The Phonetics and Phonology of Arabic Loanwords in Turkish: residual effects of gutturals

Shadiya al-Hashmi

Doctor of Philosophy

University of York
Language & Linguistic Science

November 2016

Abstract

This thesis takes the adaptation of Arabic loanwords into Turkish as a case to reflect on and contribute to the ongoing debate of loanword phonology of the Perceptual approach (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992), Phonological approach (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001, 2008; Peperkamp et al., 2008; Silverman, 1992) and a medial hybrid model of both phonetics and phonology (Kenstowicz & Suchato, 2006; Shinohara, 2004; Smith, 2006; Chang, 2008 and Dolus, 2013). The thesis includes two types of data: corpus-based and experimental. The corpus of the Arabic loanwords into Turkish comprises 1118 words from which vowel mappings and residual effects of gutturals on neighbouring vowels were identified. Based on the concept of uniformitarianism (Murray, 2013) present-day sound changes must have been governed by the same principles or laws which operated in the past. Thus, one of the goals of this work is to model the grammar of Osmanlica speakers in the perception of modern day Turkish speakers of the residual effects of vowels neighbouring gutturals.

In these effects the Arabic vowels /a/ and /u/ are adapted as /a/ and /u/ in Turkish vowels neighbouring guttural sounds (emphatics, uvulars and pharyngeals); however, the vowel /i/ is borrowed as /ɯ/ only surrounding emphatics and the uvular q and as /i/ elsewhere. It was concluded that the corpus data patterns can be best accounted for by using a hybrid model of phonetics, phonology (of both source and native language) and with the effects of orthography. In addition, the role of bilinguals as the active borrowers in the adaptation process is especially corroborated.

Contents

List of Tables

List of figures

Acknowledgments

First and foremost, I wish to thank my supervisor Dr. Sam Hellmuth for her endless guidance and support without which this thesis would not have seen the light. I am truly indebted to Sam for challenging me at many occasions to reach my potential and for encouraging me to think outside the box when all other solutions have failed. I admire Sam for her professionalism, her high research skills and her open-mindedness especially with people from different cultures. I look up to her and consider her as my role model and would love to work with her in my future research. Thank you, Sam, for everything.

In addition, I would like to extend my gratitude to each and every person who contributed to this work whether through helping with the recording sessions, filling a survey or even submitting my thesis on my behalf. In this context, I would like to thank my colleagues and friends in the department of Language and Linguistics especially Rana Almbark, Sehrish Shafi, Saudi Sadiq and Dima Malahmeh. I would also like to thank other people in the department for helping me with different linguistic and statistical questions including Dr. Márton Sóskuthy, Dr. Dom Watt and Huw Llewelyn-Jones. Moreover, I would like to thank the members of my examination panel Dr. Nancy Kula and Dr. Tamar Keren-Portnoy for their constructive feedback on the dissertation and for their challenging questions. I especially thank Tamar for her many good suggestions and recommendations during the design of the experiments and in the periodic TAP meetings we had.

I would also like to express my thanks to the Turkish community at the University of York and the city of York in addition to all the people who took the online-survey. I appreciate your time and effort. In this vein, I would like to especially thank Oğuzhan Bilgin, Soner Çubukçu, Öykü Şafak Çubukçu, Burcu Şentürk, Mercan Kaya, Ali Bostanci, Selahattin Kosunalp, Marco de Pinto, Yasemin de Pinto and Ertuğrul Uysal. Teşşekular!

Last but not the least, I would like to thank my beloved family for all their help, prayers and love.

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:

The White Rose WRT DC Conference in Leeds (2013)
8th Newcastle upon Tyne Postgraduate Conference in Linguistics (2013)
New Sounds in the North East (2015)

I hereby give consent for my thesis, if accepted, to be made available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.

1 Introduction

This thesis takes the adaptation of Arabic loanwords into Turkish as a case to reflect on and contribute to the ongoing debate of loanword phonology of the Perceptual approach (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992), Phonological approach (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001, 2008; Peperkamp et al., 2008; Silverman, 1992) and a medial hybrid model of both phonetics and phonology (Kenstowicz & Suchato, 2006; Shinohara, 2004; Smith, 2006 Chang, 2008 and Dolus, 2013). The thesis includes two types of data: corpus-based and experimental. The corpus of the Arabic loanwords into Turkish comprises 1118 words from which vowel mappings and residual effects of gutturals on neighbouring vowels were identified. Based on the concept of uniformitarianism (Murray, 2013) present-day sound changes must have been governed by the same principles or laws which operated in the past. Thus, one of the goals of this work is to model the grammar of Osmanlica speakers in the perception of modern day Turkish speakers of the residual effects of vowels neighbouring gutturals.

