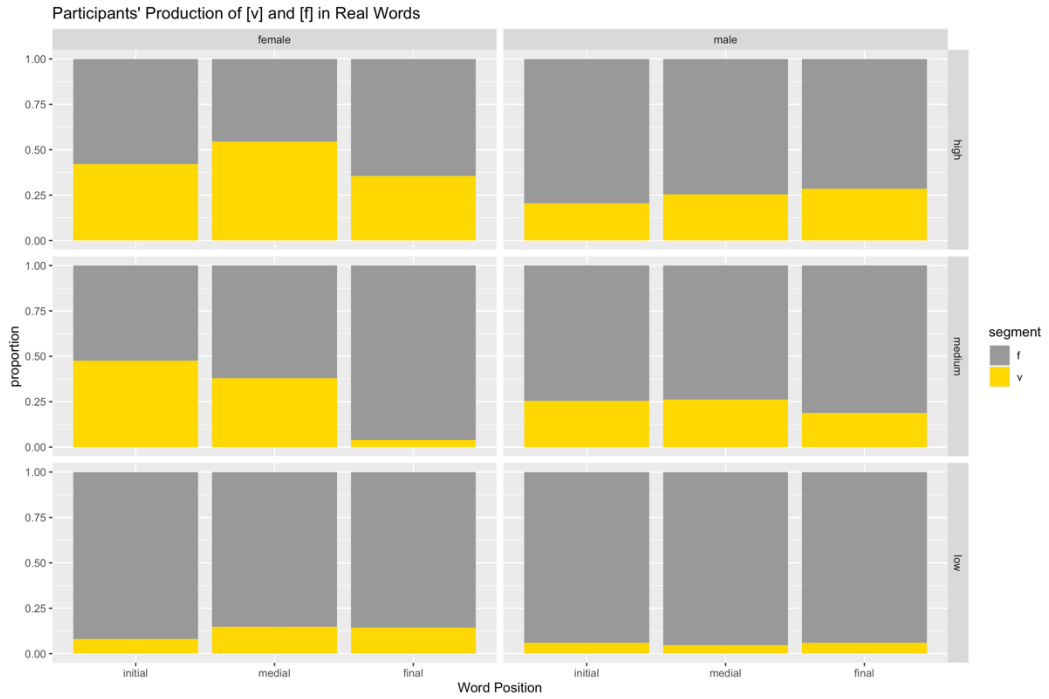


Figure 5.8: Proportion of /v/ realizations in target real words by level of English exposure, word position and gender

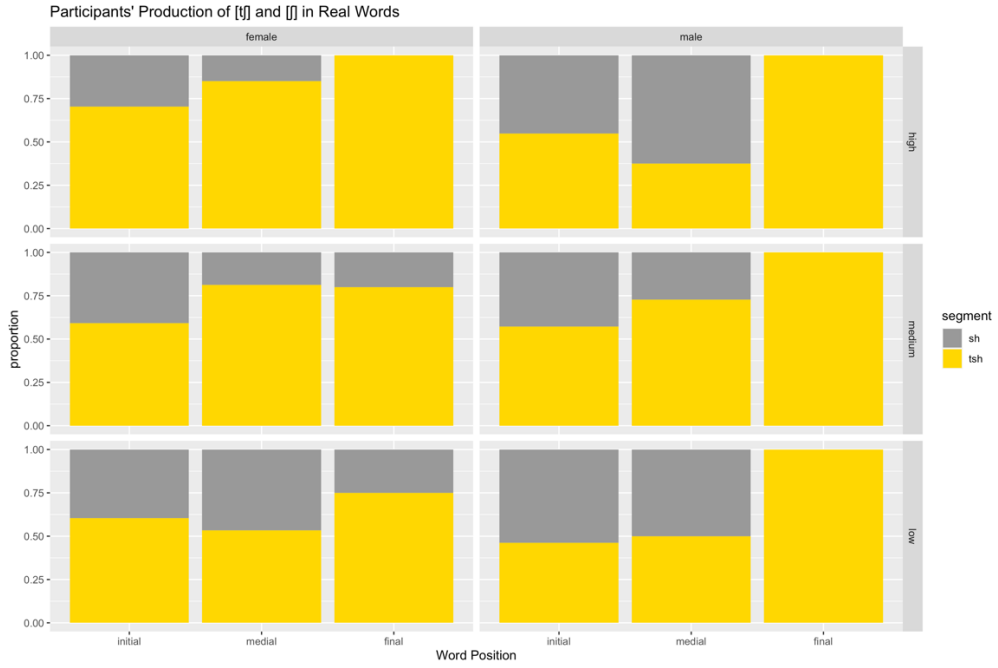


Figure 5.9: Proportion of /tʃ/ realizations in target real words by level of English exposure, word position and gender

5.10.2 Logistic Regression Analysis

Having presented the descriptive results, we turn to discuss in detail the results of the models constructed for the two target sounds. To reiterate, the target sounds were dummy coded (the imported sounds [v] and [tʃ] = 0 and the adapted sounds [f] and [ʃ] = 1). Wherever possible, random slopes were included, and interactions between fixed factors were excluded if they were not significant and did not improve the model fit.

The results of the mixed effects model[16] for target /v/ in real words are presented in Table 5.16 and visualised in Figure 5.9. The model involves gender (females and males) and a two-way interaction between word position (i.e., initial, medial, and final) and language exposure group (high, medium, and low). There are random intercepts for word and speaker. This simple model was adopted because the model with random slopes (word position by speaker and item) failed to converge. The estimate for the intercept is the estimate for low exposure group, males, and word-final position. The model shows no main effect of word position. The estimate for the reference level which is word-final position is higher than estimates for word-initial and medium positions (β = 0.9217, SE = 0.8085, z = 1.140, p = 0.254289) and (β = 0.2342, SE =0.7708, z = 0.304, p = 0.76), respectively, but the differences between the three positions are statistically not significant. Additionally, level of English exposure is not a significant predictor. The difference between the high and low exposure groups approached but did not reach statistical significance (β = -1.4952, SE = 0.8841, z = -1.691, p = 0.09). Likewise, the difference between the medium and low exposure groups is not significant (β = 0.6648, SE = 0.9941, z = 0.669, p = 0.50). In terms of the effect of gender, the difference between males and females is significant (β = -1.4761, SE = 0.5300, z = -2.785, p = 0.005350 **), suggesting that the imported sound [v] was more likely to be produced by females than males. Finally, there is an interaction between language exposure group and word position for the medium group in word-initial position (β = -3.3473, SE = 0.9244, z = -3.621, p = 0.00***) and in word-medial position (β = -2.3641, SE = 0.8940, z = -2.644, p = 0.00 **), meaning that speakers in the medium exposure group were more likely to produce the imported sound [v] in these two positions than speakers in the high and low exposure groups.

[16] glmer(segment~ context*group + gender + (1 |speaker) + (1 | word),data = loanwords_v, family = "binomial", control = glmerControl(optimizer = "bobyqa"))

Table 5.16: Summary of mixed effects logistic regression model for [v] and [f] in target real words

Fixed Effects	Estimate	SE	z value	Pr(>|z|)
Intercept	3.5937	0.8732	4.115	86e-05 ***
Initial position	0.9217	0.8085	1.140	0.254289
Medium position	0.2342	0.7708	0.304	0.761290
High group	-1.4952	0.8841	-1.691	0.090776 .
Medium Group	0.6648	0.9941	0.669	0.503611
Females	-1.4761	0.5300	-2.785	0.005350 **
Initial position: High group	-0.8092	0.8070	-1.003	0.315997
Medium position: High group	-0.8671	0.7757	-1.118	0.263674
Initial position: Medium group	-3.3473	0.9244	-3.621	0.000293 ***
Medium position: Medium group	-2.3641	0.8940	-2.644	0.008185 **

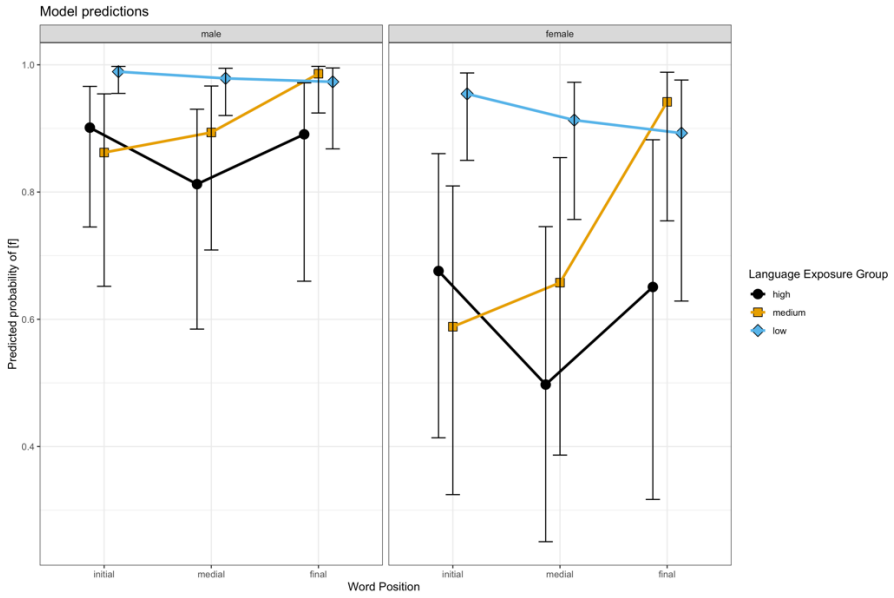


Figure 5.10: Predicted probability of [f] production in target real words by word position, language exposure group and gender

Turning to target /tʃ/ in real words, Table 5.17 and Figure 5.10 present the results of the mixed effects model[17]. The adopted model includes word position (initial, medial and high), language exposure group (high, medium and low), and gender (males and females) as fixed effects. As random effects, there are random intercepts for speaker and word. Interactions between the fixed factors were not included because they were not significant, and they did not improve the model fit. The estimate for 'intercept' is the estimate for the low exposure group, males and word-final position. The model reveals no main effect of word position, indicating that the importation rate of target /tʃ/ is similar in the three positions. However, the effect of English exposure is significant. The high exposure group (β = -1.3694, SE =0.5880, z = -2.329, p =0.01 *) and medium exposure group (β = -1.2478, SE = 0.6084, z = -2.051, p = 0.04 *) are significantly different from the low exposure group. These results suggest that the speakers in the low exposure group were more likely to produce the adapted sound [ʃ] than speakers in the other two exposure groups. Also, the model confirms the main effect of gender (β = -1.1893, SE = 0.4893, z = -2.431, p = 0.0*). Females were more likely to produce the imported sound [tʃ] than males.

Table 5.17: Summary of mixed effects logistic regression model for [tʃ] and [ʃ] in the real words

Fixed Effects	Estimate	SE	z value	Pr(>|z|)
Intercept	-2.6625	3.5543	-0.749	0.4538
Initial position	3.0709	3.8848	0.790	0.4292
Medial position	0.6681	4.5192	0.148	0.8825
High group	-1.3694	0.5880	-2.329	0.0199 *
Medium Group	-1.2478	0.6084	-2.051	0.0403 *
Females	-1.1893	0.4893	-2.431	0.0151 *

[17] glmer(segment ~ context + group + gender + (1 | speaker)+ (1 | word),data = loanwords_ch, family = "binomial")

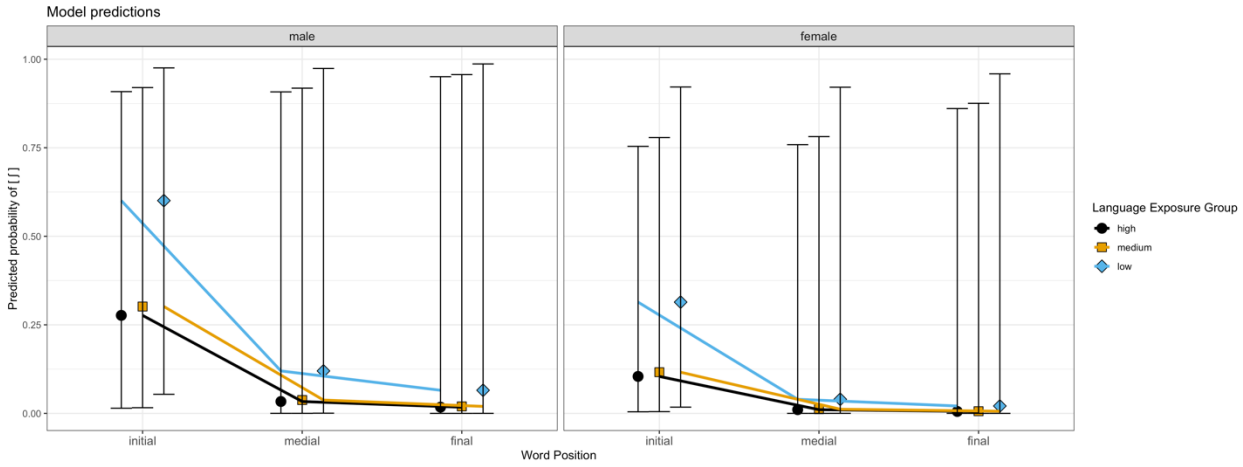


Figure 5.11: Predicted probability of [ʃ] production in target real words by word position, language exposure group and gender

5.10.3 Interim Summary

This subsection summarises the significant factors that influenced the production of target /v/ and /tʃ/ in real words. Table 5.18 shows only the statistically significant results obtained from the mixed effects logistic regression models for the two target sounds.

In real words, some factors appeared to be non-significant. For example, word position and level of English exposure did not affect the importation rate of target /v/ in real words. However, there was a significant interaction between language exposure group and word position: speakers in the medium exposure group were more likely to produce the imported sound [v] in word-initial and medial positions than in word-final position. Gender had a significant main effect: females were more likely to produce the imported sound [v] than males.

Concerning target /tʃ/ in real words, gender had also a main effect: females were more likely to produce the imported sound [tʃ] than males. Additionally, language exposure group was a significant factor: speakers in both high and medium exposure groups were more likely to produce the imported sound [tʃ] than speakers in the low exposure group.

Table 5.18: Factors influencing the production of target /v/ and /tʃ/ in real words

	Exposure Group reference = low		Word position reference = final		Gender reference = male	Exposure Group*Word Position	
	high	medium	initial	Medial	female	initial* medium	medium* medium
/v/					**	***	**
/tʃ/	*	*			*		

Asterisks refer to levels of significance: * ≤ 0.05, ** ≤ 0.01, *** ≤ 0.001
Reference refers to the reference level (intercept)

5.11 Discussion

This section aims to discuss several possible explanations for the key findings in relation to our predictions and previous literature. To reiterate, the production task in this chapter aims to examine the extent to which level of English exposure, word position and gender account for variability in the production of target /v/ and /tʃ/ in real words.

5.11.1 Effect of Language Exposure

It was predicted that speakers with high levels of exposure to English are more likely to produce the imported sounds [v] and [tʃ] than speakers with medium and low levels of exposure to English (Prediction a). The results showed that level of English exposure significantly influenced the importation rate of /tʃ/ but not /v/. Although speakers in the high exposure group descriptively showed higher percentages of importation for target /v/ than speakers in the medium and low exposure groups, the model showed that the likelihood of [v] production was not affected by level of exposure to English. Unlike target /v/, speakers in the high and medium exposure groups were significantly more likely than speakers in the low exposure group to produce the imported sound [tʃ]. These results provide only partial support for previous work showing that the importation rate becomes higher with increased exposure to the source language (e.g., Kang, 2021; Poplack, 2018). However, although language exposure was not a significant factor for target /v/, speakers in high exposure group, as shown in Table 5.16, tended to produce the imported sound [v] more than speakers in the low exposure group ($p = 0.09$).

Different explanations are possible for the non-significant impact of language exposure on the importation rate of target /v/. It is possible that speakers with the high level of English exposure paid less attention to the phonetic details in real words compared to non-words because

the acoustic and orthographic cues for target sounds in this task are unavailable. In other words, they did not hear target sounds produced by a native speaker or read their orthographic representations. This unexpected finding maybe linked to the list of real words used for the two sounds. The list of real words for /tʃ/ was small and contained more recent borrowed words (e.g, snapchat, cheesecake, and chips). In contrast, the list of real words for /v/ contained borrowed words (e.g., video, virus, receiver, and vitamins) that have been known and used in the media and public for many years ago. Some of these words are even included in dictionaries for Arabic learners (e.g., A frequency dictionary of Arabic: core vocabulary for learners, Buckwalter & Parkinson, 2014). Previous literature suggests that the longer and more commonly a loanword is used in a speech community, the more it becomes similar to native words. Thus, with frequent use of a loanword, a novel sound is more likely to be adapted to a sound in the L1 (Dohlus, 2010; Poplack, 2018).

5.11.2 Effect of Word Position

The importation rate of target /v/ and /tʃ/ was expected to be lower in word-final position than in other positions (Prediction b). Contrary to our prediction, the results showed that word position did not have a main effect on the production of target /v/ and /tʃ/ in real words, meaning that the importation rate for the two sounds is similar in the three word positions. However, there was an interaction of group with word position in target /v/ production, whereby speakers in the medium exposure group were more likely to produce the adapted sound [f] in word-final position than in other positions. However, as shown in Figure 5.9, there is a large degree of variability in word-initial and medium positions within speakers in the medium exposure group, suggesting that some speakers were also less likely to produce the imported sound [v] in these two positions.

This finding was unexpected because the results of the previous production task in Chapter 4 showed that target /v/ was more likely to be produced as [f] in word-final position. Again, the possible explanation for this finding might be particular real words used for target /v/. When a loanword becomes more widespread in a speech community, a novel sound is more likely to be replaced by a sound in the L1 (Dohlus, 2010; Poplack, 2018). The list of real words for target /v/ contained words that are repeatedly used in the media and public. Target /v/ exists in different positions in these words (e.g., virus, vitamins, and receiver); therefore, there was less room for word position to affect the production of /v/. That is, participants tended to produce /v/ as /f/ in different word positions.

5.11.3 Effect of Gender

As predicted, females were more likely to produce the imported sounds [v] and [tʃ] than males (Prediction c). Interestingly, the effect of gender was more robust in real words than in non-words. The possible reason for the impact of gender is that females and males had different attitudes towards the English language. Previous literature suggests that attitudes can influence how a speaker may select one variant rather than the other (e.g, Babel, 2010; Drager et al., 2010; Paradis and LaCharite, 2012).

As discussed in Chapter 2, the English language is prestigious (Habbash & Troudi, 2014). This finding is in line with previous studies on different Arabic dialects indicated that women tend to use novel forms more often than men (e.g., Almuhanadi, 1991; Assiri, 2014; Omari & Van Herk, 2016). This effect of gender possibly extended to the use of novel sounds in loanwords. However, it is worth noting that it is unclear if this tendency to produce [v] and [tʃ] in loanwords also includes older female speakers given that this study examined only young female speakers.

5.12 Summary

The present production task examined the impact of level of English exposure (high, medium and low), word position (initial, medium and final), and gender on the production of target /v/ and /tʃ/ in real words. All the participants in the previous perception and production tasks also performed the production task reported in this chapter. The production of the target sounds was elicited via a fill-in-the-blank task. Pictures, along with some helpful hints, were used to elicit the missing words. The stimuli were English loanwords into Arabic. Due to the strict COVID-19 ban on in-person data collection at the time, the task was implemented online using Gorilla (https://gorilla.sc). The acoustic analysis in this chapter was used to corroborate how we labelled target sounds. Main results were analysed using the mixed effects logistic regression; the dependant variables were the imported sounds [tʃ] and [v] and adapted sounds [f] and [ʃ].

The findings showed that level of English exposure had a main effect on /tʃ/ but not /v/. Speakers in the high and medium exposure groups were more likely to produce [tʃ] than speakers in the low exposure group. The null result for target /v/ could be attributed to the list of real words used for the two sounds. The list of real words for /tʃ/ involved more recent borrowed words compared to the list of real words for /v/. Unlike non-words, word position was not a

significant factor: the importation rate of target /v/ did not differ in the three positions. Once more, this result might be due to the list of real words for /v/ which contained common borrowed words that found their way into the media and public many years ago. It is possible that these words were considered as Arabic words rather than borrowed words, and hence /v/ was more likely to be produced as [f], no matter where it exists within a word. Finally, gender was a significant factor: females were more likely to produce the imported sounds [v] and [tʃ] than males. A possible explanation for this result is that females had more positive attitudes than males towards the status of English as prestigious language.

These findings suggest that language exposure and word position had more robust effect on the production of target /v/ and /tʃ/ in non-words than in real words. However, gender played a key role in the production of the two target sounds in real words. In order to disambiguate the observed impact of gender on the production of target /v/ and /tʃ/, the next chapter will examine differences in attitudes towards the English language and American culture between females and males using survey data from the same participants in this production task.