The thesis is divided into nine chapters as follows where chapter one is the introduction to the work. Chapter two provides general background vis-à-vis Arabic, Turkish and the Arabic loanwords in Turkish (henceforth, ALT) including their historical development, genetic affiliation and the geography of where Arabic and Turkish are spoken. In addition, it provides information on some linguistic topics related to both languages such as vowel harmony in Turkish, emphasis spread and ‘imala in Arabic. Moreover, the chapter presents a review of past studies on loanword phonology including the three models of phonology, perception and the hybrid model of both, and the role of orthography in loanword adaptation.

Chapter three introduces the corpus data of Arabic loanwords into Turkish through the two patterns of mapping long vowels to short vowels and the residual effects of gutturals neighbouring vowels in the Arabic loanwords in Turkish. In the latter phenomenon, /a/ is adapted as /a/ and /u/ as /u/ surrounding emphatics, uvulars or pharyngeals, otherwise /a/ is borrowed as /e/ and /u/ as /y/. However, the vowel /i/ is adapted as /ɯ/ surrounding emphatic or uvular q but not the uvulars /x/, /ɣ/ pharyngeals or plain consonants where it is adapted as /i/. Moreover, the chapter reports on the findings of a stratification task on the etymology of the corpus data. The chapter closes with the rationale for conducting the perceptual study where the perception of monolingual, bilingual and Quranic speakers of Turkish of Arabic vowel categories is tested in the presence and absence of guttural consonants into their Turkish equivalents. This is done in order to model the grammar of the Ottomans to address the role of bilingualism (phonology), perception, orthography and different channels of borrowing.

Chapter four investigates the adaptation of Turkish speakers of the residual effects of gutturals and whether they map short and long pharyngealized vowels in non-words of the form hVd onto different categories (or not). In addition, it investigates the role of Arabic phonology on the mapping of these categories. Three groups of Turkish speakers perform the Perceptual Assimilation Task (PAT), namely monolingual Turkish speakers (T), bilingual Turkish-Arabic speakers (TA) and Quranic speakers of Turkish (TQ). The participants are instructed to identify which Turkish vowel is closest to the vowel they hear. It is found in the PAT that the listeners match the corpus categories by almost 70%. In addition, the listeners exhibit the same perceptual maps and they uniformly mismatch only three vowels ($[a:]_A$>$/e/_T$ predicted as $/a/_T$; $[i^ʕ]$>$/e/_T$ instead of $/ɯ/_T$ and [u]>$/u/_T$ instead of $/y/_T$) to their predicted categories in the corpus data. The main conclusion of the PAT experiment is that perception plays the most part in the adaptations of 70% whereas phonology of Arabic (knowledge of Arabic) plays little role in the matched perceptual maps given that the listener groups yield the same mappings. However, in the remaining 30%, the role of phonology in addition to that of perception is detected based on phonological and phonetic proximity of the mismatched vowels to their predicted categories in the corpus.

Chapter five examines the adaptation of the three Turkish groups of Arabic pharyngealized vowels in both real and non-words in a Simulated Borrowing experiment (audio condition; (SB-audio)). In this task, the participants are instructed to listen to monosyllabic words and write down in Turkish spelling the word they hear. The Turkish spelling being phonetic in nature, in this task writing the responses becomes equivalent to selecting from the set of eight Turkish vowels. The findings show that i) match percentage drops to almost 50% which means that perception alone cannot be responsible for all the adaptation, ii) listeners groups reflect different perceptual maps when the stimuli are a mix of real and non-words, iii) the bilingual group displays higher degrees of match followed by the TQ group suggesting that Osmanlica speakers must have been bilingual too and vi) phonetics and phonology of both Arabic and Turkish (to a lesser extent) explain the assimilation patterns. This chapter concludes that both phonetic and phonological approximation are needed to explain the adaptation of loanwords in the SB- audio task and in the corpus data.

Chapter six considers the effect of orthography on adaptation since in the first part, the three groups are presented with real and nonsense monosyllabic words in two conditions: audio-only and audio-written. The results show that the two groups with Arabic knowledge displayed higher degrees of match in the audio-written condition, with the TA groups reflecting even higher degrees, and higher results than the T group in the audio condition. This reflects that orthography improves the adaptation rate. Furthermore, the TA and TQ groups performed a third task where the stimuli were only written to test whether one group would yield higher degrees of match than the other. The findings show that the TA group yielded higher degrees of match than the TQ group; however, both groups rendered similar mapping patterns. This may be interpreted such that the Ottoman were highly proficient in Arabic. When comparing the matching percentage in the audio-only and the written-only condition, it was found that the percentage of match was slightly higher in the written-only at 50.09 % compared with 50.75% in the audio-only. This result may be interpreted such that the channel of borrowing during the time of Ottomans must have been both spoken and written and probably used for religious purposes.

Chapter seven summarizes the main findings reached in the thesis and discusses what they mean within the field of loans phonology. The chapter concludes that the corpus data patterns can be best accounted for using a hybrid model of phonetics, phonology of both source and native language and with the effects of orthography. It also highlights the role of bilinguals as the active borrowers in the adaptation process.

2 Background and Literature review

This chapter offers some theoretical background and literature review which helped shape the research questions and methodology of the current thesis. Three main research questions are examined. **RQ1**, as raised in the Perceptual Assimilation Task (PAT) chapter (chp.4), addresses the main question of whether speakers of Turkish map the vowels of Arabic to the nearest phonetic categories of their own language, and whether the phonology of Arabic has an effect on the adaptation of the source vowels. **RQ2**, traced in the Simulated Borrowing (SB)-audio data chapter (chp.5), attempts to answer whether speakers of Turkish would generalize the residual effects of gutturals on neighbouring vowels found on the ALT to both real and nonsense borrowed words. **RQ3**, explored in the audio, audio-written and written data chapter (chp.6), investigates whether knowledge of Arabic orthography plays a role in determining the quality of vowels surrounding gutturals.

In order to answer these questions, some macro theoretical topics are addressed in relation to the two languages in contact; i.e., Turkish as the native language, Arabic as the source language and the Arabic loanwords into Turkish (ALT, from now onwards). This chapter is organized as follows. Section 2.1 presents background information on Turkish, Arabic and the ALT in addition to the acoustics of both Turkish and Arabic. Section 2.2 investigates the different loanwords adaptation models and past literature reviews on loanwords cross-linguistically. Section 2.3 summarizes the chapter.

2.1 Turkish background

2.1.1 Genetic affiliation of Turkish and geographical location of where it is spoken

Modern Standard Turkish (MST) as spoken today in Istanbul, Turkey is one of the Oghuz languages group branching from the Turkic family which includes in addition to Turkish other dialects such as Azeri spoken in Azerbaijan (a minority language spoken in north west of Iran Azerbaijan), Gagauz and varieties spoken in the Balkans, the Qashqai in south Iran and the Turkmen in Soviet Turkmenistan (Lewis, 2000; Underhill, 1986).

Underhill (1986) and Ruhlen (1994) among other linguists categorize Turkic (and in turn Turkish) as an Altaic language and a sister language to languages such as Mongol, Manchu-Tunguz, Korean and Japanese. Underhill alludes to a larger Ural-Altaic language group including Finnish, Hungarian and some Siberian languages based on similarities in agglutination, vowel harmony and absence of grammatical gender. This is corroborated by Ruhlen's classification of the world languages (1994) as illustrated by figure 2-1 below.