6 Attitudes Survey

6.1 Introduction

The most obvious finding from the production tasks was the impact of level of language exposure: the lower levels of exposure to English resulted in a lower importation rate of target /v/ and /tʃ/. It is well-known that language exposure plays an important role in L2 production and perception (e.g., Kang & Schertz, 2021; Kown, 2017; Poplack, 2018;). Individuals with low levels of L2 exposure are more likely to show low rates of L2 production and perception accuracy. However, this might not be the only reason. If someone values identity as an Arabic speaker, then the substitution of target /v/ and /tʃ/ for the native sounds [f] and [ʃ] may be more prevalent in their production. In contrast, substitution may be less prevalent if one does not highly value that identity. Another important finding of the previous production task was that female participants were more likely to produce the imported segments [v] and [tʃ] in real words than male participants. Gender, as a biological construct, does not affect production accuracy (e.g., Edwards, 2008). As mentioned in the previous chapter, a possible reason for this finding might be that female participants had more positive attitudes towards the English language and American culture than male participants. To find evidence of whether differences in attitudes could be part of the explanation for the earlier findings, the results obtained from a short attitudes questionnaire, completed by the same participants in the perception and production tasks, will be analysed in this chapter.

Previous literature on loanword phonology shows that attitudes can strongly influence the importation of novel sounds in loanwords (e.g., Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2011). For example, Paradis and LaCharite (2012) write "importations reflect a willingness to imitate the sounds of the L2, typically because this language is considered superior, friendly or dynamic" (p.98). Identity and attitudes can influence how a speaker may select one variant rather than the other (e.g., Babel, 2010; Drager et al., 2010). Substitution of native sounds for non-native sounds is also more likely to occur if the foreign language is seen as a threat to the native culture. Paradis and LaCharite (2012) cite Mackey's (1989) study which showed that English interdentals were adapted as [s] and [z] by European French speakers and as [t] and [d] by Quebec French speakers. The European French borrowers attempted to preserve the non-native segments but failed to master articulations as a result of 'flawed importation'. The production of English interdentals /θ/ and /ð/ did not reflect the impact of the native phonology

which could result in /t/ and /d/, or perception which could result in /f/ and /v/. Paradis and LaCharite (2012) noted that this flawed importation is due to the high prestige of English in France where English is not seen as threatening; whereas Quebec French speakers did not attempt to retain the English interdentals because English is imposed and threatening in Quebec.

The aim of the present chapter is to explore whether there are differences in attitudes towards the English language and American culture based on gender and language exposure. The analysis in this chapter can serve to support our explanation of the impact of gender in the real words production task; namely, that female participants were more likely to produce the imported sounds [v] and [tʃ] because they had more positive attitudes towards the English language than male participants. It may also provide further explanation for the impact of language exposure on the likelihood of producing the imported segments [v] and [tʃ]. A possible reason for the impact of language exposure may lie in participants' attitudes towards the English language and American culture. One unexpected finding in this study was that female and male participants with low levels of English exposure were less likely to perceive and produce target /tʃ/ in aural-only condition than male and female participants in the high exposure group despite the fact that all groups have high exposure to Arabic dialects that have this sound. The participants with low levels of English exposure were less likely to produce the imported segments [v] and [tʃ], probably because they had less positive attitudes than the participants with high levels of English exposure.

The remainder of this chapter discusses participants, (Section 6.2), design of the questionnaire (Section 6.3), data collection procedure (Section 6.4), data analysis (Section 6.5), results (Section 6.6), discussion (Section 6.7) and summary (Section 6.8).

6.2 Participants

All the participants in the perception and production tasks also completed the attitudes survey reported in this chapter. As a reminder, the participants were 67 Saudi speakers, including both males (31) and females (36). They were divided by their expected level of English exposure into three exposure groups: high (24), medium (22) and low (21) (for more details, see 3.2).

6.3 Design of the questionnaire

The questionnaire was relatively short; it consisted of only eight statements. If a questionnaire is long, participants may have been more likely to respond randomly, given that the preceding perception and production tasks took a long time to complete. Following

Hashimoto (2019), the statements were adapted from a national survey conducted by Te Puni (2010) about New Zealanders' attitudes, values, and beliefs towards the Māori language and culture. In Hashimoto's (2019) study, the questionnaire is divided into two sections. The first section is about speaker-specific properties whereas the second section is about word-specific properties. As for speaker-specific properties section, participants were asked to indicate their agreement with 19 statements about the Māori Language, people and culture on a 0-100 scale. This questionnaire was selected in the present study because it contains questions asking for participants' agreement or disagreement with statements about one language and its culture (Māori) that has provided another language (New Zealand English) with a large number of loanwords.

For this study, the statements were adjusted to examine Saudi Arabic speakers' attitudes towards the English language and American culture. At the start of the analysis process, one statement was deleted because it was determined to be potentially ambiguous (i.e. open to different interpretations). The remaining seven statements are presented in Table 6.1.

Table 6.1: A questionnaire on attitudes towards the English language and American culture

S1	Well-spoken English is beautiful to listen to.
S2	I like a person who speaks English fluently.
S3	It is good to teach English to all pupils from the first grade of elementary school.
S4	People should make an effort to improve their English.
S5	It is ok for people to greet others in English.
S6	It is a good thing that people speak English in public places (e.g., cafes & restaurants)
S7	It would be nice learn about American values and culture.

A Likert scale was used in the original survey (Te Puni, 2010). However, following Llamas and Watt (2014), in this study, a Visual Analog Scale (VAS), which consists of a horizontal line, was employed. The Likert scale was avoided because pre-arranged response categories may impose restrictions on a participant's ability to express their opinions or feelings (Llamas & Watt, 2014). The participants were requested to drag a bar to select a numeric value. They indicated their level of agreement or disagreement with each statement on a horizontal line ranging from 0 (strongly disagree) to 100 (strongly agree) as shown in Figure 6.1.

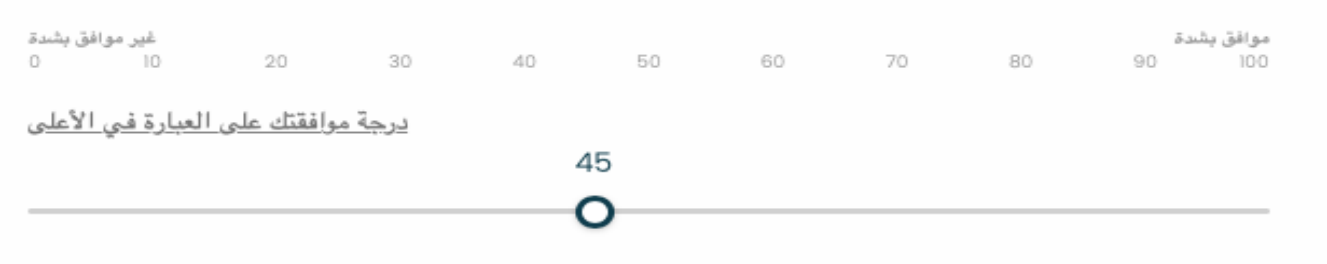


Figure 6.1: Visual analog scale

6.4 Data collection

Due to the strict COVID-19 ban on in-person data collection at the time, the questionnaire was implemented online. It is worth noting that a questionnaire was used as a tool to elicit participants' attitudes because it is easy to distribute and collect data online in a short period of time. Qualtrics (www.qualtrics.com), which is password-protected cloud-based software, was used to host the questionnaire and data collection. After performing the perception and production tasks, participants accessed the survey through a web link sent to them via the *WhatsApp* application. The link was sent to participants once they had finished the perception task.

When participants opened the link, they were requested to read an information sheet, sign the consent form and then enter the same participant code that was given to them by the researcher for use in the perception and production tasks. Participants were not allowed to proceed to answer the survey without entering their code and signing the consent form by clicking on the appropriate checkbox. The statements were presented all on one page with the same order of presentation for each participant. The expected length of time to complete the survey was approximately less than 10 minutes. However, the participants were given as much time as they needed to complete it. The survey took 5 minutes or less to complete per participant.

The survey was completed by all the participants who performed the previous perception and production tasks.

6.5 Data Analysis

R software (R Core Team, 2019) was used to run descriptive statistics and inferential statistics using linear regression. First, descriptive statistics were used to illustrate medians, means and standard deviations for the seven statements. The data was summarised and visualised using the *tidyverse* package (Wickham, 2016). Based on the raw data, the seven statements were divided into three categories as will be explained in further detail in 6.6.1. Following the

descriptive analysis, three linear regression models using the *lm* and *lmer* in the *lme4* package were run to examine whether participants' attitudes differ based on their gender and level of English exposure. Linear regression was used because the dependent variable is continuous (i.e., participants' rating for each statement ranging from 0 to 100). The predictor variables were sum coded. The *LmerTest* package (Kuznetsova et al., 2017) was used to calculate p-values. The simple model, without interactions, was adopted if the interactions were not significant and there were no significant differences between the models (p >. 05). The inclusion of interactions was assessed via model comparison using the *anova* function.

6.6 Results

This section reports descriptive statistics and presents the results of linear regression models.

6.6.1 Raw Results

Figure 6.2 shows participants' overall ratings for the seven statements. Table 6.2 provides means and standard deviations for each statement. Overall, as shown in Figure 6.2, the ratings tend to be at a high value; none of the statements had a median score lower than 50. However, there was a degree of variability in the last three statements. Whereas most participants felt that it is important to learn and be proficient in English (Statements 1-4), they differed in their views about using English in greetings and public places (Statements 5 and 6) and learning about American culture and values (Statement 7).

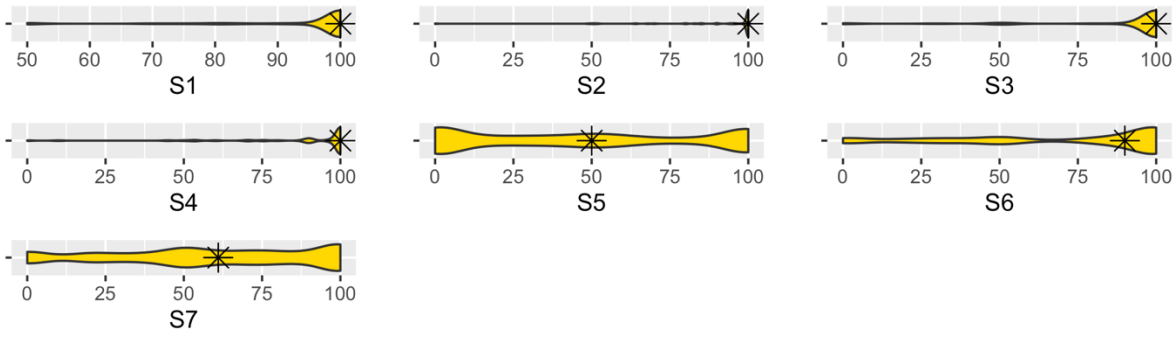


Figure 6.2: Participants' responses to the questionnaire

Table 6.2: The mean and standard deviation for each statement in the questionnaire.

Statement	Mean	SD
S1: Well-spoken English is beautiful to listen to.	95.56	11.46
S2: I like a person who speaks English fluently.	92.96	16.83
S3: It is good to teach English to all pupils from the first grade of elementary school.	92.33	21.66
S4: People should make an effort to improve their English.	90.01	19.94
S5: It is ok for people to greet others in English.	47.05	39.06
S6: It is a good thing that people speak English in public places	70.64	34.65
S7: It would be nice to learn about American values and culture.	59.93	33.13

Figure 6.3 displays the median scores of the seven statements by gender and level of English exposure. Table 6.3 presents means and standard deviations for the female and male participants' ratings in the three exposure groups.

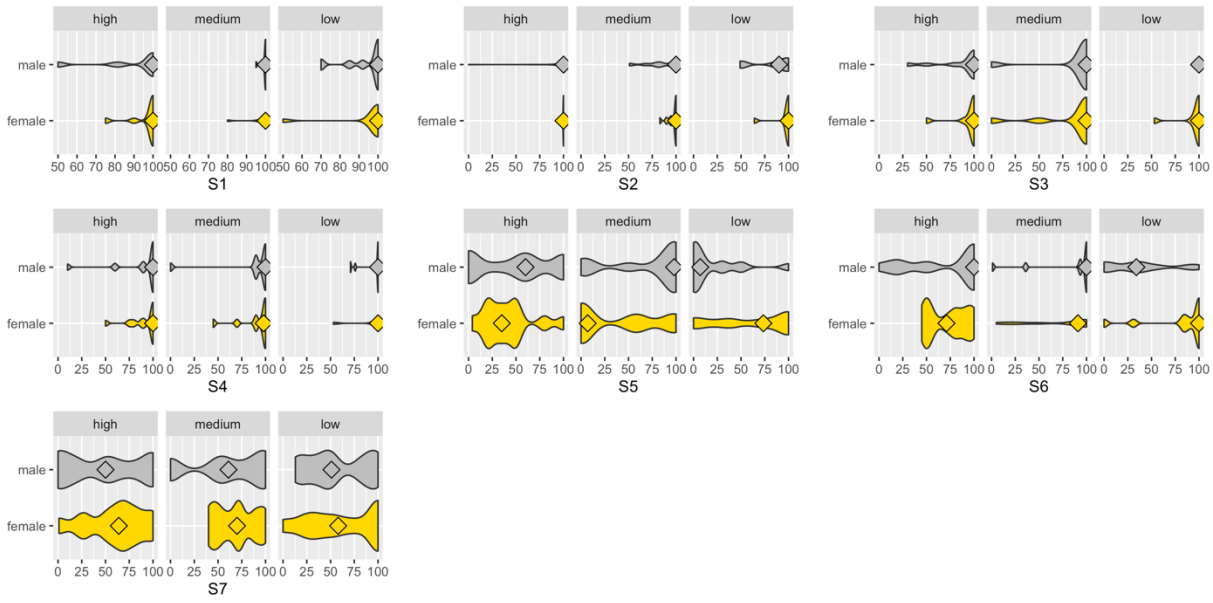


Figure 6.3: Questionnaire ratings by language exposure group and gender

Table 6.3: Means and standard deviations of the participants' ratings by group and gender[18]

	High Group		Medium Group		Low Group	
	female	male	female	male	female	male
S1: Well-spoken English is beautiful to listen to.	97.04 (7.26)	89.82 (17.72)	97.86 (5.71)	99.48 (1.52)	95.58 (14.21)	93.45 (10.56)
S2: I like a person who speaks English fluently.	99.57 (1.39)	89.57 (27.54)	96.92 (5.12)	86.27 (1.52)	96.81 (10.23)	82.85 (10.56)
S3: It is good to teach English to all pupils from the first grade of elementary school.	94.75 (14.14)	88.89 (22.06)	86.47 (30.69)	90.30 (29.64)	95.28 (13.31)	100 (0)
S4: People should make an effort to improve their English.	90.50 (14.55)	87.36 (25.57)	90.24 (16.52)	87.11 (28.91)	91.78 (17.76)	93.24 (11.40)
S5: It is ok for people to greet others in English.	41.04 (25.23)	48.23 (39.12)	38.93 (40.78)	71.18 (39.15)	61.35 (39.50)	19.30 (28.94)
S6: It is a good thing that people speak English in public places	72.11 (20.07)	70.71 (37.71)	71.82 (35.36)	82.46 (32.85)	81.34 (31.84)	37.37 (33.15)
S7: It would be nice to learn about American values and culture.	60.00 (29.39)	48.04 (39.41)	68.53 (20.24)	61.10 (37.72)	62.91 (35.42)	58.05 (31.01)

From Figures 6.2 and 6.3, it can be seen that the seven statements fall into three categories as shown in Table 6.4. It is obvious that all participants tended to hold positive attitudes towards English proficiency and learning: there was little or no variation in their ratings of the first four statements. The influence of gender on attitudes is clear in Statements 5 and 6 about using English in greetings and public places, but this influence seems to interact with the language exposure group. In the medium exposure group, the male participants showed more positive attitudes than the female participants; however, in the low exposure group, the female participants showed more positive attitudes than the male participants. Overall, participants showed more variation in Statement 7 about American culture and values, but there are no clear gender differences within each group.

[18] Values in parentheses are standard deviations

Table 6.4: The classification of the seven statements into three categories

English proficiency and learning	S1: Well-spoken English is beautiful to listen to. S2: I like a person who speaks English fluently. S3: It is good to teach English to all pupils from the first grade of elementary school. S4: People should make an effort to improve their English.
English in greetings and public places	S5: It is ok for people to greet others in English. S6: It is a good thing that people speak English in public places
American culture and values	S7: It would be nice to learn about American values and culture.

6.6.2 Linear Regression Analyses

Linear regression was performed on the participants' ratings (raw scores) of the seven statements. The first model was run on the first four statements representing attitudes towards English learning and proficiency, the second model was run on Statements 5 and 6 representing attitudes towards using English in public places and greetings, and the third model was run on Statement 7 representing attitudes towards American culture and values.

The results of the mixed effects model[19] constructed for the first four statements are presented in Table 6.5 and visualised in Figure 6.4. The estimate for intercept is the estimate for the low exposure group and males. The model includes group and gender as fixed effects and random intercepts for statement and speaker. The interaction between group and gender was not included in the model because it was not significant and did not improve the model fit. The model showed no significant differences. The female and male participants in all three groups tended to hold similar attitudes, and this can be clearly seen in Figure 6.4.

[19] M1<- lmer(response~ gender + group + (1|speaker) + (1|statement), data = attitude_a)

Table 6.5: Linear mixed effects model for statements 1-4

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	91.323	3.541	55.845	25.789	<2e-16 ***
High group	-1.572	4.019	63.000	-0.391	0.697
Medium group	-1.640	4.097	63.000	-0.400	0.690
Females	4.393	3.296	63.000	1.333	0.187

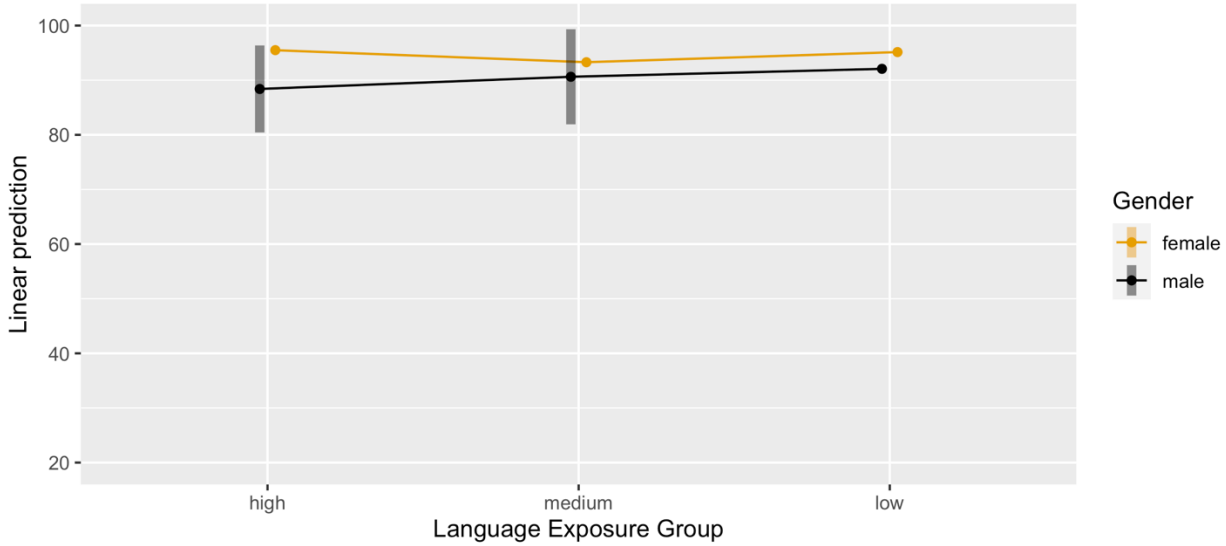


Figure 6.4: Estimated means for participants' ratings of statements 1-4 by level of English exposure and gender

Table 6.6 and Figure 6.5 present the results of the mixed effects model[20] constructed for Statements 5 and 6. It is worth noting that conducting simple linear regression models on Statements 5 and 6 separately showed similar results which confirm the classification of the statements in Table 6.4. The estimate for intercept for the mixed effects model is the estimate for the low group and males. The model involves a two-way interaction between group and gender as a fixed effect and random intercepts for statement and speaker. The results showed that language exposure had a main effect; the coefficients associated with high exposure group (β = 28.125, SE = 13.437, df = 61.000, t = 3.079, p = 0.00 **) and medium exposure group (β =

[20] M2<- lmer (response~ gender*group + (1|speaker) + (1|statement), data = attitude_b)

46.717, SE =14.001, df = 61.000, t = 2.093, p = 0.04*) are each significantly different from those associated with the low exposure group. However, this effect is likely due to the male participants; the differences were greater between the male participants than the female participants, as shown in Figure 6.5. The model also reveals a significant difference between the male and female participants (β = 41.375, SE = 13.437, df = 61.000, t = 3.337, p = 0.00**), and suggests a significant interaction between gender and group. As shown in Figure 6.5, the male participants in the medium exposure group tended to have more positive attitudes than the male participants in the low and high exposure groups and even more positive attitudes than the female participants in the same group. In the low exposure group, the female participants tended to have more positive attitudes than the female participants in the other two exposure groups and more positive attitudes than the male participants in the same group.

Table 6.6: Linear mixed effects model for statements 5-6

Fixed effects	Estimate	SE	Df	t value	Pr(>|t|)
Intercept	30.833	15.191	2.993	2.030	0.13561
High group	28.125	13.437	61.000	3.079	0.00311 **
Medium group	46.717	14.001	61.000	2.093	0.04051 *
Females	41.375	13.437	61.000	3.337	0.00145 **
High group: Females	-43.375	18.312	61.000	-2.369	0.02103 *
Medium group: Females	-62.717	18.730	61.000	-3.349	0.00140 **

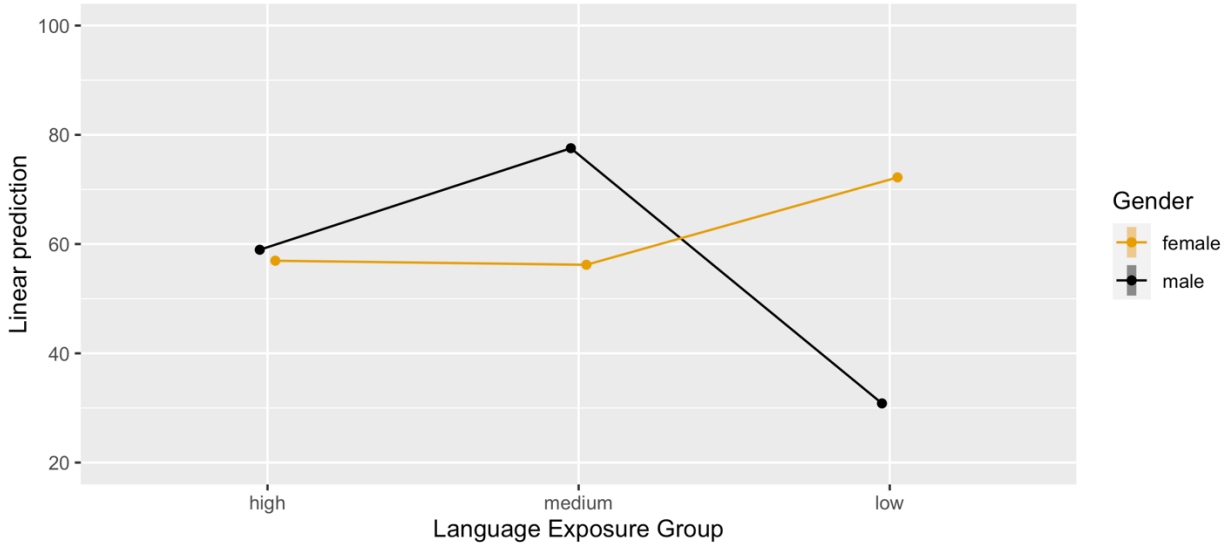


Figure 6.5: *Estimated means for participants' ratings of statements 5 and 6 by level of English exposure and gender*

Table 6.7 and Figure 6.6 present the results of the model constructed for Statement 7 representing attitudes towards American culture and values. The simple linear regression was performed on Statement 7 because there is only one score for each participant. The estimate for intercept is the estimate for the low group and males. The model includes a two-way interaction between group and gender as a fixed effect. The model showed a significant effect of language exposure group: the high exposure group significantly differ from the low exposure group (β = -10.008, SE =2.597, t =-3.854, p = 0.00***). As shown in Figure 6.6, the participants in the high exposure group showed less positive attitudes than the participants in the other two exposure groups. The model also reveals a significant interaction between group and gender. The difference between the female and male participants is greater in the high exposure group than in the medium and low exposure groups.

Table 6.7: Simple linear model for statement 7

Fixed effects	Estimate	SE	t value	Pr(>|t|)
Intercept	58.052	1.989	29.192	< 2e-16 ***
High group	-10.008	2.597	-3.854	0.00012 ***
Medium group	3.055	2.739	1.115	0.26476
Females	4.864	2.575	1.889	0.05902 .
High group: Females	7.093	3.443	2.060	0.03952 *
Medium group: Females	2.562	3.579	0.716	0.47410

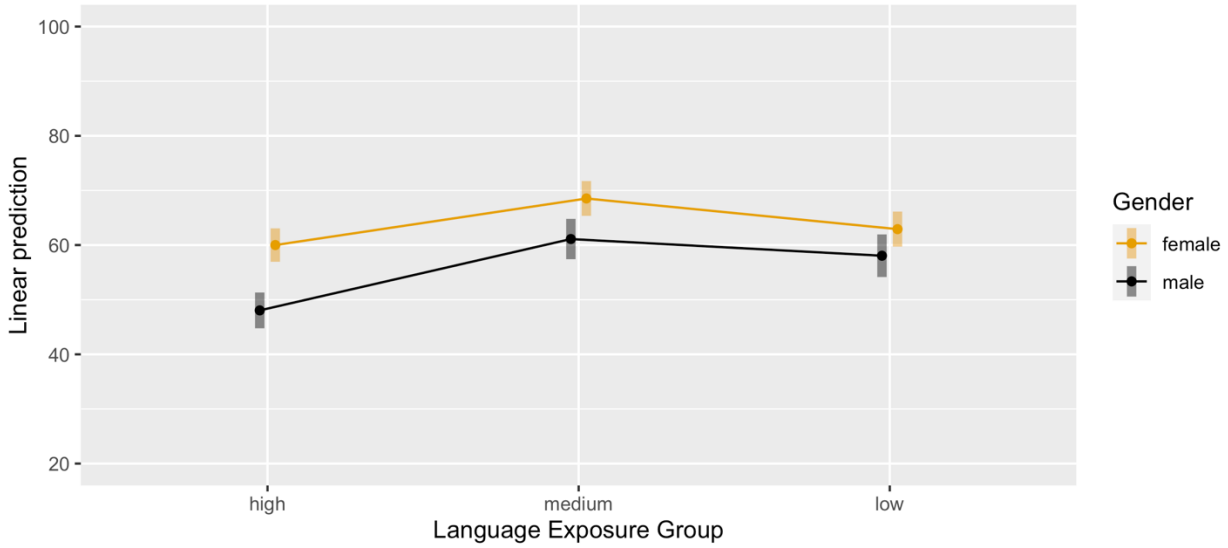


Figure 6.6: Estimated means for participants' ratings of statement 7 by level of English exposure and gender

6.7 Discussion

This section aims to discuss the survey results in relation to the results of the preceding production task. The aim of the present survey was to examine participants' attitudes towards the English language and American culture to see if any differences emerged between the three exposure groups and genders. The results showed that participants had overall positive attitudes towards English learning and proficiency (reflected in Statements 1-4), which is perhaps unsurprising given that English enjoys a high status in Saudi Arabia. English is looked upon as a passport to a better education and more job opportunities. However, interestingly, while almost

all of the participants recognized the importance of English learning and proficiency, they differed in their attitudes towards using English in greetings and public places (reflected in Statements 5-6) and learning about American culture and values (reflected in Statement 7). It is worth noting that participants were asked about American culture in Statement 7 because Saudis are more exposed to American English in the media. Thus, it is unknown if they hold different attitudes towards the culture of other English-speaking groups. The variability in participants' rating for the statements about using English in greetings and public places may suggest that participants may have had an ideology regarding language use; Arabic may be thought of as more 'appropriate' in certain places or situations.

Although previous literature demonstrates that attitudes affect the production of novel sounds in loanwords; namely, positive attitudes can result in a higher importation rate (e.g., Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2011), the survey results did not show this effect. Recall the results of the preceding production tasks. The female participants had a higher importation rate of /v/ and /tʃ/ in the real words than the male participants. In addition to gender, language exposure had a main effect on the importation rate of both /v/ and/ tʃ/; the higher levels of exposure to English resulted in a higher importation rate of the two target segments. Based on the previous literature, it was expected that females and participants in the high exposure group would have more positive attitudes than males and participants in the medium and low exposure groups.

However, in the first four statements, the results showed no significant differences based on level of English exposure or gender. The instrumental value of English was high for both female and male participants in the three exposure groups. In Statements 5 and 6, unlike the previous statements, the results showed significant differences between participants based on their level of English exposure and gender. Interestingly, the female participants in the low exposure group showed more positive attitudes than the male participants; in contrast, the male participants in the medium exposure group showed more positive attitudes than the female participants. The difference between the female and male participants in the high exposure group was small. In response to Statement 7, gender had no main effect and the participants in the high exposure group showed less positive attitudes than the participants in the low exposure group.

When examining the data for the female and male participants separately in both Figures 6.5 and 6.6, one can see that group differences among the male participants across the three

exposure groups were greater than among the female participants. As shown in Figure 6.5, the male participants in the medium exposure group showed more positive attitudes towards using English in greetings and public places than the participants in the other two exposure groups. As mentioned previously, the participants in the medium exposure group are students in the department of Computer Science. It is possible that males in the medium exposure group had more positive attitudes than males in the other groups because of the need of English for employment; private sector companies and factories, where men are more likely to work than women, may not only require good computer skills, but also proficiency in English. Whatever the reasons, the fact is that the participants who are shown by the survey to hold more positive attitudes were not more likely to produce [v] and [tʃ]. It is surprising that the female and male participants in the high exposure group did not show more positive attitudes than the participants in the other two exposure groups towards the English language and American culture. That is, having a higher level of English exposure does not necessarily promote more positive attitudes.

According to these results, it can be inferred that attitudes do not provide a satisfactory explanation for the effect of gender and language exposure on the importation rate of the target segments. The participants with the high level of English exposure were more likely to produce the imported segments [v] and [tʃ] because they simply had more opportunities to hear and read the imported segments. It still remains a puzzle, however, why the female participants had a higher importation rate than the male participants. We shall come back to this point in the general discussion in Chapter 7 where the three studies of this thesis are tied together.

6.8 Summary

This section briefly summaries this chapter. The aim of this chapter was to explore whether differences in attitudes towards the English and American culture could be part of the explanation for the significant impact of gender and language exposure on the importation rate of target /v/ and/ tʃ/. The findings of the production tasks showed that females and participants with the high level of exposure to English were more likely to produce the imported segments [v] and [tʃ] than males and participants in the other two exposure groups. A short questionnaire consisting of seven statements was adapted from Te Puni (2010) to explore participants' attitudes towards the English language and American culture. All the participants in the perception and production experiments also completed the present attitudes survey. Qualtrics

(www.qualtrics.com) was used to distribute and collect data online due to the strict COVID-19 ban on in-person data collection at the time.

The findings showed that participants tended to hold similar attitudes towards English learning and proficiency. The was a significant interaction between language exposure group and gender in attitudes towards using English in greetings and public places. The female participants had more positive attitudes than the male participants in the low exposure group; however, the male participants had more positive attitudes than the female participants in the medium exposure group. The most surprising aspect of data is that the participants in the high exposure group, who had the highest rate of importation, showed less positive attitudes towards American culture than the participants in the medium and low exposure groups. In general, these results suggest that attitudes cannot be part of the explanation for the significant effect of gender and language exposure on the importation rate of target /v/ and/ ʧ/. That is to say, the positive attitudes did not increase the likelihood of producing [v] and [ʧ]. These survey results, together with the results of the perception and production experiments, will be discussed thoroughly in the next chapter.

7 General Discussion

This chapter aims to pull together all four studies in this thesis (Section 7.1), present k-means clustering analysis across all perception and production tasks (Section 7.2), discuss the key findings (Section 7.3), conclude the thesis (Section 7.4) and outline limitations, contributions, and suggestions for future research (Section 7.5).

7.1 Summary of Key Findings

The aim of this thesis is to examine factors governing the variable adaptation of novel structures in loanwords generally. The interest of this thesis lies especially in the variable adaptation of two novel sounds in English loanwords into Arabic: namely /v/ and /tʃ/. As explained in Chapters 1 and 2, previous literature showed that target /v/ and /tʃ/ are sometimes retained [v] and [tʃ] and sometimes replaced by [f] and [ʃ], respectively (e.g., Abu Guba, 2016; Aloufi, 2016). As a reminder, a production of [v] or [tʃ] is treated as a case of being imported from English whereas a production of [f] or [ʃ] is a case of being adapted to Arabic phonology.

In pursuit of our aim, participants performed three tasks (i.e., one perception task and two production tasks) and completed a short survey on attitudes towards the English language and American culture. As pointed out in Chapter 1, in data collection, the two production tasks preceded the perception task to prevent participants from being familiarized with auditory inputs, since the same non-words were used in the perception task as in the production task for non-words. However, the perception task was presented first in this thesis because its results were first analysed and used as predictions for analysis of the non-word production task. Recall that the participants in the four studies were 67 Saudi speakers of Arabic, split by gender (31 males and 36 females) and by level of English exposure (high, medium and low). In what follows, we summarise the four studies and present their key findings one by one.

In Chapter 3, we examined the discrimination accuracy of the /v-f/ and /tʃ-ʃ/ contrasts. The concrete question addressed in this chapter was: *do participants find it difficult to perceive target /v/ and /tʃ/ as different from their L1 counterparts /f/ and /ʃ/?* Participants performed an oddity task in which they were asked to decide which stimulus is different, with the option to say that the three auditory stimuli are the same if they do not hear an odd stimulus. The stimuli were 24 CVCVC non-words with target segments embedded in three positions: initial, intervocalic and final. The vowels were held constant with /æ/ in the first syllable and /ɪ/ in the second

syllable. Target non-words were represented in six different trials (AAB, ABA, ABB, BAA, BBA, BAB). The key findings of this task are:

a. Overall, the discrimination accuracy of the /v-f/ contrast was poor while the discrimination accuracy of the /tʃ-ʃ/ contrast was good.
b. Level of English exposure had a main effect on the discrimination accuracy of the two contrasts: the participants in the high exposure group were more likely to discriminate the two contrasts than the participants in the medium and low exposure groups.
c. Word position had a main effect on the discrimination accuracy of the /v-f/ contrast but not the /tʃ-ʃ/ contrast: the /v-f/ contrast was less likely to be discriminated in word-final position than other in other positions.
d. Gender did not affect the discrimination accuracy of the two contrasts: there were no significant differences in discrimination performance between the female and male participants.

In Chapter 4, we considered the likelihood of importation vs. adaptation in non-words. The main question addressed in this chapter was *to what extent do input modality, level of English exposure, word position and gender account for variability in the production of target /v/ and /tʃ/ in non-words?* Participants' productions of target non-words were elicited in three conditions: aural-only (listen-say), written-only (read-say), and aural-written (listen-read-say). The same 24 CVCVC non-words used in the perception task were split across the three conditions. In each condition, participants were requested to produce six non-words twice. The key findings of this task are:

a. Overall, input modality had a main effect on the importation rate of /v/ and /tʃ/: auditory inputs facilitated the production of [tʃ] but hindered the production of [v]. In contrast, written inputs facilitated the production of [v], but hindered the production of [tʃ].
b. Level of language exposure had a main effect on both /v/ and /tʃ/ importation rates. The participants in the high and medium exposure groups were more likely to produce [v] than the participants in the low exposure group. The participants in the high exposure group were more likely to produce [tʃ] than the participants in the low and medium exposure groups.
c. Gender had a main effect on the importation rate of /v/ but not /tʃ/: females were more likely to produce [v] than males. There was a significant interaction of gender with input

modality for /v/ production, whereby females were more likely to produce [v] with the availability of aural inputs than males.

d. Word position had a main effect on the importation rate of /v/ but not /tʃ/: participants were more likely to produce /v/ as [f] in word-final position across the board.

In Chapter 5, we considered the likelihood of importation vs. adaptation in real words. The main question addressed in this chapter was *to what extent do level of English exposure, word position and gender account for variability in the production of target /v/ and /tʃ/ in real words?* The production of /v/ and /tʃ/ in real words was elicited using a fill-in-the-blank task in which the participants were asked to read sentences and then fill in the blanks with 29 loanwords. The key findings of this task are:

a. Level of language exposure had a main effect on the importation rate of /tʃ/ but not /v/: the participants in the high and medium exposure groups were more likely to produce [tʃ] than the participants in the low exposure group.
b. There was a significant interaction of group with word position in /v/ production, whereby the participants in the medium exposure group were more likely to produce [f] in word-final position than in other positions.
c. Gender had a main effect on the importation rate of /v/ and /tʃ/: females were more likely to produce [v] and [tʃ] than males.

In Chapter 6, we analysed a short survey consisting of seven statements. The original aim of this survey was to examine participants' attitudes towards the English language and American culture. The analysis in this chapter aimed also to explore attitudes as possible explanation of gender effects in Chapters 4 and 5; namely, that female participants may possibly hold more positive attitudes towards the English language and/or American culture than the male participants. The key findings of this survey are:

a) Participants of both genders tended to hold similar attitudes towards English learning and proficiency.
b) The male participants in the medium exposure group showed more positive attitudes towards using English in greetings and public places than the female participants in the same group.
c) The female participants in the low exposure group showed more positive attitudes towards using English in greetings and public places than the male participants in the same group.

d) Overall, the female participants did not show more positive attitudes towards the American culture than the male participants.
e) The female and male participants in the high exposure group, who had the highest rate of importation in production, in fact displayed less positive attitudes towards the American culture than the participants in the medium and low exposure groups.

Table 7.1 summarises the significant effects found across the four main tasks that we shall discuss in Section 7.3. Recall that in the statistical model the reference levels are the following: low exposure group, word-final position, males, and aural-only condition. As a reminder, the first four statements in the attitudes survey represent attitudes towards English learning and proficiency, statements 5 and 6 represent attitudes towards using English in public places and greetings, and statement 7 represent attitudes towards American culture and values. In Table, 7.1, SG denotes a statistically significant effect ($p < 0.05$) for the relevant variable and task. SGI indicates that a variable was involved in a statistically significant interaction with another variable ($p < 0.05$). The levels of the categorical variables that are significantly different from the reference level in each case are indicated between parentheses.

Table 7.1: Summary of findings in the perception and production tasks and attitudes survey (SG: statistically significant variable ($p < 0.05$); SGI: statistically significant interaction ($p < 0.05$))

Variable	Ch.3 Perception Task		Ch.4 Production Task (Non-words)		Ch.5 Production Task (Real Words)		Ch.6 Attitudes Survey		
	/v/	/tʃ/	/v/	/tʃ/	/v/	/tʃ/	Ss 1-4	Ss 5-6	S 7
Level of English Exposure	**SG** (high)	**SG** (high)	**SG** (high & medium)	**SG** (high)		**SG** (high & medium)		**SG** (high & medium) **SGI** (high*females) (medium*females)	**SG** (high) **SGI** (high* females)
Word Position	**SG** (initial & intervocalic)		**SG** (initial & intervocalic)		**SGI** (initial* medium group) (medium* medium group)				
Gender			**SG** (females)		**SG** (females)	**SG** (females)		**SG** (females)	
Input Modality			**SG** (written) **SGI** (written* females)	**SG** (written & aural-written)					

7.2 K-means clustering

Before discussing the above findings, in this section we explore grouping of the participants based on their performance on the perception and production tasks using K-means clustering. Under Covid conditions, we were not able to reliably assess participants' proficiency in English because the assessments took place in unsupervised settings. For example, a participant might ask someone to take the test for them or search for answers online.

Therefore, as explained in Chapters 1 and 3, we used medium of instruction as a proxy and recruited the participants in three groups (high, medium and low) based on their expected level of English exposure (i.e. based on their academic program). The purpose of conducting k-

means clustering in this section is to check whether participants form clusters that closely match the three exposure groups.

K-means is a method of clustering analysis that is used to group similar data points into clusters. The *k* denotes the number of clusters in the data. The recommended number of clusters for k-means clustering can be determined using a variety of techniques, such as the elbow, silhouette, and gap statistics. The clustering of data points is done in a way that minimizes sum of distances between data points within one cluster (Zubair et al., 2022). The analysis was conducted in the results for each participant along three parameters – that is, their performance in the perception task, in the production task for real words, and in the production task for non-words. First, the number of target sounds discriminated correctly by each participant in the perception task, and the number of imported sounds [v] and [tʃ] produced by each participant in the first production task (non-words) and in the second production task (real words) were calculated. As shown in the previous section, level of English exposure, as operationalized in terms of groups, plays an important role. The participants in the high exposure group were more likely to discriminate the target contrasts and produce the imported sounds [v] and [tʃ], followed by the participants in the medium and low exposure groups.

The k-means analysis was performed separately for target /v/ and /tʃ/. The results support the adopted method of recruiting participants as a good proxy for level of English exposure. The results showed that participants form two clusters based on their performance across the three tasks. Participants who are assigned to the same cluster are similar to each other in their performance in the three tasks. High-proficiency participants and low proficiency participants were assigned to different clusters. Participants in the high exposure group who tended to perceive and produce target /v/ and /tʃ/ were assigned to the same cluster. Future studies, however, might measure participants' levels of English exposure using a continuous scale.

For target /v/, most participants in the high exposure group belong to cluster 1 while most participants in the low exposure group belong to cluster 2. The participants in the medium exposure group were divided between the two clusters. For target /tʃ/, most participants in the high exposure group belong to cluster 2; however, the participants in the low and medium exposure groups were divided between the two clusters.

The number of participants in each cluster for target /v/ and /tʃ/, split by their proxy available exposure group (high, medium and low), are visualized in Figures 7.1 and 7.2. The blue

cluster in the two figures includes mostly high-proficiency participants (i.e., those who tended to perceive and produce the target sounds) while the red cluster includes mostly low-proficiency participants. Participants who are in the blue cluster for target /v/ are also in the blue cluster for target /tʃ/ (except two female participants in the medium group and two male participants in the high group). Tables 7.2 and 7.3 present participants in cluster 1 and cluster 2 for target /v/ and /tʃ/.