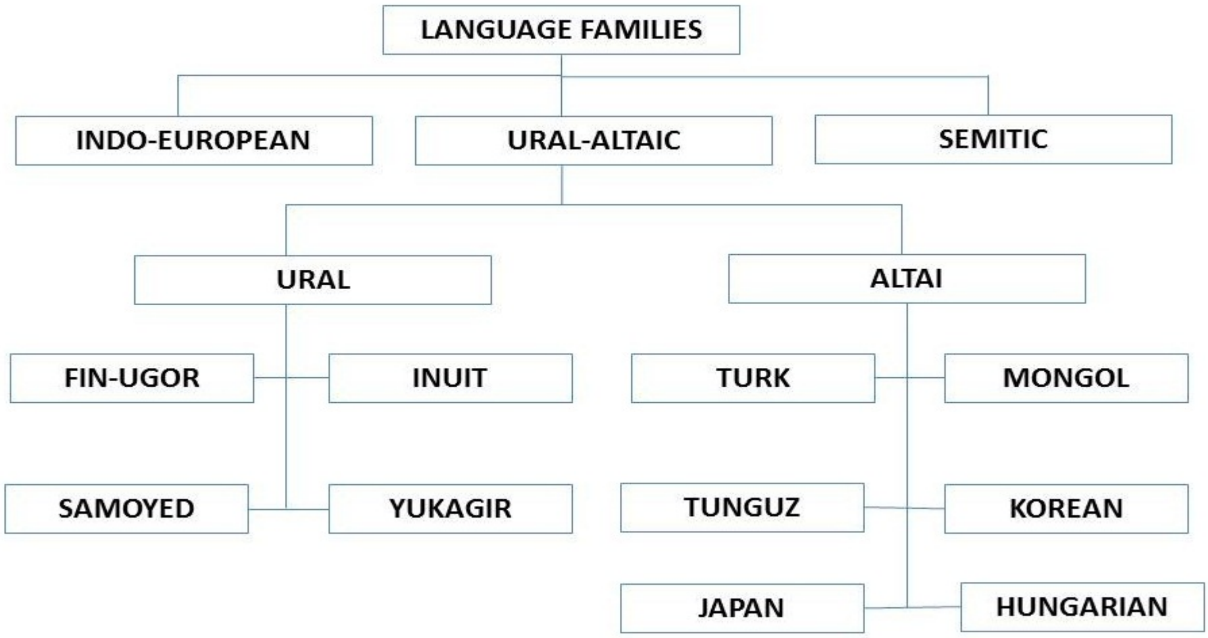


Figure 2-1: Ruhlen's classification of world language adapted from (Saydam, 2008)

Lewis (2000), however, weakens the widely held concept of *Ural-Altaic* family grouping on the grounds that the Turkic family might not be a branch of the Altaic family. At any rate, as Underhill (1986) contends many similarities exist between Altaic and Uralic languages but he then suggests that these might have occurred as a result of continual cultural contact among the people of these languages.

'Turkey Turkish' is the official language of Turkey, a country which stretches between western Asia (Anatolia) and South Eastern Europe, hence the term 'Eurasia.' According to Göksel & Kerslake (2005), no statistics are available to show how many speakers have Turkish as their mother tongue since many bilinguals in turkey belong to some ethnic minorities including Kurds and Arabs.

In 2013, Turkish was spoken by more than 76 million speakers in Turkey itself while many more speak it in other countries. Some of these are Germany, Siberia, Russia, Greece, Bulgaria, Macedonia, Cyprus, other parts of Eastern Europe and parts of the Middle East including Northern Iraq, Syria and Lebanon, the USA, Canada and Australia. Figure 2-2 shows a map of Turkey ("Turkey political map," 2016).

Figure 2-2: Map of Turkey

2.1.2 Historical account of Turkish and Arabic loanwords in Turkish

Modern Standard Turkish as spoken today in Istanbul is the official language of the Republic of Turkey which is sometimes referred to by Turcologists as ‘Turkey Turkish’ to distinguish it from other Turkic dialects spoken outside of Turkey. Prior to adopting MST in Turkey, two Turkish varieties were in use: a Turkic vernacular used on a daily basis by the Turkish uneducated masses among themselves and a high register known as the Ottoman language (*Osmanlı Dil*; Osmanli). Today Osmanli refers to Classical Ottoman which was used from the

16th until early 19th century during the reign of the Ottoman dynasty[1]. Arabic and Persian influence was prevalent in Osmanli, which was a synthetic language composed of mainly Turkish with a large number of vocabulary words and idiomatic expressions from Persian and Arabic. The Ottoman rulers themselves were trilingual (speaking Turkish, Arabic and Farsi). During that time, Arabic and Farsi words were seen as erudite and using them was a sign of prestige; however, the masses spoke ordinary Turkish since they had no access to learning these two languages. It is said that the Ottomans used Arabic as a language of religion and politics, Persian as a language of art, which is reflected in the poetry and literary works of that era and Osmanli as a language of administration (Göksel & Kerslake, 2005). This influence was so strongly felt that some scholars claim that Arabic and Persian words constituted around 65-75% of Osmanli (Stein, 2006).