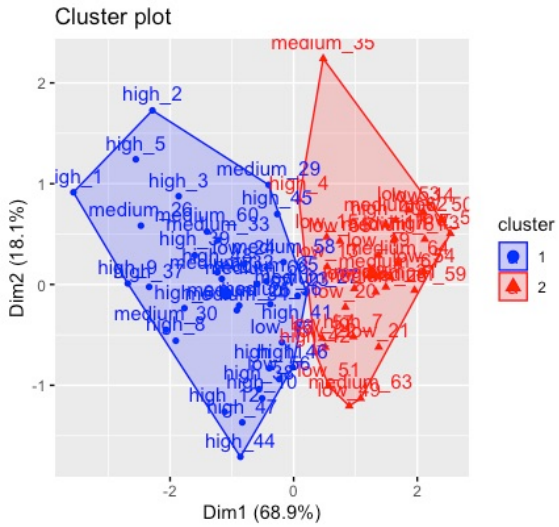


Figure 7.1: Cluster plot of participants' performance in the perception and production tasks for target /v/

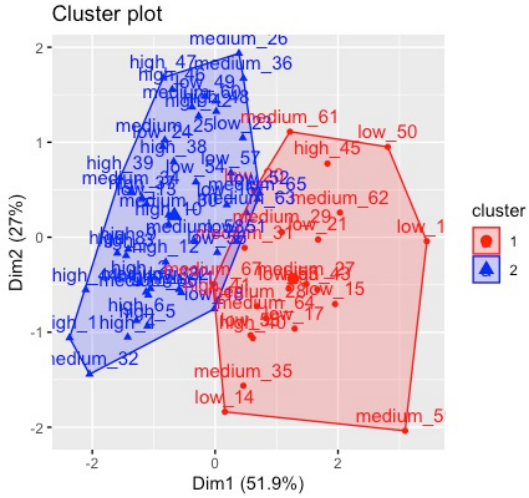


Figure 7.2: Cluster plot of participants' performance in the perception and production tasks for target /tʃ/

Table 7.2: K-means clustering results of participants' performance in the perception and production tasks for target /v/

	high group	medium group	low group
cluster 1	1. FH01 2. FH02 3. FH03 4. FH05 5. FH06 6. FH08 7. FH09 8. FH10 9. FH11 10. FH12 11. MH01 12. MH02 13. MH03 14. MH05 15. MH08 16. MH09 17. MH10 18. MH11 19. MH12	1. FM01 2. FM02 3. FM03 4. FM06 5. FM07 6. FM09 7. FM10 8. FM11 9. FM13 10. MM01 11. MM03 12. MM09 13. MM11	1. FL04 2. FL11 3. FL12 4. ML11
cluster 2	1. FH04 2. FH07 3. MH04 4. MH06 5. MH07	1. FM04 2. FM08 3. FM12 4. MM02 5. MM04 6. MM05 7. MM06 8. MM08 9. MM12	1. FL01 2. FL02 3. FL03 4. FL05 5. FL06 6. FL07 7. FL08 8. FL09 9. FL10 10. ML01 11. ML02 12. ML03 13. ML04 14. ML05 15. ML07 16. ML10 17. ML12

Table 7.3: K-means clustering results of participants' performance in the perception and production tasks for target /tʃ/

	high group	medium group	low group
cluster 1	1. MH04 2. MH05 3. MH07	1. FM03 2. FM04 3. FM06 4. FM08 5. FM12 6. MM02 7. MM04 8. MM05 9. MM08 10. MM12	1. FL03 2. FL05 3. FL07 4. FL08 5. FL09 6. ML02 7. ML05 8. ML10 9. FL02
cluster 2	1. FH01 2. FH02 3. FH03 4. FH04 5. FH05 6. FH06 7. FH07 8. FH08 9. FH09 10. FH10 11. FH11 12. FH12 13. MH01 14. MH02 15. MH03 16. MH06 17. MH08 18. MH10 19. MH11 20. MH12	1. FM01 2. FM02 3. FM07 4. FM09 5. FM10 6. FM11 7. FM13 8. MM01 9. MM03 10. MM06 11. MM09 12. MM11	1. FL01 2. FL04 3. FL06 4. FL10 5. FL11 6. FL12 7. ML01 8. ML03 9. ML04 10. ML07 11. ML11 12. ML12

The participants in the high and low exposure groups fall in clusters that closely match their exposure groups for /v/. This match is somewhat less for /tʃ/ than /v/, and the differences between participants were greater for /v/ than /tʃ/. A possible reason for this difference is that /tʃ/ exists in other Arabic dialects, so is easier to perceive and produce. The cluster analysis displayed positive but small silhouette values, however, so caution must be applied. The

silhouette value is a measure of how well samples are clustered. The silhouette values range from -1 to 1 (i.e., the best value is 1 while the worst value is -1). Large positive values indicate that clusters are separated perfectly while negative values indicate overlapping clusters (Janert, 2010). The average silhouette width was 0.40 for the /v/ dataset and 0.33 for the /tʃ/ dataset (see Appendix E). In the discussion that follows, we will however continue to use the three exposure groups since we did not have an alternative means to assess participants' English proficiency.

7.3 What Matters: Factors Affecting the Importation Rate of Target /v/ and /tʃ/

In this thesis, we explored a range of factors affecting the likelihood of selecting the imported sounds [v] and [tʃ] vs. the adapted sounds [f] and [ʃ]. In this section, we discuss the key findings of this study in relation to our predictions and previous literature.

7.3.1 Effect of Input Modality

The first factor we considered in this study is input modality. The results revealed a significant effect of input modality, and the direction of this effect is as predicted. Presence of auditory input was predicted to increase the importation rate of /tʃ/ but not /v/. Target /tʃ/ exists in some Arabic dialects and hence it involves salient acoustic features for speakers of Saudi Arabic. As predicted, auditory inputs facilitated the production of [tʃ], but hindered the production of [v]. In contrast, written inputs facilitated the production of [v], but hindered the production of [tʃ]. In written-only condition, participants were more likely to control their articulation and produce acoustic features consistent with target /v/. However, participants' production of /tʃ/ in written-only condition may have been affected by their prior experience of letter-to-sound correspondences in Arabic which are transparent. Written inputs might be less helpful for Arabic speakers in supporting the production of [tʃ], because English spelling is inconsistent in mapping between <ch> and /tʃ/. In English, /tʃ/, /ʃ/ and even /k/ can be represented by the same diagraph <ch>. In this way, these findings confirm previous studies suggesting a negative impact of L2 orthography on speakers whose L1 has a transparent writing system as the case for Arabic (e.g., Bassetti, 2017), and further suggest that even if the L1 and L2 use different scripts (Arabic script vs. Roman script), a sound might be mispronounced if the L2 spelling is inconsistent in mapping between sounds and graphemes.

As discussed in the literature review in Chapter 2, perception is known to play a key role in the likelihood of adaptation vs. importation of novel structures. The overall discrimination accuracy of the /v-f/ and /tʃ-ʃ/ contrasts may explain why the effect of input type varies for the

two sounds. Consistent with PAM predictions (Best, 1995; Best & Tyler, 2007) (see 3.5), the discrimination accuracy of the /tʃ-ʃ/ contrast (category goodness assimilation - CG) was higher than that of the /v-f/ contrast (Single category - SG). Native language experience might be one source of this discrepancy; Saudi Arabic speakers have more potential exposure to the acoustic cues of /tʃ/ than of /v/, because /tʃ/ exists in certain Arabic dialects spoken in Gulf countries (e.g, Kuwait and Bahrain), while /v/ does not. The perception accuracy of /v/ was poor; hence the availability of auditory input did not support the production of [v]. In contrast, the perception accuracy of /tʃ/ was good; hence the availability of auditory input did support the production of [tʃ].

Given the effect of auditory inputs, I speculated that there is a link between participants' perception and production. To verify this speculation, we examined the relationship between participants' perception and non-word production using a series of Kendall's tau non-parametric tests. Kendall's tau non-parametric test is used to measure the strength of association between two variables when data is not normally distributed. Recall that we used the same non-words in the perception and production tasks. We did not include participants' production of real words in this analysis. The reason for this is that a speaker's production of a real word may depend on its stored representation (how the word should sound based on their experience of its use in Arabic context). For example, they may tend to produce the initial sound in 'virus' as [f] even though they are able to produce [v].

Prior to analysis, we established a by-participant score for each task. For the perception task, we calculated the number of target sounds discriminated correctly by the participants. For the non-word production task, we calculated the number of imported sounds produced by the participants across the three conditions (aural-only/ written-only/ aural-written). Overall, we found some evidence of a positive correlation between participants' perception and production: it seems that the participants who produced more [v] and [tʃ] also discriminated the /v-f/ and /tʃ-ʃ/ contrasts better. Kendall's tau was 0.4424188 ($z = 5.1179$, $p = 3.09e-07$) for /v/, and 0.2122293 ($z = 2.4082$, $p = 0.01603$) for /tʃ/. However, it is worth noting that these correlations are not large. According to Field et al. (2012), coefficients greater than 0.5 represent a large correlation, coefficients between 0.3 and 0.5 represent a medium correlation, and coefficients less than 0.3 represent a small correlation. The correlation between production and perception can thus be characterised as medium for /v/ and small for /tʃ/. As shown in Figures 7.3 and 7.4, the data

points for /tʃ / are more spread out than the data points for /v/, namely that the correlation between perception and production for target /v/ is stronger than the correlation between perception and production for target /tʃ/. The small but significant correlation for /tʃ/ could indeed be driven by some individual participants, since it seems that most participants showed no link between their perception and production.

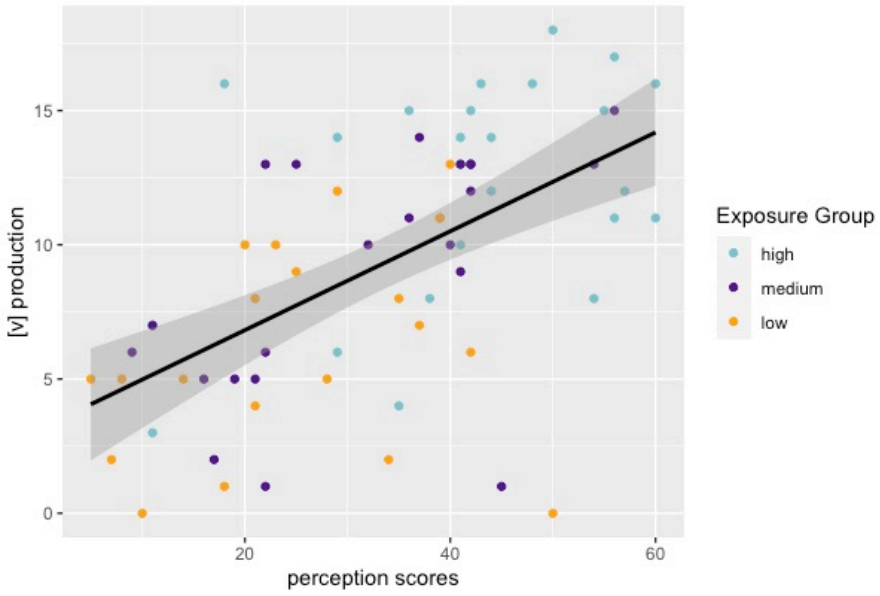


Figure 7.3: The relationship between participants' perception and production of target /v/. The x-axis represents participants' perception scores, and the y-axis represents their use of the imported sound [v] across the three conditions.

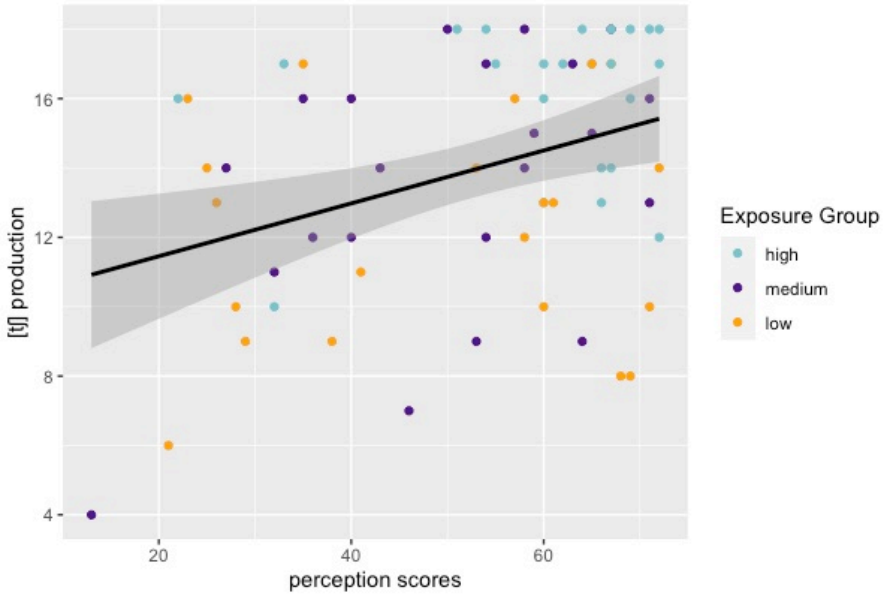


Figure 7.4: The relationship between participants' perception and production of target /tʃ/. The x-axis represents participants' perception scores, and the y-axis represents their use of the imported sound [tʃ] across the three conditions.

Overall, the correlational analysis showed positive correlations between participants' perception and production of target non-words. Additionally, the k-means analysis in 7.2 showed that participants can be divided into two non-overlapping clusters based on their performance in the perception and production tasks. These results suggest a close relationship between perception and production. That is, participants who used imported sounds [tʃ] and [v] more in production were also better at perceiving target contrasts /tʃ-ʃ/ and /v-f/.

However, participants' perception accuracy might not correlate consistently with their non-word production in the three input conditions. Therefore, we examined whether participants' perception accuracy correlated with their production within each input condition in a further series of kendall's tau non-parametric test on relevant subsets of the data. As shown in Table 7.4, participants' perception is mirrored in their production in aural-only condition for target /v/ but not /tʃ/. In aural-only condition, auditory inputs were not helpful for target /v/ because it was not easy for participants to attend to the acoustic differences between /v/ and /f/.

Table 7.4: The relationship between participants' perception accuracy and their production in the three conditions

Segment	condition	correlation (tau)	z-value	p-value
/v/	Aural condition	0.4333552	4.8335	(p <0.001)
	Aural-written condition	0.3999067	4.4604	(p <0.001)
	Written Condition	0.2799588	3.1219	0.001797
/tʃ/	Aural condition	0.1026628	1.0764	0.2817
	Aural-written condition	0.2031645	2.1056	0.03524
	Written Condition	0.2025004	2.2404	0.02506

Overall, this explanation of the relationship between perception and production provides only partial support for the claim that perception is the basis for the production of non-native structures (e.g, Boersma & Hamann, 2009; Dupoux et al., 1999; Peperkamp et al. 2008**)**. However, these findings support the argument made here that the effect of input modality can vary for different target sounds. Our argument is bolstered by the fact that the correlation between participants' perception accuracy of the /v-f/ contrast and their use of [v] in aural condition is significant, and earlier descriptive and inferential results showing a main effect of

input modality such that the auditory inputs resulted in an increase in the importation rate of /tʃ/ but not /v/.

7.3.2 Effect of Language Exposure

The second factor we considered in this study is level of exposure to the source language. For non-words, the effect of the level of exposure to English was as expected. The participants expected to have high levels of English exposure were more likely to produce [v] and [tʃ] than those with lower levels of English exposure. A possible reason for this is that the participants with high levels of English exposure were likely to have read, heard and produced more English words than those with low levels of English exposure. In this way, our results are in line with the findings of previous studies on loanword phonology (e.g., Kang & Schertz, 2021; Kwon, 2017) showing the significant effect of language exposure on importation rates of novel sounds. According to the Lexical Entrenchment Hypothesis (Diependaele et al., 2013), variable exposure to the L2 has important consequences for the processing of L2 words. A high amount of exposure to the L2 can reduce frequency effects; speakers with high levels of L2 exposure process low-frequency L2 words more quickly and accurately than speakers with low levels of L2 exposure. That is, differences in processing low-frequency L2 words are due to differences in exposure to the L2. In the present study, participants in the three exposure groups have never encountered the target non-words before. However, these non-words contain linguistic elements (sound/graphemes) that may be familiar to speakers with high levels of English exposure but may be quite unfamiliar to speakers with low levels of English exposure.

The importation rate of the two target sounds was low in real words compared to non-words. There was a main effect of level of English exposure on the importation rate of /tʃ/ but not of /v/ in real words. The participants in the high and medium exposure groups were more likely to produce the imported sound [tʃ] than the participants in the low exposure group. However, we speculate that this null result for target /v/ might be due to specific properties of the set of real words employed in the production task. Due to the rise of Internet-mediated communication, borrowing from English into Arabic has increased. New words borrowed into Arabic are often in the areas of science, technology, economy, politics and food (Alhussami, 2020). The list of real words with /tʃ/ was small (8 words) and contained several recently borrowed words (e.g, snapchat, cheesecake, and chips). However, the list of real words with /v/ (21 words) contained older-borrowed words that have found their way into the media many years

ago and were included in Arabic lexicons (e.g., contemporary Arabic lexicon; Omar, 2008) and dictionaries for Arabic learners (e.g., A frequency dictionary of Arabic: core vocabulary for learners, Buckwalter & Parkinson, 2014).

With frequent use of a loanword, a novel sound is less likely to persist (Poplack, 2018). According to Daland et al. (2015), in commonly used borrowed words, the adaptation process becomes stabilised, and individuals may converge on one or other form of production. Dohlus (2010) suggests:

> "monolingual borrowers may use an adaptation form based on misperception, but bilingual speakers may perform adaptations yielding phonological similarity. One of these adaptation forms will sooner or later prevail due to standardizations that occur when words are used repeatedly in the media or public, are recorded in dictionaries and are adjusted to a different writing system" (p.146).

Speakers with different levels of English exposure may therefore tend to produce the adapted sound [f] in some words for target /v/ because these words are much more similar to native words than loanwords, especially that English acoustic and orthographic cues for the target sounds are absent in the task.

It is unknown when the selected English borrowed words came into Arabic; however, Table 7.5 shows real words for target /v/ and /tʃ/ that are included in a contemporary Arabic lexicon (Omar, 2008) and/or a dictionary for Arabic learners (Buckwalter & Parkinson, 2014). The letter A in circle brackets refers to words included in the contemporary Arabic lexicon (Omar, 2008) while the letter B refers to words included in the dictionary of Arabic for learners (Buckwalter & Parkinson, 2014). Table 7.6 shows Arabic pronunciation and orthographic representations[21] for the subset of target words included in the contemporary Arabic lexicon and the dictionary of Arabic for learners.

[21] The standardised words have the same orthographic representations in the Arabic lexicon and dictionary for Arabic learners

Table 7.5: Target real words included in the contemporary Arabic lexicon (Omar, 2008) and the dictionary for Arabic learners (Buckwalter & Parkinson, 2014). The letter A refers to words included in the lexicon while the letter B refers to words included in the dictionary

	Target words included in the lexicon and dictionary	Target words not included in the lexicon and dictionary
/v/	video (A/B), vanilla (A), vaseline (A), vitamins (A/B), virus (A/B), vase (A), veto (A)	van, receiver, cover, avocado, boulevard, lavender, red velvet, seven up, caravan, mauve, glove, microwave, live
/tʃ/	chocolate (A), ketchup (A), chimpanzee (A)	Cheesecake, chips, cappuccino, snapchat, clutch

Table 7.6: Arabic pronunciation and orthographic representations of standarised target loanwords

English word		IPA transcription	Romanised transliteration[22]	Arabic orthographic Representation
/v/	video	[fidju:]	<fidyu:>	<فديو>
	vanilla	[fanilla]	<fa:nylla:>	<فانيلا>
	vaseline	[fazli:n]	<fa:zli:n>	<فازلين>
	vitamins	[fitami:n]	<fyta:mi:n>	<فيتامين>
	virus	[firu:s]	<fa:yru:s>	<فايروس>
	vase	[faːz ɑh]	<fa:zah>	<فازة>
	veto	[fi:tu:]	<fytu:>	<فيتو>
/tʃ/	chocolate	[ʃu:ku:latah]	<shu:ku:la:tah>	<شوكولاتة>
	ketchup	[katʃab]	<katshab>	<كتشب>
	chimpanzee	[ʃambanzi]	<sha:mbanzi:>	<شامبنزي>

Additionally, it is important to note here that a possible reason for the non-significant impact of language exposure on the importation rate of target /v/ is Arabic orthography. In words included in Arabic lexicons, e.g., 'chocolate' <شوكلاتة> and 'veto' <فيتو> , novel /v/ and /tʃ/ are

[22] The romanised transliteration was adopted from the Intonational Variation in Arabic Corpus (Hellmuth & Almbark, 2017).

written using Arabic letters <ف> and <ش> representing /f/ and /ʃ/, respectively. The standardization process is very slow in Arabic countries (Alabdaly & Metwally, 2021). Some recently borrowed words, however, (e.g., 'boulevard'<بولفارد> [bu:li:fard] and 'snapchat' <سناب شات> [snabʃat]) have found their way into the written media. In these words, /v/ is often represented by <ف> /f/. In Arabic, two additional dots <ڤ> can be added to the grapheme <ف> to refer to English /v/; however, this is rare given that <ڤ> does not appear on the standard Arabic keyboard.