Modern Turkish was reengineered after the establishment of the Republic of Turkey in 1923 by Mustafa Kemal (Atatürk) who led a nationalist, secularist campaign to purify the Turkish language from language impurities, i.e., Arabic and Farsi, in order to preserve the Turkish identity and transform the country into a modern and western state. Thus, he romanized the Turkish alphabet in 1928 and as a consequence the direction of writing was changed to be from left to right, similar to Western Latin systems. Moreover, he ordered and supervised the establishment of the Turkish Language Society (Türk Dil Kurumu, TDK) in 1932 which consisted of linguists, philologists and Turkish scholars who were missioned to replace Arabic and Persian loanwords and phrases with Turkish ones from Anatolian dialects and other Turkic languages. When there were gaps, the TDK scholars had to derive new words from the roots and stems of Old Turkic, old ottoman words, Turkic dialects and western languages or at times even coin new words altogether (Göksel & Kerslake, 2005).

Despite these efforts to reduce the influence of Arabic and Persian on Turkish, not all Arabic and Farsi loanwords were eliminated. Underhill (1976) reports that in a Turkish textbook, Arabic and/or Persian loanwords formed 35% compared to 62% Turkic and 3% European ones. Aksan (1993) as cited in Versteegh (2001) states that the percentage of Arabic loans in

[1] The Ottoman language can be divided into three periods: *Old Ottoman* (14th– mid 15th century), *Classical Ottoman* (16th-19th centuries) and *late Ottoman* to the *Tanzimat* period (1839-1896) which is characterized by more exposure to Western literature and culture (Saydam, 2008).

Turkish newspapers dropped from 51% in 1931 to 26% in 1965. Many argue that Atatürk's efforts to reform Turkey and the Turkish language were successful, even *catastrophic* as Lewis (1999) puts it in his book title, "*The Turkish Language Reform: A Catastrophic Success.*" This is because the Turks, according to Lewis, have been detached from their past and their literary heritage by means of the new version of Turkish which was rendered simpler and easier to learn by all the masses.

Today, the term 'Arabic loanwords' for most Turkish speakers connotes both Arabic and Persian words and is not limited to original Arabic words. Most Linguists and Turcologists contend that Arabic words were first introduced to Turkish via Persian. Perry (1984) states that one etymological link of Arabic in Turkish to Persian is that of the Arabic loans borrowed with feminine suffixes (-at and -a) which have the same forms and meanings as those in Persian. He, however, refers to Tietze (1958 and 1999) who compiled a large list of words from Arabic vernaculars. Most linguists agree that the first contact between Turkish and Arabic dates back to the 9th century upon the Turks embracing Islam. This accounts for the large number of Arabic religious words in Osmanli and later in Turkish. Titeze (1992) as cited in Johanson (2006) maintains that Arabic words in Turkish were introduced over two stages of language contact. The first stage through Persian during the 9th, 10 and 11th centuries and hence these words were affected by the Persian forms of Arabic. The loanwords of this stage were adapted to Turkish phonology with long vowels '*partly*' being shortened and subject to palatalization in accordance with the rules of Turkish vowel harmony. Additionally, emphatic consonants lost their emphasis and became signals for velarization on neighbouring vowels. General or common words were borrowed during this stage such as words related to Islam, household items and cultural terms and everyday words. Stein (2006, p. 153) collected many words from a 17th century Turkish manual. She attributed many words to the first stage including *'cultural words'* and words of everyday use such as *adam* (Arab. "man"), *avrat* (Arab. "'aurat, privy parts"), *hasta* (Pers. "ill"), *ayna* (Pers. "mirror"), *almas* (Pers. "diamond"), *ramazan* (Arab. "Ramadan; the fasting month") among others. The second stage, Tietze notes, took place through contact with big cultural and religious centres such as madrasas (Qur'anic schools) under the Ottoman empire and involved correction of the older lexical words rendering them similar to Arabic ones. Hence, loanwords of this stage were borrowed directly from Arabic. According to Tietze (1992), the Ottoman language adapted its phonology to a certain extent to that of 'foreign elements', probably to both Arabic/Persian phonology and Western languages' influence especially that of syllable structure as will be shown when discussing the features of both adaptation stages.

It does not seem clear when the two stages began and ended. However, Tietze (1992) contends that the second stage was never completed (p. 350) and that many words resisted the adaptation process. This might be the reason for the many disharmonic forms of Arabic and Persian loanwords in modern Turkish which Tietze (p.358) labels as '*archaic words'*.