On the other hand, there is high variation in how target /tʃ/ is written: /tʃ/ is sometimes represented by <ش>/ʃ/ or <ت> /t/ followed by <ش> /ʃ/. This orthographic variation can have a considerable impact on how a speaker might produce /tʃ/ as in 'cheesecake'. A speaker might be exposed to the source word in English <cheesecake>, in Arabic with a grapheme <ش> representing the adapted sounds /ʃ/ <شيزكيك> [ʃi:zkeık] or in Arabic with two graphemes representing the imported sound /tʃ/ <تشيزكيك> [tʃi:zkeık]. Note that even in 'ketchup', which is included in Arabic lexicons, there is <ت> /t/ followed by <ش> /ʃ/. However, Arabic <ت> /t/ represents grapheme <t> in 'ketchup'.

Before ending this subsection, it is worth noting again that level of English exposure was operationalised in this study in terms of three groups (high, medium and low). However, the cluster analysis in Section 7.2 showed that participants overall form only two clusters based on their performance across the perception and production tasks. Future studies, which take this variable into account, should ideally measure participants' levels of English exposure using a continuous scale and divide them into only two groups (a high exposure group and a low exposure group). This would facilitate comparison of participants based on their levels of English exposure and there would be less overlap between groups.

7.3.3 Effect of Word Position

The third factor addressed in the study is word position. The importation rate of target /v/ and /tʃ/ was predicted to be lower in word-final position. Cross-linguistically, in word-final position, voiced obstruents are more marked than their voiceless counterparts and affricates are more marked than fricatives (Eckman, 1991). In non-words, word position was a major factor influencing the importation rate of /v/ but not /tʃ/; participants were more likely to produce /v/ as [f] in word final position. Likewise, word position significantly influenced the perception accuracy of the /v-f/ contrast but not the /tʃ-ʃ/ contrast; the /v-f/ contrast was less likely to be

discriminated in word-final position than other positions. /tʃ-ʃ/ contrast in manner of articulation while /v-f/ contrast in voicing. Voicing contrasts in English are unstable in word-final position. Final /v/ was only partially voiced in auditory inputs, which is a feature of English word-final fricatives (Ogden, 2009). That is, /v/ and /f/ do not differ much in voicing. With this in mind, participants' tendency to produce /v/ as [f] in word-final position could perhaps reflect their perception, because final /v/ was even more similar to /f/.

However, in written-only condition, in which auditory inputs are absent, the importation rate of /v/ was also low in word-final position. The tendency to produce /v/ as [f] in word-final position also reflects cross-linguistic tendencies (markedness) which might be due to articulatory factors. Voiced obstruents, cross-linguistically, tend to be devoiced in word final position (Eckman & Iverson, 1994), probably because the aerodynamic difficulty of producing friction and voicing simultaneously (Ohala, 1983). Initial and intervocalic voiced fricatives are likely to be followed by vowels or other sonorant segments, and hence voicing is likely to be retained or maintained.

In real words, word position did not affect the production of either /v/ or /tʃ/. A possible reason why word position did not have a main effect on the production of target /v/ in real words may be that participants considered some of the borrowed words for target /v/, which are commonly used in the media and public for many years (e.g., receiver, virus and vitamins), as native words rather than foreign words borrowed from another language. Therefore, these words were produced with [f] instead of [v], resulting in an importation rate of target /v/ in real words that was similar in the three different positions.

7.3.4 Effect of Gender

The fourth and final factor we addressed is gender. It was predicted that females would be likely to produce the imported sounds [v] and [tʃ] than males because English is prestigious in Saudi Arabia. As discussed in Chapter 2, some sociolinguistic studies on Arabic dialects showed that women tend to use novel forms more often than men (e.g, Almuhanadi, 1991; Assiri, 2014; Omari & Van Herk, 2016). Although the effect of gender was reported in these earlier studies, it was yet to be explored with regard to loanword phonology. In real words, our results showed a robust main effect of gender on the likelihood of adaptation vs. importation, and the direction was as predicted; females were more likely to produce [v] and [tʃ] than males. The hypothesised reason for this finding was that females would have more positive attitudes than males towards

the English language and American culture. Previous literature has argued that attitudes can strongly influence the importation of novel sounds in loanwords; positive attitudes can result in higher importation rates (e.g., Lev-Ari & Peperkamp, 2014; Paradis & LaCharite, 2012). Surprisingly, the survey results did not support this hypothesis. Overall, the female participants in the high and medium exposure groups did not display more positive attitudes towards the English language and American culture than the male participants in the same groups.

However, the instrumental value of English for procuring a job and becoming financially independent is perhaps more important for women than men. One of the main objectives of Saudi Arabia's 2030 vision is to enable Saudi women to move into jobs that were previously male-dominated (Shalhoub, 2017). Nowadays, there are more opportunities for women to join the workforce than in the past. Another possible reason for this discrepancy could be the nature of the real words used in the study. The recently borrowed words, used in the list of real words for /v/ and /tʃ/, are potentially used more by women (e.g., lavender, mauve, valentine, and cheesecake) than men.

In the non-word production task, the effect of gender was less robust: there was no effect of gender on /tʃ/ production. There was a significant interaction of gender with input condition for /v/ production, whereby females were more likely than males to produce [v] when aural input is available. We speculate that this unexpected finding may be due to female participants' willingness to accommodate to the female native speaker. Female participants may, quite simply, accommodate more to the native speaker's production of target sounds than male speakers do. Further research is needed to examine the impact of speech accommodation on the production of novel sounds in loanwords.

7.4 Putting It All Together

Finally, we take steps to summarize what this study reveals about variation in loanword phonology. When combined, the results demonstrate that the likelihood of adaptation vs. importation of target sounds is attributable to multiple factors, ranging from input modality (written and/or auditory information), to level of exposure to the source language or even word position and certain properties of the lexical items. These results are important in that they show that loanword adaptation is a multifaceted process; variation cannot be explained by one factor alone.

Based on an exhaustive literature review, Hayes-Harb and Barrios (2021) identified four factors that modulate the effect of orthography on L2 phonological learning; namely, systematicity, perceptuality, familiarity and congruence. *Systematicity* relates to the transparency of the L2 writing system in terms of whether or not a novel sound is systematically represented in the L2 writing system. *Perceptuality* refers to learners' ability to perceive a novel sound that is systematically represented in the L2 writing system. *Familiarity* refers to whether or not a speaker is familiar with the L2 writing system, while *congruence* concerns whether or not the letter to sound correspondence is alike in the L1 and L2 (e.g., <m> maps to /m/ in both L1 and L2). These four factors are also expected to influence patterns of adaptation of novel sounds in loanwords. However, familiarity and congruence do not apply in our case because English and Arabic have different writing systems (Arabic script vs Roman script), and Saudi speakers have some knowledge of the English alphabet system because English is a compulsory school subject. We believe that the competing factors, shown in Figure 7.5, influence the adaptation patterns available for each of the two target sounds.

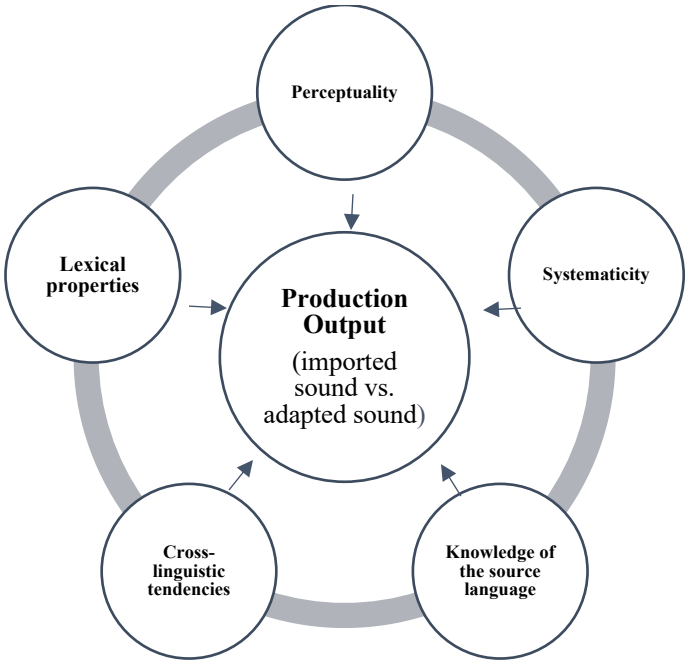


Figure 7.5: *Factors affecting the variability of novel sounds in loanwords*

The five factors in this figure are defined as follows:

- **Perceptuality:** The ability to perceive a novel sound as different from a sound in the L1 (Hayes-Harb & Barrios, 2021).
- **Systematicity:** Whether or not novel sounds correspond systematically to letters in the source language writing system (ibid).
- **Knowledge of the source language:** The amount of knowledge of source language phonological categories and writing system.
- **Cross-linguistic tendencies:** Patterns that are cross-linguistically common (e.g., final devoicing).
- **Lexical properties:** Whether or not a word is recently entered into the borrowing language and whether or not a word is frequently used by the entire speech community.

The adaptation patterns of novel sounds can differ depending on the linguistic information (written and/or auditory) made available to borrowers (Vendelin & Peperkamp, 2006). The influence of *perceptuality* and *systematicity* depend on the input modality of a particular item. *Perceptuality* is an important factor when relying on auditory information. Auditory information will be less helpful in the case of a single category (Best, 1995; Best & Tyler, 2007) in which two segments are captured by a single L1 phonemic category. *Systematicity,* however, plays a key role if exposed to orthographic information. Orthographic information can be less helpful if the orthography of the borrowing language is more systematic or transparent than that of the source language (for example, yielding borrowing language-induced expectations that one letter only represents one sound). Whether a novel sound is spelled with one letter or different letters can have a significant effect on the choice between two variants (imported sound vs. adapted sound) (e.g., Vokic, 2011; Bassetti, 2017). When both orthographic and auditory information are available, the literature suggests that speakers may tend to rely on one or the other. For example, orthographic information can assist with clarifying phonological distinctions between two sounds in some cases (Davidson, 2010). However, auditory information serves as a better authority if a novel sound is spelled in a way that does not align with the speaker's prior experience of letter-to-sound correspondences in the L1 (Bassetti & Atkinson, 2015).

An additional factor that conditions the realization of novel sounds is borrowers' *knowledge of the source language.* Adaptation patterns can differ according to the extent to which a borrower is exposed to the source language (e.g., Kang & Schertz, 2021; Kwon, 2017). Borrowers may have adequate knowledge of the source language's sounds and/or graphemes. This is in contrast with previous work indicating that loanword adaptation is a function solely of non-native speech perception, which assumed that borrowers are naïve listeners with no knowledge of the source language (e.g., Boersma & Hamann, 2009). The greater one's knowledge of the source language, the more likely one is to utilise orthographic and auditory information, demonstrating more target-like production (importation). In other words, speakers with more experience and greater proficiency tend to make better use of orthographic and auditory information to produce a novel sound in cases where the novel sound is *perceptible* to them and/or is *systematically* represented by the L2 writing systems. As borrowers gain more knowledge of the source language, their production presumably becomes less affected by *perceptuality* and *systematicity*.

However, this is a simplification, as the situation is much more complex, due to other forces that play an important role. The likelihood of adaptation vs. importation can be influenced by *cross-linguistic tendencies* which have roots in articulatory and perceptual factors. For example, cross-linguistically, voiced obstruents tend to be devoiced in word-final position (Eckman & Iverson, 1994). Voicing is more likely to be maintained in word- initial and intervocalic positions than in word-final position because of following vowels or sonorant consonants. That is, there is a cross-linguistic tendency that perceptual cues are weaker for a voicing contrast in word-final position than in word- initial and intervocalic positions.

Interestingly, Repiso-Puigdelliura et al. (2021) demonstrate that the influence of orthography can be constrained by cross-linguistic tendencies. Repiso-Puigdelliura et al. (2021) explored variation in the production of the voiced palatal obstruent /ʝ/ by Spanish speakers of Mexican descent living in the USA. The Spanish /ʝ/ is produced with a stronger constriction than /j/, its most similar sound in English. Both sounds are orthographically represented with the same grapheme <y> in Spanish and English. Overall, participants tended to produce the grapheme <y> as [j] more often than as [ʝ]. However, <y> was more likely to be produced as [ʝ] when preceded by high vowels. The authors attributed these finding to the cross-linguistic tendency for stronger constrictions after the utterance of high vowels.

Lexical properties, including frequency and recency, have also been shown to be relevant (Poplack, 2018). The frequent usage of a borrowed word might stabilize its production, speakers will tend to converge on one or another form of production (Daland et al., 2019). As Dohlus (2010) explained, even if a sound is initially produced variably in a loanword, borrowers with different levels of exposure to the source language will eventually converge on a particular mode of production, due to the process of standardisation that occurs when a loanword is frequently used within a speech community. A borrowed word will become filtered through the writing system of the borrowing language because of its repeated use and re-use in the media or public and will be eventually recorded in dictionaries. Hence, individuals will have much more exposure to the adapted forms than the source forms. That is, they are more likely to hear, read and use the adapted form than the source form. In addition to frequency and recency, the relative prestige of lexical items in the semantic domain can influence the likelihood of importation vs. adaptation. Lev-Ari and Peperkamp (2014) examined French speakers' production of novel /dʒ/ in two Italian products; namely, ice cream and beer. These two products showed Italy's relative prestige in France: ice cream has high prestige, while beer is rated relatively low. The authors believed this explained why /dʒ/ was more likely to be imported as [dʒ] when pronouncing ice cream rather than beer.

Let us now have a look at the role that these factors play in how Saudi speakers of Arabic variably produced target /v/ and /tʃ/ in this study. The type of input (auditory and/or orthographic) significantly affected the production of target /v/ and /tʃ/, and this effect varied for the two sounds. Auditory inputs facilitated the production of [tʃ], but hindered the production of [v]. In contrast, written inputs facilitated the production of [v], but hindered the production of [tʃ]. *Perceptuality* and *systematicity* can explain variation in the impact of input type on the two target sounds.

A *perceptually* difficult contrast such as /v-f/ imposes a particular challenge for Arabic speakers, because, as shown in this study, /v/ is less likely to be perceived as different from Arabic /f/. In word-final position, target /v/ may be more likely to be perceived as /f/ because final /v/ can be partially or fully devoiced, rendering it even more similar to /f/ (e.g., Bayley & Holland, 2014; Ogden, 2009). With the availability of orthographic information, target /v/ is more likely to be imported as [v] because grapheme <v> *systematically* maps to /v/.

Conversely, a contrast like /tʃ-ʃ/, which exists in some Arabic dialects, is easier to discriminate, with the result that auditory information alone can afford sufficient data for Arabic speakers to produce the imported sound [tʃ]. With the availability of orthographic information, target /tʃ/ is more likely to be adapted to [ʃ]. The reason for this is differences in *systematicity* between English and Arabic; grapheme <ch> maps to multiple sounds in English.

However, any beneficial orthographic and auditory effects may not be available to Arabic speakers when they are exposed to both auditory and orthographic inputs. The findings of the non-word production task showed that the difference between aural-only and aural-written conditions for target /v/ is not significant, indicating that the importation rate was similar in the two conditions. In contrast, the difference between aural-only and aural-written conditions for target /tʃ/ is significant. These results suggest that the relative difficulty with production of [v] and [tʃ] may not be moderated by the availability of orthographic inputs alongside auditory inputs.

Speakers' *knowledge of the source language* is most evident in how *perceptuality* and *systematicity* contribute to the production of target /v/ and /tʃ/. Speakers with greater prior exposure to the source language were more likely to inhibit the impact of their L1, thereby making better use of auditory and orthographic information to produce the imported sounds. The findings of the non-word production task showed that speakers with high levels of English exposure were more likely to produce [v] and [tʃ] than speakers with low levels of English exposure.

Speakers of Saudi Arabic in this study, irrespective of their level of exposure to English or input modality, were less likely to produce the imported sound [v] in word-final position than in the other positions. As mentioned earlier, there is a *cross-linguistic tendency* to devoice final voiced obstruents. In word-final position, voiced obstruents lose their voicing partially or completely because of the aerodynamic difficulty of maintaining friction and voicing (Ohala, 1983). As a result, the perceptual cues for /v/ are weaker in word final position compared to word-initial and intervocalic positions.

Finally, we believe that *lexical properties* of target real words affected the importation rate of /v/ and /tʃ/. The findings of the real word production task showed that level of exposure to English significantly affected the production of /tʃ/ but not /v/. Speakers with high levels of English exposure were more likely to produce [tʃ] than speakers with low levels of English

exposure. We attributed this null result for target /v/ to the nature of the real words used in the task. The list of real words used for target /tʃ/ is small (8 words), and contained new borrowed words (e.g., snapchat and cheesecake). In contrast, the list of real words used for target /v/ contained very common words (e.g., vitamins, virus, receiver and video) that appear in the Saudi written media, such as *Aljazira* and *Alriyadh* (popular Saudi newspapers). Some of these words are even contained in contemporary dictionaries for Arabic learners (e.g., A frequency dictionary of Arabic: core vocabulary for learners, Buckwalter & Parkinson, 2014).

The findings of the real word production task also showed that gender had a main effect on the production of /v/ and /tʃ/; females were more likely to produce the imported sounds [v] and [tʃ] than males. As shown in Chapter 6, participants' attitudes towards the English language and American culture did not provide an explanation for gender differences in the production of the two target sounds. Overall, females did not display more positive attitudes than males. We speculated that the *lexical properties* of real words used in the task may also account for the observed impact of gender. The recently borrowed words (e.g., mauve and cheesecake) are potentially used by women more than men.

To conclude, we have confirmed and explained that loanword adaptation is not a unitary phenomenon: it cannot be explained by a single factor. Notably, adaptation depends chiefly on how borrowers are first exposed to the source word. *Perceptuality* and *systematicity* both have significant effects on the production of novel sounds. However, these effects are influenced by speakers' *knowledge of the source language.* Speakers with greater prior exposure to the source language are more likely to make better use of auditory information in cases where a novel sound is *perceptible* to them and to utilise orthographic information in cases where a novel sound is *systematically* represented by the L2 writing system. *Cross-linguistic tendencies* also affect the likelihood of adaptation vs. importation in terms that certain patterns are cross-linguistically common (e.g., final voiceless obstruents).

We do not know when or how borrowers will be exposed to source words. We also do not know whether such words will be adapted first by individuals with high or low levels of English exposure; technical words can be adapted even by individuals with low levels of English exposure through reading. However, we hypothesise that the factors discussed above may primarily come into play at the initial stages of the adaptation process. Then, standardisation, resulting from frequency of use and recency, can play a key role in stabilising the production of

borrowed words, leaving less scope for other factors to affect their production. When a borrowed word is commonly used in a speech community, it will be adjusted to the writing system of the borrowing language and eventually become included in dictionaries (Dohlus, 2010). That is, English novel sounds must then be replaced by L1 sounds in order to be represented by the available Arabic graphemes.

In the current digital era, social media (e.g., Twitter and Facebook) may serve to speed the process of standardisation; borrowed words are filtered through the L1 writing system and become increasingly widespread. For example, a word like 'valley' could be written as <فالي> by one user and hence would be pronounced as [fæli] by other users. Novel /v/ is here replaced by the only available Arabic letter <ف> /f/. If such a word becomes widespread in social media, speakers are more likely to produce /v/ as [f] because of their repeated encounters with the Arabic orthographic representation.

7.5 Limitations, Contributions and Future Research

This section presents the limitations of the thesis, points out future research areas, and highlights its main contributions.

The study has some limitations that need to be addressed. First, the present study focuses only on the production of /v/ and /tʃ/ in English loanwords into Arabic. However, previous studies on Arabic loanword phonology also showed variation in the production of other English sounds /ʒ, p, ŋ/ (e.g., Abu Guba, 2016). A subsequent study focusing on other novel sounds could produce interesting findings that account for variability in the production of novel structures in loanwords. In English, target /v/ and /tʃ/ are often represented by <v> and <ch>, respectively. However, multiple graphemes represent /ʒ/ (<si> and <su>), /p/ (<p> and <pp>) and /ŋ/ (<ng> and <nk>). Arabic speakers, especially those with low levels of English exposure, might be misled in cases where different graphemes represent one sound. For example, we expect that novel /p/ will be adapted to [b] by Arabic speakers. However, if novel /p/ is represented by both <p> and <pp> in the stimuli, we expect that /p/ might be adapted as a singleton [b] when it is written with one letter <p> and as a geminate /bb/ when it is written with two letters. This prediction is motivated by previous literature (e.g., Abu Guba, 2016) showing that /p/ is often replaced by Arabic /b/ and the fact that Arabic orthography has a transparent writing system, meaning one letter represents only one sound. Gemination could be caused by English orthography, i.e., /p/ can be spelled with two graphemes <pp>.

Second, all participants in this study were young speakers of Saudi Arabic. Therefore, it is not possible to generalize the results to other populations. Future studies might explore whether similar results would be obtained if the study were replicated, using the same stimuli and procedures, with older Saudi speakers or speakers with different levels of education or speakers of other Arabic dialects. The proposed studies are expected to yield different results. We expect, for example, that orthography would have a different impact on speakers of Moroccan Arabic who are highly exposed to French, in which the grapheme <ch> is used to represent /ʃ/. Older speakers and speakers with lower levels of education might be less likely to utilise orthographic and auditory information to produce the target sounds because they might be less proficient in English compared to young speakers and speakers with higher levels of education.

Third, the perception and production tasks in this study had to be conducted online, as in-person data collection was made impossible by the strict ban on in-person gathering at the time of study due to the COVID-19 pandemic. There was no control over the setting in which the participants engaged with the tasks; the participants used a range of different devices, such as Android phones, iPhones, iPads, and laptops, to participate. We had to cancel one production task because of the pandemic; we originally intended to examine whether a speaker's likelihood of importation vs. adaptation is influenced by interlocutors, using a game similar to the one used by to Lev-Ari and Peperkamp (2014). In their study, the authors used a card game called Go Fish (see 2.3.2 for more details). In future work, it would be interesting to see if speakers are more likely to adjust their production to become more similar or dissimilar to other speakers who tend to use adapted or imported forms.

Fourth, there were six trials for each target and filler item in the perception task. In the pilot study, participants were able to identify the target contrasts of interest because there were no filler items. For this reason, filler items were added to the main study, which made the perception task very long. Thus, it might have been better to reduce the number of trials in order to make the task shorter. However, to maintain statistical power, a higher number of participants would have been needed to participate in the task to get a large number of observations. This was not possible because recruiting participants for the full sequence of tasks online was not easy.

Fifth, in the production tasks, participants were asked to produce each target and filler item twice. Despite this, some participants did not produce all the target items and some

participants produced each item only once. Some tokens were excluded because they were unclear due to background noise. Conducting a similar study under the researcher's in-person supervision would have provided richer data (i.e., more tokens and higher quality recordings).

Finally, phonological working memory was considered for inclusion in this study, but not measured because the perception and production tasks took a long time to complete, and it was already difficult to incentivise participants to complete the full sequence of tasks.

However, despite these limitations, this study contributes considerably to the growing literature on Arabic loanword phonology. Prior studies clearly established the existence of variation between speakers in the production of target /v/ and /tʃ/ in English loanwords into Arabic (e.g., Al-Athwary, 2017; Abu Guba, 2016; Aloufi, 2016; Jarrah, 2013; Saaida, 2015), but none of these studies considered the possible impact of perception, orthography, word position and level of English exposure in an experimental setting.

The present study also adds to the general literature on loanword phonology. The findings of this study show clearly that the adaptation of loanwords is a multifaceted process influenced by different factors, confirming previous evidence of independent perceptual and orthographic effects on loanword adaptation (e.g., Boersma & Hamann, 2009; Dupoux et al. (1999); Kang, 2009; Peperkamp et al. 2008; Vendelin & Peperkamp, 2006). The findings of this study also provide evidence of the impact of cross-linguistic tendencies (Eckman, 1991) and speakers' knowledge of the source language (Kang & Schertz, 2021; Kwon, 2017) on the variable production of novel sounds in loanwords. While confirming that perception, orthography, cross-linguistic tendencies, and language exposure influence the variable production of novel sounds, this study also suggests that lexical properties play a major role in adaptation of borrowed words.

In conclusion, this study provides new insights into the mechanisms that govern variability in producing novel sounds in loanwords. The findings of this study confirm that loanword adaptation is a complex dynamic process; it cannot be explained merely by one factor. Instead, the outcome depends on the interplay of perceptuality, systematicity, cross-linguistic tendencies, lexical properties and borrowers' knowledge of the source language.

Appendices

Appendix A: Norming Studies

Two norming studies with English and Arabic native speakers were carried out to select the appropriate stimuli for the main study. None of the participants took part in the main study. In the first norming study, English native speakers rated the naturalness of a list of English non-words. In the second norming study, Arabic native speakers rated the extent to which they were familiar with a list of English loanwords in Arabic.

A1: Naturalness Rating

The goal of the naturalness rating is to select the target non-word stimuli for the perception and production experiments. A list of 78 CVCVC non-words was used to conduct this norming study. Three questionnaires were administered to 46 participants who speak English as their first language, one as a preliminary questionnaire and the two as follow- up questionnaires.

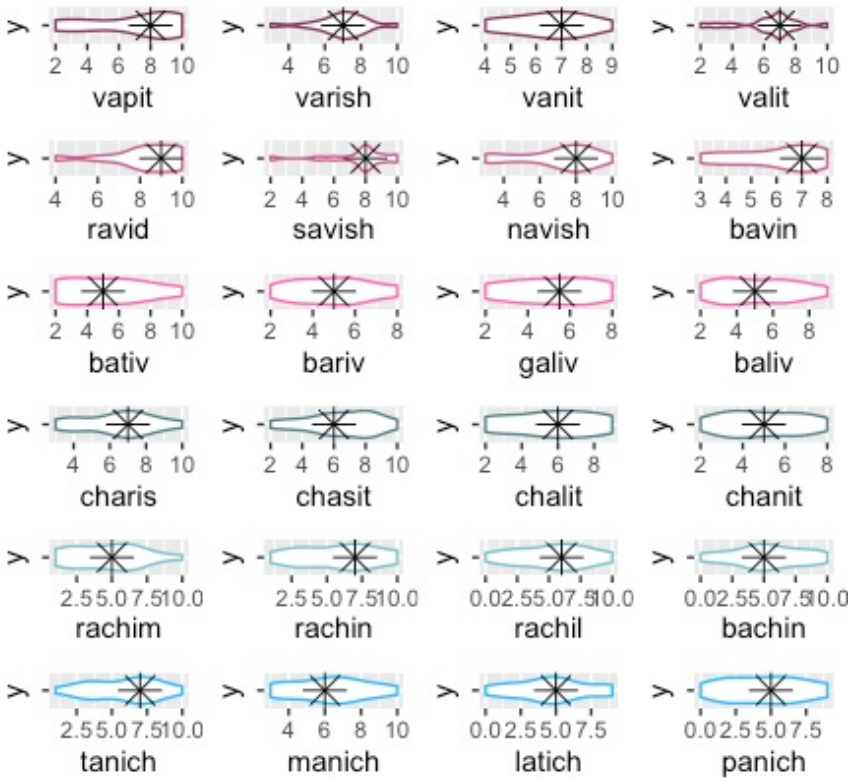
Participants were invited to participate by distributing the link using Twitter, Reddit, and email. The participation was voluntary and anonymous. They received no payment for participation and they were not asked to provide their demographic information or details relating to their accent. To ensure that participants spoke English as their first language, they were asked the following question, "what is your first language", giving the options of "English" and "others". Responses provided by participants, who do not speak English as their first language, were excluded.

The first follow-up questionnaire was distributed three days after the preliminary one. The second follow-up questionnaire was distributed two weeks after the first one. The follow-up questionnaires were carried out to increase the number of candidates for the target stimuli because the results from the first questionnaire showed that some of the candidates had very low ratings. Fifteen participants participated in the first questionnaire, sixteen participated in the second questionnaire, and fifteen participated in the third questionnaire.

The non-word pool involved two word groups. The first group included 52 non-words containing /v/ and the second group includes 26 words containing /tʃ/. The two target phonemes occurred in the three positions: onset, intervocalic, and coda. All the non-words were presented orthographically with their phonetic transcriptions (/CaCɪC/). The questionnaires were carried out online using Qualtrics (www.qualtrics.com). Participants were informed that they would read a list of non-words and would be asked to judge how natural each non-word on a scale ranging from 0 (extremely odd) to 10 (extremely natural)

using the button box. The median is used instead of the mean because it is less susceptible to extreme values. The selected non-words received a median score of 5 or higher. The results showed adequate inter-rater reliability. Raters' Cronbach's alpha is 0.97 for the first questionnaire, 0.71 for the second questionnaire and 0.95 for the third questionnaire. The graph below shows the target items. The star inside the violin plot is the median.

Target Items[23]



[23] Colours represent target words in different word positions. For example, pink = target words containing /v/ in word-final position

Detailed Results of Naturalness Ratings for /v/'s Non-words

Word-Initial	vapit (8)	varish (7)	vanit (7)	valit (7)	vamit (6)	vasin (6)	vanik (4)
Word-Inter-vocalic	ravid (9)	savish (8)	navish (8)	bavin (7)	ravil (6)	mavik (6)	savik (6)
	lavit (5)	wavin (5)	lavik (4)				
Word-final [24]	bariv (5)	galiv (5)	bativ (5)	baliv (5)	madiv (4)	taliv (4)	naliv (4)
	madiv (4)	kariv (4)	ganiv (4)	nariv (4)	Dariv (4)	kaliv (4)	gariv (4)
	radiv (4)	sativ (4)	paliv (4)	saliv (4)	makiv (4)	daliv (3)	laniv (3)
	raniv (3)	sariv (3)	zariv (3)	mariv (3)	paniv (3)	tariv (3)	rabiv (3)
	maniv (3)	tamiv (3)	rariv (2)	raliv (2)	wariv (2)	lariv (2)	zamiv (2)
	maliv (2)						

Detailed Results of Naturalness Ratings for /tʃ/ in Non-words

Onset	charis (6)	chasit (6)	charit (6)	chalit (6)	chanit (6)	chatil (5)	chafit (4)
Inter-vocalic	rachin (7)	rachil (6)	bachin (5)	rachim (5)	tachit (4)	pachik (4)	rachib (2)
	pachib (2)						
Coda	tanich (7)	manich (6)	panich (5)	habich (5)	Latich (5)		

Note: the number in parenthesis under each non-word is the median.

[24] Non-words in word-final position had very low ratings in comparison to non-words in word-initial and intervocalic positions, which is driven by the fact that words ending with /ɪv/ are spelt as <ive> in English.

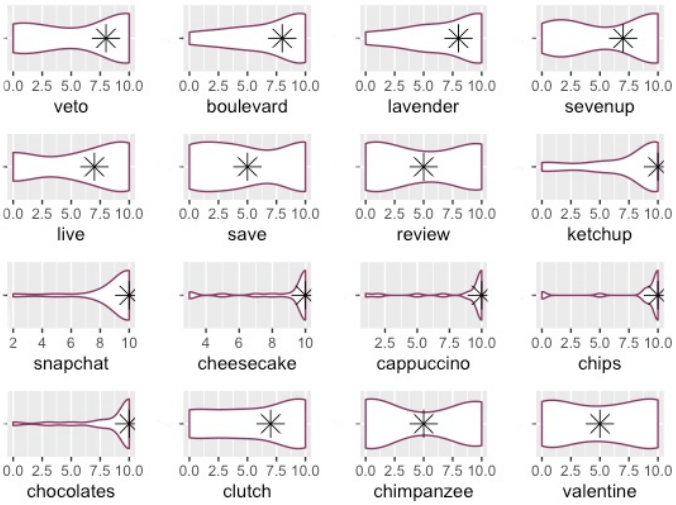
A2: Familiarity Rating

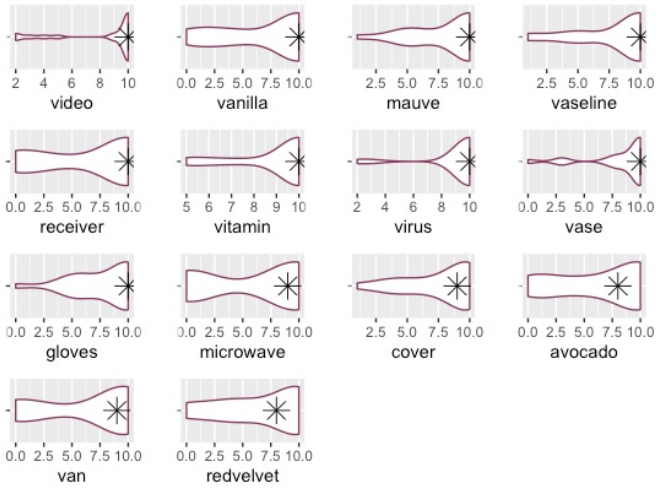
This norming study consists of a familiarity rating for a list of 41 English loanwords in Arabic. The list includes the loanwords that contain either /v/ (n=28) and /tʃ/ (n=13). The loanwords were collected from different resources, such as daily conversations, Arabic websites (e.g., forums, online newspapers, and online shops), networking sites (e.g., Twitter and Instagram), and previous studies (e.g, Abu-Guba, 2016; Aloufi, 2016).

The goal of the familiarity rating is to select the loanwords that are frequently used in society for the production task. To evaluate the frequency of use of the English loanwords in Arabic, 21 Saudi Arabic participants with age ranging from 21 to 55 (mean = 34.80; SD= 9.70) were involved. The participants were 7 males and 14 females who were almost monolingual (i.e., having a very limited knowledge of English). It is difficult to find monolingual speakers as English is taught at schools in Saudi Arabia. All the loanwords were presented individually and in sentences.

The study was conducted online, again using Qualtrics(www.qualtrics.com), and the link was distributed via the WhatsApp application. The participation was voluntary, participants were not paid for participating in the study. They were requested to rate how often they use, hear or read each loanword on a scale ranging from 0 (very seldom) to 10 (very often) using the button box. Raters' responses were consistent, raters' Cronbach's alpha is 0.94. The selected loanwords received a median score of 5 or higher. The graph below shows all the target loanwords.

Target Real Words





D*etailed Results of Familiarity Ratings*

	Loanwords	Median
/v/	1. video	10
	2. vanilla	10
	3. mauve	10
	4. Vaseline	10
	5. receiver	10
	6. vitamin	10
	7. virus	10
	8. vase	10
	9. gloves	10
	10. microwave	9
	11. cover	9
	12. van	9
	13. avocado	8
	14. veto	8
	15. boulevard	8
	16. lavender	8
	17. red velvet	8
	18. seven up	7
	19. live	7
	20. valentine	5
	21. caravan	5
	22. save	5
	23. review	5
	24. movie	4
	25. villa	4
	26. cv	4
	27. caviar	2
	28. vegan	1
		0
		0

/tʃ/	1. ketchup 2. snapchat 3. cheesecake 4. cappuccino 5. chips 6. chocolates 7. clutch 8. chimpanzee 9. check in 10. coach 11. sketches 12. chopsticks 13. matcha	10 10 10 10 10 10 7 5 4 4 1 0 0

Appendix B: Filler Items

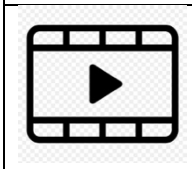
Filler items used in the perception task.

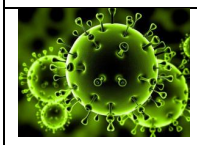
	/t-d/		/m-n/	
/tæCɪC/ vs. /dæCɪC/	tarish	darish	masit	nasit
/mæCɪC/ vs. /næCɪC/	tarik	darik	masik	nasik
	tasin	dasin	malis	nalis
	tasik	dasik	mafis	nafis
/CætɪC/ vs. /CædɪC/	latis	ladis	ramik	ranik
/CæmɪC/ vs. /CænɪC/	ratis	radis	ramit	ranit
	natis	nadis	lamis	lanis
	satip	sadip	samid	sanid
/CæCɪt/ vs. /CæCɪd/	lasit	lasid	narim	narin
CæCɪm / vs. /CæCɪn /	masit	masid	ralim	ralin
	ranit	ranid	barim	barin
	safit	safid	fasim	fasin

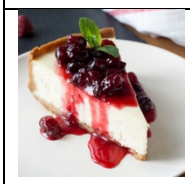
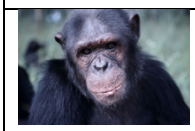
Filler items used in the production task for non-words.

ranik	sarik
rafit	sarish
ranit	lasik
ralif	rasik
ranib	pasik
radis	masit
ladis	rasit
lanib	lasit
sarit	tarish
salin	narin
safit	darif
daris	warif

Appendix C: Pictures and Sentences Used in the Production Task for Real Words

Picture	Sentence	Translation
	سجلت..... على اليوتيوب مدته ثلاث دقايق.	I recoded a 3-minute ……...on YouTube
	احتاج اشري باص او	I need to buy a bus or ……...
	هذا .....يوفر قنوات رياضية كثيره.	This……...offer many sport channels.
	احتاج اشري ..... الشتاء هذا.	I need to buy ……...this winter
	أحب .....اكثر من البنفسجي.	I like……...more than purple.
	راح احط .....لجوالي.	I will put a……...on my phone
	مهمه لصحة الشعر والبشرة.....	……...are essential for healthy hair and skin
	حط الورد في	put flowers in the ……...
	روسيا استخدمت ..... في مجلس الأمن.	Russia used the ……...in UN
	تبي تروح ونترلاند ولا ..... في موسم الرياض.	Do want to go to winter land or…….. in Alriyadh season

	اشرب..... كل يوم.	I drink …….. everyday
	جدتي تستخدم ..... كل يوم.	My grandmother use …….. daily
	أحب ريحة.....	I love the smell of ……..
	استخدم معقم حتى تحمي نفسك من كورونا	Use a sanitizer to be protected from corona ……..
	شريت .....أمس.	I bought this …….. yesterday
	ما أحب اكل	I do not like to eat ……..
	أحب اشوف المبارة.....	I like to watch the match ……..
	استخدم.....قبل تدهن الجدار.	Use before……..painting the wall
	أحب نكهة..... في الآيسكريم.	I love ……..flavour ice-cream
	يوجد في اقصى اليسار	The ……..is on the left
	أفضل تطبيق عندي.....	……..is my best application

	انا احط .....في البطاط المقلي.	I put ……...on fries
	حلاي المفضل	……...is my best dessert
	انا أحب	I love ……...
	فيه..... بحديقة الحيوانات.	There is a ……...in the zoo
	نحتفل .....كل سنه.	we celebrate……...every year.
	اشرب..... كل يوم.	I drink ……...everyday
	انا شريت	I bought ……...
	انا ما أحب كيكة الزعفران. أحب ..... كيكة	I do not like saffron cake. I like ……...

Appendix D: Information Sheets

D1: Perception Experiments' information sheet

Differences in the Production of Words Borrowed from English into Arabic

Study Information

Researcher: Areej Alenazi (aa2138@york.ac.uk)

What is the research about?

The research project aims to examine how foreign words can be produced differently when they are borrowed into another language.

What does the study involve?

In this experiment, you will listen to three sound files on each part of the test. You will listen to same word produced by the same speaker in each of the sound files. You should decide which word is different, with the option to say that the three words are the same if you do not hear an odd word. The expected length of this experiment is approximately 45 minutes.

What will happen to the data I provide?

The Data will be only used for the purposes of this research project. Your participation is voluntary, and you are free to withdraw from the study at any time. Your data and personal information will be anonymized. The data will be kept strictly confidential and stored securely at the Department of Language and Linguistic Science at the University of York. If you withdraw from the study, we will destroy your data and will not use it in any way.

Note: This study is run in compliance with the University's General Privacy Notice and will respect all your rights as described therein.

If you have questions, suggestions or need further information regarding this study, please feel free to contact: aa2138@york.ac.uk

Participant Consent

By clicking " I agree " button below you confirm that you have read and understood the information above, and that you agree to take part in the study.

الأختلافات في نطق المفردات المستعارة من اللغة الإنجليزية في اللغة العربية

الباحثة: اريج العنزي

(aa2138@york.ac.uk)

ما الهدف من الدراسة؟

تبحث هذه الدراسة في كيفية نطق الكلمات بشكل مختلف عندما يتم استعارتها في لغة اخرى

ماهي الاجراءات؟

سوف تتسمع الى تسجيل يحتوي على ثلاث كلمات مسجلة من نفس الشخص. يتوجب عليك ان تحدد الكلمة التى تحتوي على صوت مختلف مع خيار ان تقول ان الكلمات نفسها اذا لم تسمع كلمة مختلفة . اداء هذه التجربة لايتجاوز ٤٥ دقيقة

ملاحظة : يوجد اختبار تجريبي يبين كيفية اتمام المهمة المطلوبة قبل البدء في التجربة

ماذا سيحدث للبيانات التي أقدمها؟

المعلومات التي ستقدمها لن تستخدم إلا لغرض هذه الدراسه. مشاركتك تطوعية في هذه التجربة ويمكن الأنسحاب في أي وقت علماً انه سيتم التخلص من جميع المعلومات في حال قررت عدم المشاركة. جميع معلوماتك الشخصية والمعلومات التي تقدمها سيتم اخفاءها وستظل سرية. سيتم الاحتفاظ بالمعلومات بشكل امن في قسم اللغة وعلوم اللسانيات في جامعة يورك علماً ان هذه المعلومات لن تستخدم إلا لغرض هذه الدراسة.

ملاحظة: يتم إجراء هذه الدراسة وفق وثيقة الخصوصية الخاصة بجامعة يورك وسوف تحترم جميع حقوقك كما هو موضح فيها

شاكرين لكم تعاونكم ومقدرين دعمكم ووقتكم . لمقترحاتكم واستفسارتكم أو اذا كنت تحتاج مزيد من المعلومات حول هذه الدراسة، يرجى التواصل على البريد إلكتروني

aa2138@york.ac.uk

بالنقر على كلمة موافق في الأسفل، تؤكد نك توافق على المشاركة في هذه الدراسة وقد قرأت وفهمت جميع المعلومات المذكورة أعلاه

☐ موافق

D2: Production Experiments' Information Sheet

Differences in the Production of Words Borrowed from English into Arabic

Study Information

Researcher: Areej Alenazi (aa2138@york.ac.uk)

What is the research about?

The research project aims to examine how foreign words can be produced differently when they are borrowed into another language.

What does the study involve?

The expected length of this experiment is between 45 to 60 minutes. You will be asked to make short audio recordings. You will see, hear and/or read some words and international brand names. Produce the words aloud within the given sentences.

What will happen to the data I provide?

The data will be only used for the purposes of this research project. Your participation is voluntary and you are free to withdraw from the study at any time. Your data and personal information will be anonymized. The data will be kept strictly confidential and stored securely at the Department of Language and Linguistic Science at the University of York. If you withdraw from the study, we will destroy your data and will not use it in any way.