Despite the ambiguity surrounding the timeframe of the two borrowing stages of Arabic/Persian loanwords into Turkish, Tietze (1992) outlines the general linguistic features of both stages which lasted for hundreds of years.

He first states that Arabic features were adapted to Persian rules before loanwords were borrowed into Turkish during the first phase. Some of these include depharyngealization or emphatics becoming non-emphatic, Hamza (glottal stop /ʔ/) being assimilated into /y/ before /i/ and word-medial and word-final consonant clusters de-gemination. Some difficulties persisted and became characteristics of the first phase as explained below. All the examples are taken from Tietze (1992, p.351).

1. **Representation of foreign phonemes according to Turkish phonology:**
    a. At the level of consonants, Persian /ʒ/ which was not present in Turkish was adapted either as /ʃ/ or /tʃ/ in Turkish.
    b. At the level of vowels, long vowels could not be borrowed since Turkish does not allow long vowels and were thus shortened.
2. **Turkish syllable structure rules related to clusters:**
    a. Certain consonants were not allowed word-initially, hence vowel insertion took place such as ('rûze' → Turkish '**<u>o</u>**ruč' (fasting)).
    b. Vowels were inserted to break syllable initial clusters (e.g. Persian 'brâdar' → Turkish 'b**<u>u</u>**rader, b**<u>i</u>**râder' (brother)) and some syllable-final clusters such as Arabic 'qatl' → Turkish 'qat**<u>i</u>**l' (murder/ murderer)).

c. Word medial clusters were also not allowed. (e.g. Persian 'pâ**dš**âh' → Turkish 'pâd**i**šâh' (sultan).

3. **Loanwords with palatal and velar consonants within the same word:**

   The example provided was that of 'lûtfen/lütfen' from Arabic /lutˁfan/ ('please') because of the presence of Turkish palatal /l/ which triggers fronting and the Arabic /emphatic /ṭ/ which triggers backing. Such words did not follow the Turkish rules of vowel harmony since the palatal and velar consonants would signal either palatalization or velarization in a word but not both. Velarization/velar harmony (backing) happens when a word has a back consonant such as [/g/, /ɣ/, /k/, /q/ and /x/ in the Arabic word] which spreads its [+back] feature to surrounding vowels. On the other hand, palatalization/palatal harmony (fronting) is triggered by front consonants such as [/l/, /ʃ/, /kʲ/, /gʲ/ and /j/ in the Arabic word] which spread the [-back] feature to neighbouring vowels. Thus, the ALT /lytfen/ *lütfen* 'please' reflects only palatal harmony due to the presence of /l/ despite the fact that the Arabic source word /lutˁfan/ also has /tˁ/ which is a signal of back harmony. However, Turkish phonology only permits either palatalization or velarization in the same word but not both.

As for the second stage, most of the problems from the first phase were overcome. However, some of them were not, resulting in exceptions or disharmonic forms.

1. a. The Persian phoneme /ʒ/ was integrated into Turkish inventory.

   b. Some inconsistencies were rendered vis-à-vis vowel length, e.g. **short** /a/ in 'mal' (Arab. 'property') but **long** /a/ in the phrase 'mal etmek' ('to produce something at a stated cost') since it is pronounced as /ma:l etmek/.
2. Some inconsistencies related to syllable structure including
   a. Words starting with Cs which previously could not appear in initial position during the first phase were incorporated in Turkish during the second phase. Thus, words sometimes started with vowels such as in the colloquial 'Urum' along with the high standard variant 'Rum' (Greek).

b. Syllable medial clusters had forms with syllable medial clusters ('rüzgâr' (wind, breeze)) and no-medial clusters (e.g., 'rûzigâr' with vowel inserted).
c. Syllable final clusters were accepted in Turkish grammar and syllable initial clusters were accepted as a result of contact with European languages.

3. Palatal-velar inconsistencies such as 'lûtfen/lütfen' from Arabic /lutˤfan/ ('please') still persisted. Tietze (1992, p.352) mentions the following consonants signaling velarization in words, /g/, /ġ/ (/ɣ/), /k/, /q/ and /x/, but /l/ as signaling palatalization. In the same vein, Stein (2005, p.149) names 9 emphatic, velar or laryngeal Arabic consonants signaling back harmony including /tˤ/, /ðˤ/, /sˤ/, dˤ/, /zˤ/ /q/, /ɣ/, /x/ and /ħ/, but /l/, /ʃ/, /kʲ/, /gʲ/ and /j/ as consonants signaling palatalization (fronting) (p.146).