Note: This study is run in compliance with the University's General Privacy Notice and will respect all your rights as described therein.

If you have questions, suggestions or need further information regarding this study, please feel free to contact: aa2138@york.ac.uk

Participant Consent

By clicking " I agree " button below you confirm that you have read and understood the information above, and that you agree to take part in the study.

(الأختلافات في نطق المفرادات المستعارة من اللغة الإنجليزية في اللغة العربية)

الباحثة: اريج العنزي

(aa2138@york.ac.uk)

ما الهدف من الدراسة؟

تبحث هذه الدراسة في كيفية نطق الكلمات بشكل مختلف عندما يتم استعارتها في لغة اخرى

ماهي الاجراءات؟

اداء هذه التجربة قد يستغرق من ٤٥ الى ٦٠ دقيقة. تتطلب التجربة القيام بعدد من التسجيلات القصيرة سوف تشاهد ، تقرأ، او تستمع الى كلمات وأسماء لعلامات تجارية عالمية. المطلوب هو ملء الفراغ في جملة مكتوبة وذلك بإستخدام الكلمة التي ترمز لها الصورة او الكلمة التي قرأتها اوسمعتها في الملف الصوتي

ماذا سيحدث للبيانات التي أقدمها؟

المعلومات التي ستقدمها لن تستخدم إلا لغرض هذه الدراسه. مشاركتك تطوعية في هذه التجربة ويمكن الأنسحاب في أي وقت علماً انه سيتم التخلص من جميع المعلومات في حال قررت عدم المشاركة .جميع معلوماتك الشخصية والمعلومات التي تقدمها سيتم اخفاءها وستظل سرية. سيتم الاحتفاظ بالمعلومات بشكل امن في قسم اللغة وعلوم اللسانيات في جامعة يورك علماً ان هذه المعلومات لن تستخدم إلا لغرض هذه الدراسة.

ملاحظة: يتم إجراء هذه الدراسة وفق وثيقة الخصوصية الخاصة بجامعة يورك وسوف تحترم جميع حقوقك كما هو موضح فيها

شاكرين لكم تعاونكم ومقدرين دعمكم ووقتكم . لمقترحاتكم واستفسارتكم أو إذا كنت تحتاج مزيد من المعلومات حول هذه الدراسة، يرجى التواصل على البريد إلكتروني

aa2138@york.ac.uk

بالنقر على كلمة موافق في الأسفل، تؤكد أنك توافق على المشاركة في هذه الدراسة وقد قرأت وفهمت جميع المعلومات المذكورة أعلاه

☐ موافق

D3: Attitudes Survey's Information Sheet

Differences in the Production of Words Borrowed from English into Arabic

Study Information

Researcher: Areej Alenazi (aa2138@york.ac.uk)

What is the research about?

The research project aims to examine how foreign words can be produced differently when they are borrowed into another language.

What does the study involve?

In this questionnaire, you will be asked to indicate your level of agreement or disagreement with each statement on a horizontal line ranging from 0 (strongly disagree) to 100 (strongly agree) The expected length of time to complete this questionnaire is approximately five minutes.

What will happen to the data I provide?

The data will be only used for the purposes of this research project. Your participation is voluntary, and you are free to withdraw from the study at any time. Your data and personal information will be anonymised. The data will be kept strictly confidential and stored securely at the Department of Language and Linguistic Science at the University of York. If you withdraw from the study, we will destroy your data and will not use it in any way.

Note: This study is run in compliance with the University's General Privacy Notice and will respect all your rights as described therein.

If you have questions, suggestions or need further information regarding this study, please feel free to contact: aa2138@york.ac.uk

Participant Consent

By clicking " I agree " button below you confirm that you have read and understood the information above, and that you agree to take part in the study.

الأختلافات في نطق المفرادات المستعارة من اللغة الإنجليزية في اللغة العربية

الباحثة: اريج العنزي
(aa2138@york.ac.uk)

ما الهدف من الدراسة؟

.تبحث هذه الدراسة في كيفية نطق الكلمات بشكل مختلف عندما يتم استعارتها في لغة اخرى

ماهي الاجراءات؟

في هذا الأستبيان ستسأل عن مدى موافقتك على مجموعة من العبارات وذلك عن طريق اختيار اي رقم من 0 (غير موافق بشدّة) الى 100 (موافق بشدّة) وذلك عن طريق تمرير السهم في الخط الافقي الموجود اسفل كل عبارة.مدة تعبئة هذا الاستبيان لا تتجاوز ٥ دقائق

ماذا سيحدث للبيانات التي اقدمها؟

مشاركتك تطوعية في هذا الاستبيان ويمكن الأنسحاب من التجربة في أي وقت علماً انه سيتم التخلص من جميع المعلومات في حال قررت عدم المشاركة. جميع معلوماتك الشخصية والمعلومات التي تقدمها سيتم اخفاءها وستظل سرية. سيتم الاحتفاظ بالمعلومات بشكل امن في قسم اللغة وعلوم اللسانيات في جامعة يورك علماً ان هذه المعلومات لن تستخدم إلا لغرض هذه الدراسة

ملاحظة: يتم إجراء هذه الدراسة وفق وثيقة الخصوصية الخاصة بجامعة يورك وسوف تحترم جميع حقوقك كما هو موضح فيها

.شاكرين لكم تعاونكم ومقدرين دعمكم ووقتكم . لمقترحاتكم واستفسارتكم أو اذا كنت تحتاج مزيد من المعلومات حول هذه الدراسة، يرجى التواصل على البريد إلكتروني

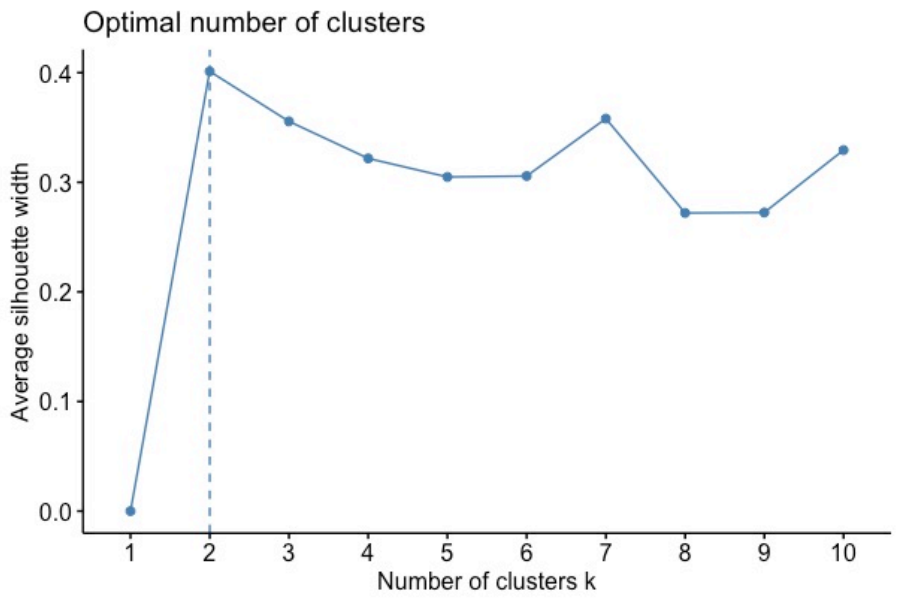
aa2138@york.ac.uk

بالنقر على كلمة موافق في الأسفل، تؤكد نك توافق على المشاركة في هذه الدراسة وقد قرأت وفهمت جميع المعلومات المذكورة أعلاه

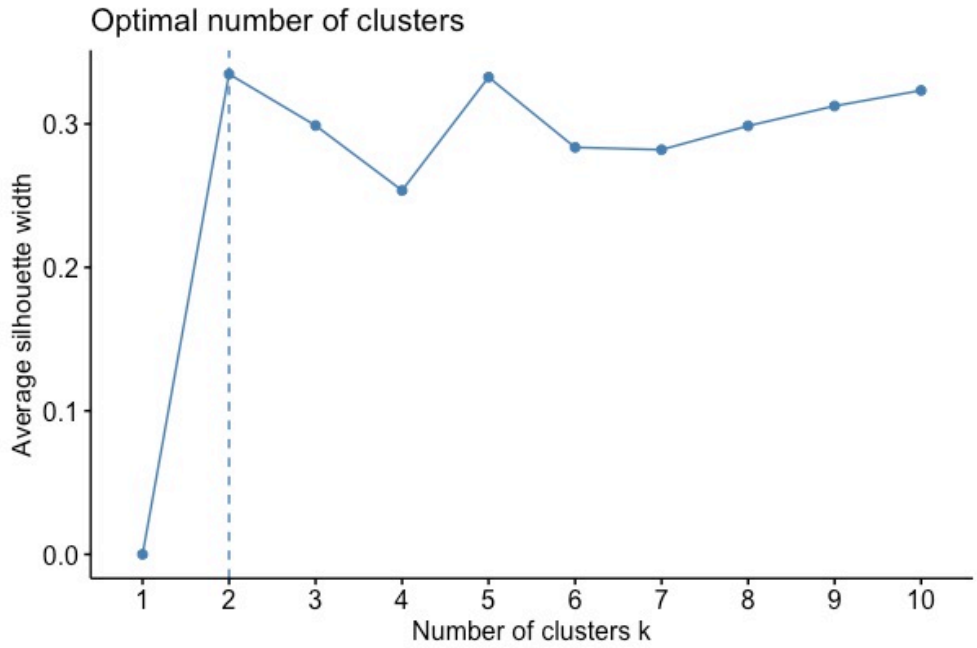
- موافق
- غير موافق

Appendix E: K-means Clustering

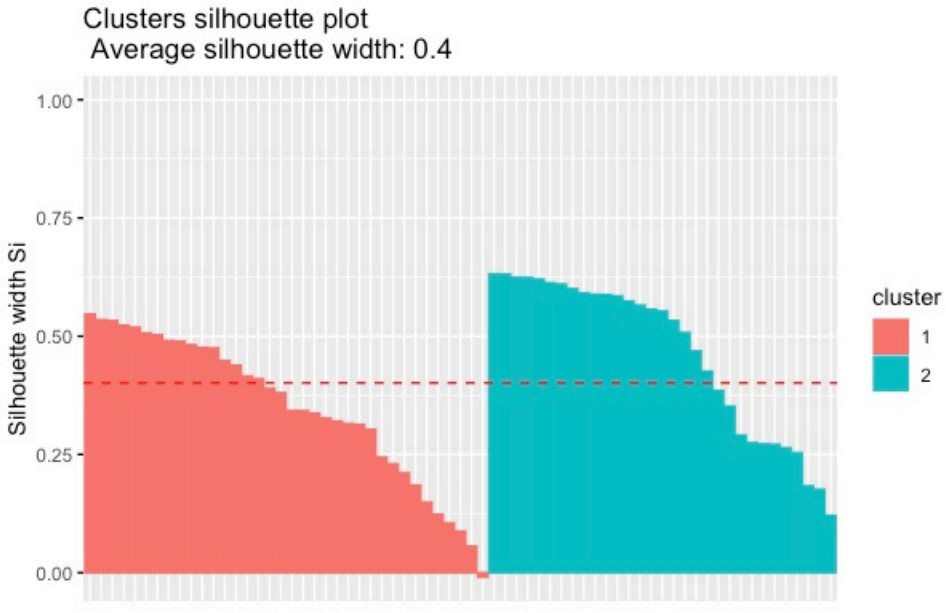
The Optimal Number of Clusters in /v/ Dataset Using the Silhouette Method.



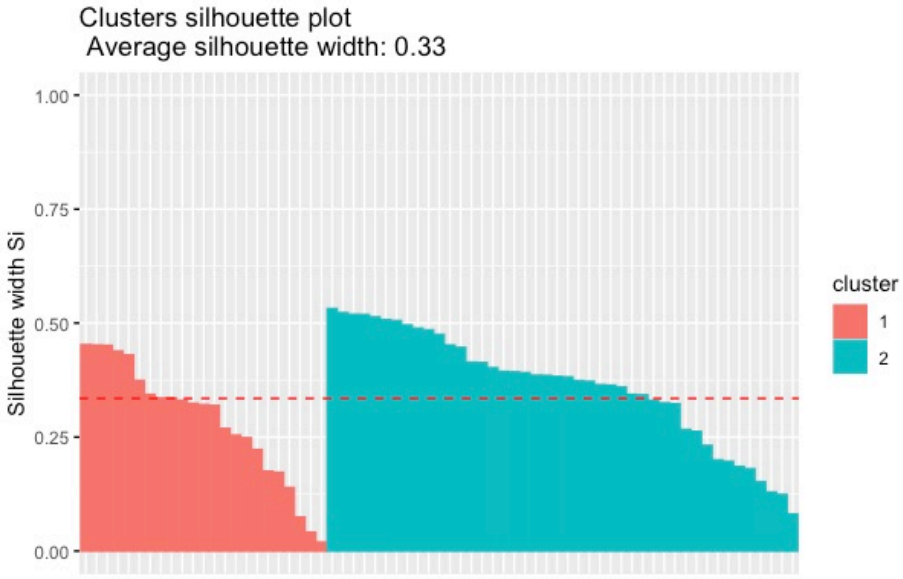
The Optimal Number of Clusters in /tʃ/ Dataset Using the Silhouette Method.



The Average Silhouette Width for /v/ Dataset[25]



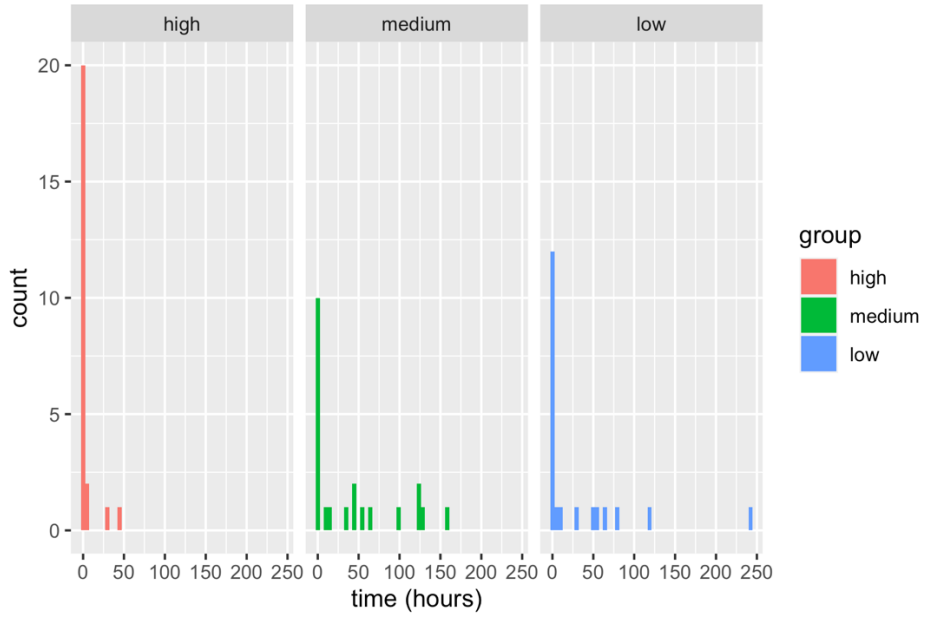
The Average Silhouette Width for for /tʃ/ Dataset.



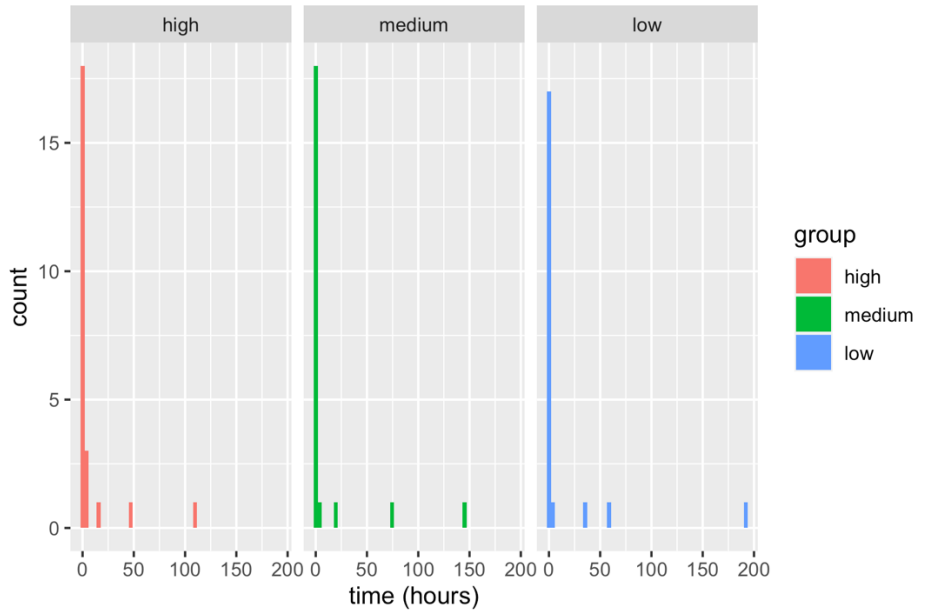
[25] Each color in the figure represents one cluster. The closer the coefficient is to 1, the better the observations are grouped.

Appendix F: Histograms[26] of the distribution of time taken by the participants to complete the perception and production tasks.

The Time Taken by Each Participant to Complete the Perception Task in Hours

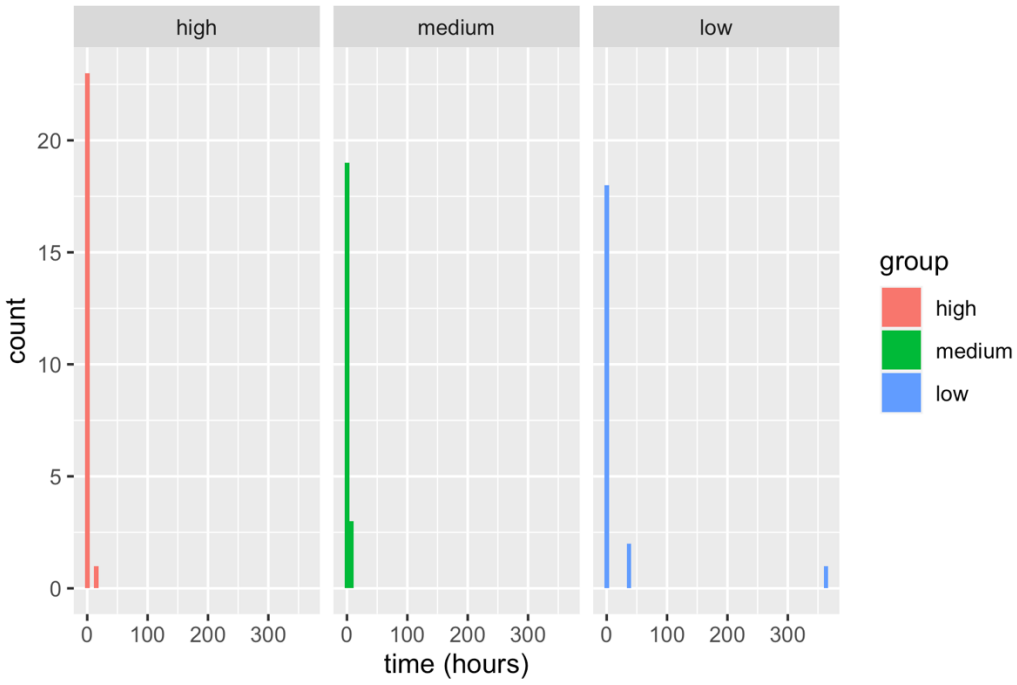


The Time Taken by Each Participant to Complete the Production Task for Non-words in Hours



[26] The numbers on the horizontal axis indicate the number of hours spent on completing the tasks. 0, for example, indicates that the task was completed in less than one hour.

The Time Taken by Each Participant to Complete the Production Task for Real Words in Hours



References

Abu Guba, M. N. (2016). *Phonological adaptation of English loanwords in Ammani Arabic* [Unpublished Doctoral Thesis]. University of Salford. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.698167

Adamson, H. D., & Regan, V. M. (1991). The Acquisition of community speech norms by Asian immigrants learning English as a second language: A Preliminary study. *Studies in Second Language Acquisition*, *13*(1), 1-22. https://doi.org/10.1017/S0272263100009694

Adler, A. N. (2006). Faithfulness and perception in loanword adaptation: A case study from Hawaiian. *Lingua*, *116*(7), 1024-1045. https://doi.org/10.1016/j.lingua.2005.06.007

Alabdaly, A. & Metwally, A. (2021). Arabicization in Saudi Arabia: Procedures and Implementations. *International Journal of Linguistics*, *13* (6), 60-75. https://doi.org/10.5296/ijl.v13i6.19317

Al-Athwary, A. A. (2017). The Phonotactic adaptation of English loanwords in Arabic. *Arab World English Journal (AWEJ)*, *8* (3), 392-406. https://dx.doi.org/10.24093/awej/vol8no3.25

Al-Braik, M. (2007). Performance of KFU English major students. *Scientific Journal of King Faisal University*, *8*(2), 647-677.

Alhussami, A. (2020). *Mutual linguistic borrowing between English and Arabic.* Cambridge Scholars Publishing.

Al-Johani, H. M. (2009). *Finding a way forward: The impact of teachers' strategies, beliefs and knowledge on teaching English as a foreign language in Saudi Arabia* [Unpublished Doctoral Thesis]. University of Strathclyde.

Al-Muhannadi, M. (1991). *A sociolinguistic study of women's speech in Qatar* [Unpublished Doctoral Thesis]. University of Essex.

Al-Saqqaf, A. H. (2006). The linguistics of loanwords in Hadrami Arabic. *International Journal of Bilingual Education and Bilingualism*, *9*(1), 75-93.

Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *International Journal of English Linguistics, 4*(5), 17-26. 10.5539/ijel.v4n5p17

Alshahrani, M. (2016). A brief historical perspective of English in Saudi Arabia. *Journal of Literature, Languages and Linguistics*, *26*(2), 43-47.

Alnamer, A. S. M., & Alnamer, S. A. S. (2018). The use of loanwords in Emirati Arabic according to speakers' gender, educational level, and age. *International Journal of Applied Linguistics and English Literature*, *7*(4), 158-176.

Aloufi, A. (2017). *The Phonology of English loanwords in UHA* [Unpublished Doctoral Thesis]. University of Sussex.

Alzaaq, A. Y. (2017). *Arabic adaptation of loanwords: An empirical examination of pharyngealization and vowel epenthesis* [Unpublished Master's Thesis] California State University.

Ashalhoub, L. (2017, February 22). Remarkable week for Saudi women as females conquer top financial jobs. https://www.arabnews.com/node/1058026/saudi-arabia

Assiri, A. (2014). Sociolinguistic variation in Rijaal Almaʕ, Saudi Arabia: a dialectological study. *King Khalid University Journal for Humanities*, *23*(2), 74-122.

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*, 437-456. https://doi.org/10.1017/S0047404510000400

Bahumaid, S. (2015). Lexical borrowing: The case of English loanwords in Hadhrami Arabic. *International Journal of Language and Linguistics*, *2*(6), 13-24.

Bańko, M., Witalisz, A., & Hansen, K. (2022). Linguistic purism and loanword adaptation techniques: the case of Polish. *Language Awareness*, *31*(1), 95–116. https://doi.org/10.1080/09658416.2021.1990306

Bassetti, B. (2017). Orthography affects second language speech: Double letters and geminate production in English. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(11), 1835-1842. https://doi.org/10.1037/xlm0000417

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics*, *36*(1), 67-91. https://doi.org/10.1017/S0142716414000435

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1 - 48. https://doi.org/10.18637/jss.v067.i01

Bator, M. (2006). Scandinavian loanwords in English in the 15th century. *Studia Anglica Posnaniensia*, *42*, 285-299.

https://link.gale.com/apps/doc/A167977578/LitRC?u=uniyork&sid=bookmark-LitRC&xid=48374381

Bayley, R. & Holland, C. (2014). Variation in Chicano English: The case of final (z) devoicing. *American Speech*, *89*(4): 385–407. https://doi.org/10.1215/00031283-

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ED.), *Speech perception and linguistic experience* (pp. 171-206). York Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775-794. https://doi.org/10.1121/1.1332378

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345.

Best, C. T., & Tyler, M. (2007). Nonnative and second-language speech perception: commonalities and complementarities. In O. Bohn & M. Munro (Eds), *Language Experience in Second language speech learning* (pp.13-34*)*. Amsterdam: John Benjamins Publishing Company.

Boersma, P., Hamann, S. (2009). Loanword adaptation as first-language phonological perception. In A. Calabrese & W. Wetzels (Eds), *Loan phonology* (pp. 11–58). Benjamins. https://doi.org/10.1075/cilt.307.02boe

Boersma, P., & Weenink, D. (2016). Praat, a system for doing phonetics by computer. Retrieved from http://www.praat.org/

Brasington, R. (1997). Cost and benefit in loanword adaptation. *Working Papers in Linguistics*, *3*, 1-19.

Buckwalter, T., & Parkinson, Dilworth B. (2011). *A frequency dictionary of Arabic : core vocabulary for learners / by Tim Buckwalter, Dilworth Parkinson.* Routledge.

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabilary size is associated with second-language vowel perception performance in adult learner. *Studies in Second Language Acquisition*, *33*(3), 433-461. https://doi.org/10.1017/S0272263111000040

Burenhult, N. (2001). Loanword phonology in Jahai. *Lund University Department of Linguistics Working Papers, 48*, 5-14.

Chang, C. B. (2009). English loanword adaptation in Burmese. *Journal of the Southeast Asian Linguistics Society*, *1*, 77-94.

Clyne, M. (2003). *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511606526

Coetsem, F. v. (1988). *Loan phonology and the two transfer types in language contact.* Dordrecht: Foris Publications.

Cruttenden, A., & Gimson, A. C. (2001). *Gimson's pronunciation of English*. Arnold.

Daland, R., Oh, M., & Kim, S. (2015). When in doubt, read the instructions: Orthographic effects in loanword adaptation. *Lingua*, *159*, 70 - 92. https://doi.org/https://doi.org/10.1016/j.lingua.2015.03.002

Davidson, L. (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics*, *38*(2), 272-288. https://doi.org/https://doi.org/10.1016/j.wocn.2010.01.001

Davidson, L. (2018). Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association*, *48*(3), 331-356. https://doi.org/10.1017/S0025100317000330

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology*, *66*(5), 843-863. https://doi.org/10.1080/17470218.2012.720994

Dohlus, K. (2010). *The Role of phonology and phonetics in loanword adaptation*. Peter Lang Verlag. https://www.peterlang.com/document/1136127

Drager, K., Hay, J., & Walker, A. (2010). Pronounced rivalries: Attitudes and speech production. *Reo Te*, *53*, 27-53.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology. Human Perception and Performance*, *25*(6), 1568-1578. https://doi.org/10.1037/0096-1523.25.6.1568

Eckman, F. (1984). Universals, typologies and interlanguage. In W. Rutherford, *Language universals and second language acquisition* (79-105). John Benjamins.

Eckman, F., & Iverson, G. (1994). Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In M. Yavaş, *First and second language phonology* (251-265). Singular Press.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning*, *27*(2), 315-330. https://doi.org/https://doi.org/10.1111/j.1467-1770.1977.tb00124.x

Eckman, F. R. (1981). On the naturalness of interlanguage phonological rules. *Language Learning*, *31*(1), 195-216. https://doi.org/https://doi.org/10.1111/j.1467-1770.1981.tb01379.x

Eckman, F. R. (1991). The Structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, *13*(1), 23-41. https://doi.org/10.1017/S0272263100009700

Edwards, J. (2008). Social factors and variation in production in L2 phonology. In J. Edwards and M. Zampini, *Phonology and second language acquisition (251-279)*. John Benjamins.

Eisenbeiss, S. (2010). Production methods in language acquisition research. In E. Blom and S. Unsworth, *Experimental methods in language acquisition research*, 11-34. John Benjamins.

Elyas, T. (2008). The attitude and the impact of American English as a global language within the Saudi education system. *Novitas-ROYAL, 2*(1), 28-48.

Faruk, S. (2013). English language teaching in Saudi Arabia: A world system perspective. *Scientific Bulletin of the Politehnica University of Timişoara Transactions on Modern Languages, 12*(1-2), 73-80.

Field, A, Miles, M, & Field Z. (2012). *Discovering statistics using R.* Sage Publications.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*(1), 47-65. https://doi.org/https://doi.org/10.1016/S0095-4470(19)30537-6

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, *92*, 233-277.

Foucart, A., & Frenck-Mestre, C. (2013). Language processing. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp.

394-416). Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9781139051729.024

Galal, M. (2004). An OT approach to loanword adaptation in Cairene Arabic. *Kansas Working Papers in Linguistics*, *27*, 1-20. https://doi.org/10.17161/KWPL.1808.1241

Gardner-Chloros, P. (2009). *Code-switching.* Cambridge University Press. https://doi.org/10.1017/CBO9780511609787

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task, *Perception & Psychophysics*, *66*(3), 363-376. https://doi.org/10.3758/BF03194885

Habbash, M., & Troudi, S. (2015). The discourse of global English and its representation in the Saudi context: A postmodernist critical perspective. In R. Raddawi (Eds.), *Intercultural communication with Arabs: Studies in educational, professional and societal contexts* (pp. 57-75). Springer Singapore. https://doi.org/10.1007/978-981-287-254-8_5

Hafez, O. (1996). Phonological and morphological integration of loanwords into Egyptian Arabic. *Égypte Monde Arabe,* (27-28), 383-410.

Hamann, S., & Colombo, I. E. (2017). A formal account of the interaction of orthography and perception. *Natural Language & Linguistic Theory*, *35*(3), 683-714. https://doi.org/10.1007/s11049-017-9362-3

Hamdi, S. (2017). Lexical borrowing in Arabic and the role of orthography. *International Journal of Language and Linguistics*, *4*(2), 17-28.

Hashimoto, D. (2019). Sociolinguistic effects on loanword phonology: Topic in speech and cultural image. *Laboratory Phonology*, *10*(1). https://doi.org/10.5334/labphon.187

Haspelmath, M., & Tadmor, U. (2009). The loanword typology project and the world loanword database. In M. Haspelmath & U. Tadmor (Eds.), *Loanwords in the world's languages* (pp. 1-34). De Gruyter Mouton. https://doi.org/doi:10.1515/9783110218442.1

Haugen, E. (1950). The analysis of linguistic borrowing. *Language*, *26*(2), 210-231. https://doi.org/10.2307/410058

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching*, *54*(3), 297–326. https://doi.org/10.1017/S0261444820000658

Hayward, K. (2000). *Experimental phonetics / Katrina Hayward*. Longman.

Hellmuth, S. & Almbark, R. (2017). *Intonational Variation in Arabic Corpus*. UK Data Archive, Colchester, Essex.

Holden, K. (1976). Assimilation rates of borrowings and phonological productivity. *Language*, *52*(1), 131-147. https://doi.org/10.2307/413213

Horslunda, C. S., Ellegaardb, A. A., & Bohnc, O.-S. (2015). Perceptual assimilation and identification of English consonants by native speakers of Danish. *In Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.

Hsu, B., & Jesney, K. (2017). Loanword adaptation in Québec French: Evidence for weighted scalar constraints. *In Proceedings of the 34th West Coast Conference on Formal Linguistics* (pp. 249-258).

Huang, H.-H., & Lin, Y.-H. (2016). To epenthesize or not? Adaptations of English coda [m] in standard Mandarin loanwords. *Proceedings of the 2015 Annual Meetings on Phonology*. Linguistic Society of America.

Ingham, B. (1994). *Najdi Arabic : central Arabian / Bruce Ingham*. J. Benjamins Pub. https://doi.org/10.1075/loall.1

Ito, J., & Mester, A. (1995). Japanese phonology. In J. Goldsmith (Eds), *The Handbook of Phonological Theory* (pp. 817-838). Blackwell.

Jarrah, A. S. I. (2013). English loan words spoken by Madinah Hijazi Arabic speakers. *AWEJ Special issue on Translation*, 2, 67-85.

Janert, P. K. (2010). *Data analysis with open source tools: a hands-on guide for programmers and data scientists*. O'Reilly Media Inc.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, *108*(3), 1252-1263. DOI: 10.1121/1.1288413

Kang, Y. (2008). Tensification of voiced stops in English loanwords in Korean. *Harvard Studies in Korean Linguistics*, *12*, 179-192.

Kang, Y. (2009). English /z/ in 1930s Korean. *Proceedings of the 2nd International Conference on East Asian Linguistics*. Simon Fraser University.

Kang, Y. (2011). Loanword Phonology. In M. Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 1-25). https://doi.org/https://doi.org/10.1002/9781444335262.wbctp0095

Kang, Y., & Schertz, J. (2021). The influence of perceived L2 sound categories in on-line adaptation and implications for loanword phonology. *Natural Language & Linguistic Theory*,*39*(2), 555-578. https://doi.org/10.1007/s11049-020-09477-9

Kenstowicz, M., & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua*, *116*(7), 921-949. https://doi.org/https://doi.org/10.1016/j.lingua.2005.05.006

Kim, J. (2021). Perception of foreign segments in loanword phonology. *Lingua*, *262*, 103-160. https://doi.org/10.1016/j.lingua.2021.103160

Kim, Hyunsoon (2009). Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. In Calabrese, A., & Wetzels, W. L, *Loan Phonology* (pp. 155-180). John Benjamins Publishing Company. https://doi.org/10.1075/cilt.307.06kim

Kubozono, H., Ito, J., & Mester, A. (2008). Consonant gemination in Japanese loanword phonology [paper presentation]. 18th International Congress of Linguistics, Seoul.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1 - 26. https://doi.org/10.18637/jss.v082.i13

Kwon, H. (2017). Language experience, speech perception and loanword adaptation: Variable adaptation of English word-final plosives into Korean. *Journal of Phonetics*, *60*, 1-19. https://doi.org/https://doi.org/10.1016/j.wocn.2016.10.001

LaCharité, D., & Paradis, C. (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic inquiry, 36*(2), 223-258.

Laver, J. (2003). Linguistic Phonetics. In M. Aronoff & J. Ress-Miller (Eds.), *The handbook of linguistics* (pp. 150-179). Oxford: Blackwell.

Lee, P. (2013). The impact of borrowed sounds and neutralization on Korean contrasts: An entropy-driven analysis. *Indiana University Linguistic Club Working Papers ,(13). https://scholarworks.iu.edu/journals/index.php/iulcwp/article/view/26096*

Lenth, R. (2016). Least- squares means: The R package lsmeans. *Journal of Statistical Software, 69*(1), 1-33. doi:10.18637/jss.v069.i01

Lev-Ari, S., & Peperkamp, S. (2014). An experimental study of the role of social factors in language change: The case of loanword adaptations. *Laboratory Phonology*, *5*(3), 379-401. https://doi.org/10.1515/lp-2014-0013

Lev-Ari, S., San Giacomo, M., & Peperkamp, S. (2014). The effect of domain prestige and interlocutors' bilingualism on loanword adaptations. *Journal of Sociolinguistics*, *18*(5), 658-684. https://doi.org/10.1111/josl.12102

Llamas, C., & Watt, D. (2014). Scottish, English, British?: Innovations in attitude measurement. *Lanuage and Linguistics Compass*, *8(11)*, 610-617. https://doi.org/10.1111/lnc3.12109

Louriz, N., & Kenstowicz, M. (2009). Reverse engineering: Emphatic consonants and the adaptation of vowels in French loanwords into Moroccan Arabic. *Brill's Journal of Afroasiatic Languages and Linguistics*, *1*(1), 41-74.

Mackey, W. (1989). La genèse d'une typologie de la diglossie. *Revue Québécoise de Linguistique Théorique et Appliquée 8*(2):11–28.

Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design, second edition*. New York: Routledge. https://doi.org/10.4324/9781315750606

Mahboob, A., & Elyas, T. (2014). English in the kingdom of Saudi Arabia. *World Englishes*, *33*(1), 128-142. https://doi.org/10.1111/weng.12073

Martin, A. & Van Heugten, M. & Kager, R. & Peperkamp, S., (2022). Marginal contrast in loanword phonology: Production and perception. *Laboratory Phonology* 13(1). doi: https://doi.org/10.16995/labphon.6454

Mok, P. P. K., Lee, A., Li, J. J., & Xu, R. B. (2018). Orthographic effects on the perception and production of L2 mandarin tones. *Speech Communication*, *101*, 1-10. https://doi.org/10.1016/j.specom.2018.05.002

Myers-Scotton, C. (2006). *Multiple voices :An introduction to bilingualism*. Blackwell Pub.

Nagle, C. L. (2021). Revisiting perception–production relationships: Exploring a new approach to investigate perception as a time-varying predictor. *Language Learning*, *71*(1), 243-279. https://doi.org/https://doi.org/10.1111/lang.12431

Nagy, N. (1994). Language contact: Italian (?) geminates in Faetar. *Belgian journal of linguistics, 9*(1), 111-128. https://doi.org/10.1075/bjl.9.08nag

Naim, S. (1998). The adventure of borrowed words in Arabic-Borrowings and pharyngealization. *LINGUISTIQUE*, *34*(2), 91-102.

Niblock, T. (2006). *Saudi Arabia: Power, legitimacy, and survival*. Routledge.

Nimz, K., & Khattab, G. (2020). On the role of orthography in L2 vowel production: The case of Polish learners of German. *Second Language Research*, *36*(4), 623–652. doi:10.1177/0267658319828424

Ogden, R. (2009). *An Introduction to English Phonetics*. Edinburgh University Press. https://doi.org/10.3366/j.ctt1g0b2j9

Ohala, J. J. (1983). The Origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Eds.), *The Production of speech* (pp. 189-216). Springer New York. https://doi.org/10.1007/978-1-4613-8202-7_9

Omar, M. K., & Nydell, M. K. (1975). *Saudi Arabic--urban Hijazi Dialect: Basic Course*. Foreign Service Institute, Department of State.

Omari, O., & Van Herk, G. (2016). A sociophonetic study of interdental variation in Jordanian Arabic. *The Jordanian Journal of Modern Languages and Literature*, *8*(2), 1-21.

Paradis, C. (1987). ON Constraints and repair strategies. *Linguistic Review*, *6*(1), 71-97. https://doi.org/10.1515/tlir.1987.6.1.71

Paradis, C., & LacharitÉ, D. (2008). Apparent phonetic approximation: English loanwords in Old Quebec French. *Journal of Linguistics*, *44*(1), 87-128. https://doi.org/10.1017/S0022226707004963

Paradis, C., & Lacharité, D. (2011). Loanword adaptation: From lessons learned to findings. In J. Goldsmith, J. Riggle, A. Yu (Eds.), *The Handbook of phonological theory* (pp. 751-778). https://doi.org/https://doi.org/10.1002/9781444343069.ch23

Paradis, C., & LaCharité, D. (2012). The influence of attitude on the treatment of interdentals in loanwords: Ill-performed importations. *Catalan Journal of Linguistics*, *11*, 0097-0126. https://raco.cat/index.php/CatalanJournal/article/view/292408

Peperkamp, S., Vendelin, I., & Nakamura, K. (2008). On the perceptual origin of loanword adaptations: experimental evidence from Japanese. *Phonology*, *25*(1), 129-164. https://doi.org/10.1017/S0952675708001425

Perfetti, C. (2007). Reading Ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383. https://doi.org/10.1080/10888430701530730

Poplack, S., & Sankoff, D. (1984). Borrowing: the synchrony of integration. *Linguistics, 22*(1), 99-136. https://doi.org/doi:10.1515/ling.1984.22.1.99

Poplack, S. (2018). *Borrowing : loanwords in the speech community and in the grammar*. Oxford University Press.

Prince, A. & Smolensky P. (1993). *Optimality theory: Constraint interaction in generative grammar.* University of Colorado.

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Repiso-Puigdelliura, G., Benvenuti, I., & Kim, J. Y. (2021). Heritage speakers' production of the Spanish voiced palatal obstruent /j/: A closer look at orthography and universal phonetic principles, *Heritage Language Journal*, *18*(1), 1-30. doi: https://doi.org/10.1163/15507076-12340005

Roever, C., & Phakiti, A. (2017). *Quantitative methods for second language research: A problem-solving approach*. Routledge.

Sa'aida, Z. A. M. (2015). Aspects of the phonology of English loanwords in Jordanian urban Arabic: A distinctive feature, moraic, and metrical stress analysis [Unpublished Doctoral Thesis]. University of Leeds.

Shehata, A (2018). Native English speakers' perception and production of Arabic consonants. In M. Alhawary Eds, *The routledge handbook of Arabic second language acquisition* (pp. 56-69). Routledge. DOI: 10.4324/9781315674261-4.

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). *Afex: Analysis of factorial experiments*. Retrieved from https://CRAN.R-project.org/package=afex

Te Puni, K. (2010). 2009 Survey of attitudes, values, and beliefs towards the Māori language. https://thehub.swa.govt.nz/resources/2009-survey-of-attitudes-values-and-beliefs-towards-the-maori-language.

Thomason, S. G., & Kaufman, T. (1988). Language contact, creolization, and genetic linguistics. University of California Press. https://doi.org/doi:10.1525/9780520912793

Vendelin, I., & Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua*, *116*(7), 996-1007. https://doi.org/10.1016/j.lingua.2005.07.005

Vokic, G. (2011). When alphabets collide: Alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Language Research*, *27*(3), 391–417. https://doi.org/10.1177/0267658310396627

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis.* Springer-Verlag.

Winford, D. (2003). *An introduction to contact linguistics.* Blackwell.

Winford, D. (2010). Contact and borrowing. In R. Hickey (Eds.), *The handbook of language contact* (pp. 170-187). Wiley-Blackwell. https://doi.org/10.1002/9781444318159.ch8

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. *arXiv preprint arXiv:1308.5499*. https://arxiv.org/pdf/1308.5499.pdf

Winter, B. (2019). *Statistics for linguists : an introduction using R.* Routledge.

Yang, P., & Golston, C. (2001). White Hmong loanword phonology. *Proceedings of HILP* 5 (pp. 40–57). University of Potsdam. http://zimmer.csufresno.edu/~chrisg/index_files/HmongLoans.pdf

Zubair, M., Iqbal, M. A., Shil, A., Chowdhury, M. J. M., Moni, M. A., & Sarker, I. H. (2022). An Improved K-means Clustering Algorithm Towards an Efficient Data-Driven Modeling. *Annals of Data Science*, 1-20