In an effort to draw some conclusions about the palatal-velar pronunciations of Arabic and Persian loanwords (the 3rd point above in stage II), Tietze (1992) studied an Ottoman text of 276 stems from the mid-15th century by reference to their Turkish suffixes with the back harmony signal consonants /g/, /ġ/(ɣ), /k/, /q/ and /x/ taken into consideration. The vowel quality of the suffix was determined by the vowel in the immediately preceding syllable in accordance with the rules of vowel harmony. Tietze concluded that the material he inspected dated back to the 2nd stage of adaptation. He argued that the 'words whose vowel quality was not determined by the presence of a signal consonant were not assigned to the palatal category as was the case later on but were classified as velar' (p.357), meaning that the default was velarization/backing.

Tietze's findings about the palatal-velar pronunciations of Arabic loanwords and the adaptation process can be summarized as follows.

1. Words with /a:/ or /u:/ or /o:/ were adapted as /a/, /u/ and /o/ (as back vowels; preserving vowel quality) most of the time with exceptions, whereas those with /i:/ as /i/, i.e. a front/palatal vowel. Vowel length was already adapted to Turkish vowel shortening during the first stage.
2. Fronting in a large number of words might have been triggered by a tense and front /a:/ allophone compared to the lax and back /a:/ variant which triggered backing as

mentioned in 1. Tietze (1992) states that this particular variant was used in Turkish during World War I and was still used in some Arabic loanwords with the feminine ending /a/ when it is realized as /a/ not /e/ before a final /t/. In addition, he specifies the environment of this variant as one in which the /a/ occurs before an emphatic (tezahurat (Arabic taðˤa:hura:t) 'demonstrations'), a geminate (hamal (Arabic ḥammâl) 'porter') or a combination of two consonants where the first element is an /r/ (ayyar (Arabic 'ayyâr) 'crafty; schemer').

3. Arabic loanwords with feminine ending of /a/ and /at/ are generally velar in nature. Tietze points out that most of the exceptions, i.e., words with palatal vowels are Persian and comments that Arabic and Persian words might have been adapted into Turkish differently.

2.1.3 Sound system of MST

Phonemically, Modern Standard Turkish has 8 asymmetric vowels; namely /i/, /ɯ/, /e/, /u/, /y/, /o/, /œ/ and /a/ (figure 2-3) which contrast in the three distinctive dimensions of height, backness and roundness (table 2-1). Despite the fact that the vowel /a/ is phonetically represented as a front vowel, phonologically it behaves as a back vowel. Likewise, the high unrounded vowel /ɯ/ has been debated in the Turkish phonetics literature as being either central or back. However, phonologically it too behaves as a back vowel, hence the use of the IPA symbol /ɯ/ (Kiliç & Öğüt, 2004). Turkish short vowels are exemplified in 1) below followed by the distinctive vowel features (feature combination) of Turkish as given in table 2-2.

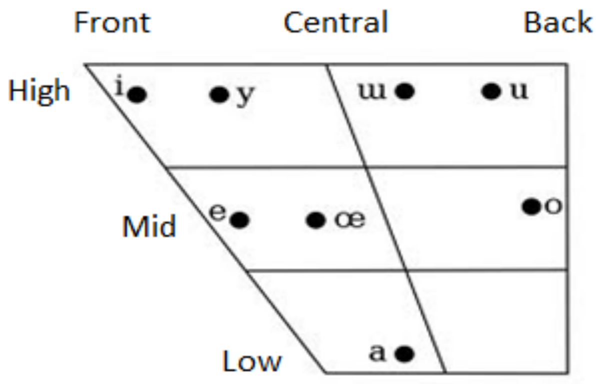


Figure 2-3: Turkish vowels in IPA (International Phonetic, 1999, p. 155)

	Front		Back	
	Unrounded	Rounded	Unrounded	Rounded
High	i [i]	ü [y]	ı [ɯ]	u [u]
Non-high	e [e]	ö [œ]	a [a]	o [o]

Vowel	[front]	[high]	[round]
/a/	-	-	-
/i/	+	+	-
/u/	-	+	+
/e/	+	-	+
/o/	-	-	+
/ɯ/	-	+	-
/y/	+	+	+
/œ/	-	-	+

Table 2-1: Orthographic representation of the Turkish vowels along with IPA symbolism and their feature specifications

(1) Examples of short Turkish vowels embedded in Turkish words

/a/	/bal/	*bal*	'honey'
/i/	/bir/	*bir*	'one'
/u/	/tur/	*tur*	'tour; round'
/e/	/ev/	*ev*	'house'
/o/	/tʃodʒuk/	*çocuk*	'child/infant'
/ɯ/	/kɯz/	*kɯz*	'daughter'
/y/	/syt/	*süt*	'milk'
/œ/	/œ:retmen/	*öğretmen*	'teacher'

Vowel length is not phonemic in pure MST words although according to Kornfilt (1997, p. 501) it can be heard auditorily only through compensatory lengthening as the minimal pair in (2) shows. In addition, residual vowels are witnessed in Turkish as a result of compensatory vowel lengthening where an /h/ is deleted before a fricative or a nasal such as kahve~ka:ve 'coffee' (Kenstowicz, 1994).

(2) Vowel length through compensatory lengthening (Kornfilt, 1997)

dağ 'mountain' [da:] da 'also, too' [da]

Diphthongs are absent in MST although a word like *ay* /aj/ 'moon/ month' for instance may sound like the English word *eye* /aɪ/. However, in English, the vowel combination acts as a diphthong (a single vowel) whereas in MST it does not since the second vowel becomes the onset of the next syllable by the syllable structure rules of Turkish as in *ay* becoming 'a-yı' when the suffix –ı is added to it (Balpinar, 2011).

As for the acoustics of Turkish vowels, F1 and F2 values were plotted using PRAAT (Boersma & Weeink, 2009). Two native Turkish female speakers were recorded, one from Ankara and the other from Gaziantep, reading 48 real monosyllabic Turkish words (8 vowels X 6 words). These are given in table 2-2.

The choice of the speakers to be female was done for uniformity purposes since the Arabic speaker who did the Arabic recordings was also female. Figure 2-4 represents the mean vowel positions of 2 Turkish female speakers where we can see that the three vowels /y, ɯ and œ/ are centralized and the vowel /a/ is almost back. Figure 2-5 shows the mean of Turkish vowel positions for male speakers according to Kiliç (2003) as cited in Kiliç and Öğüt (2004) where /y and œ/ are front, /ɯ/ (HUTV; High unrounded Turkish vowel, (Kiliç & Öğüt, 2004)) is centralized and /a/ is back. The F1 and F2 mean value readings of the two Turkish speakers are given in table 2-3.

word 1	word 2	word 3	word 4	word 5	word 6
had /had/ 'limit'	Can /ʤan/ 'soul'	dar /dar/ 'tight'	bal /bal/ 'honey'	kap /kap/ 'cover'	bağ /baa/ 'orchard'
hiç /hitʃ/ 'never'	bir /bir/ 'one'	diş /diʃ/ 'tooth'	fil /fil/ 'elephant'	bin /bin/ '1000'	Kim /kim/ 'who'
huy /huj/'nature'	çul /tʃul/ 'clothes'	muz /muz/ 'banana'	dul /dul/ 'widow'	sur /sur/ 'fence'	kum /kum/ 'sand'
hem /hem/ 'Both, and'	geç /getʃ/ 'late'	çek /tʃek/ 'check'	şef /ʃef/ 'chief'	ben /ben/ 'I'	her /her/ 'each'
çöl /kœl/ 'desert'	kör /kœr/ 'blind'	göz /gœz/ 'eye'	yön /jœn/'side'	köy /kœj/ 'village'	gök /gœk/ 'sky'
hol /hol/ 'hall'	doz /doz/ 'dose'	toy /toj/ 'immature'	mor /mor/ 'purple'	boş /boʃ/ 'empty'	fon /fon/ 'fund'
hız /hɯz/ 'speed'	kın /kɯn/'sheath'	yıl /jɯl/ 'year'	sır /sɯr/ 'secret'	dış /dɯʃ/ 'outside'	tıp /tɯp/ 'medicine'
hür /hyr/ 'free'	düş /dyʃ/'dream'	yüz /jyz/ '100'	kül /kyl/ 'ash'	gün /gyn/ 'day'	süt /syt/ 'milk'

Table 2-2: Turkish stimulus material for acoustics plotting

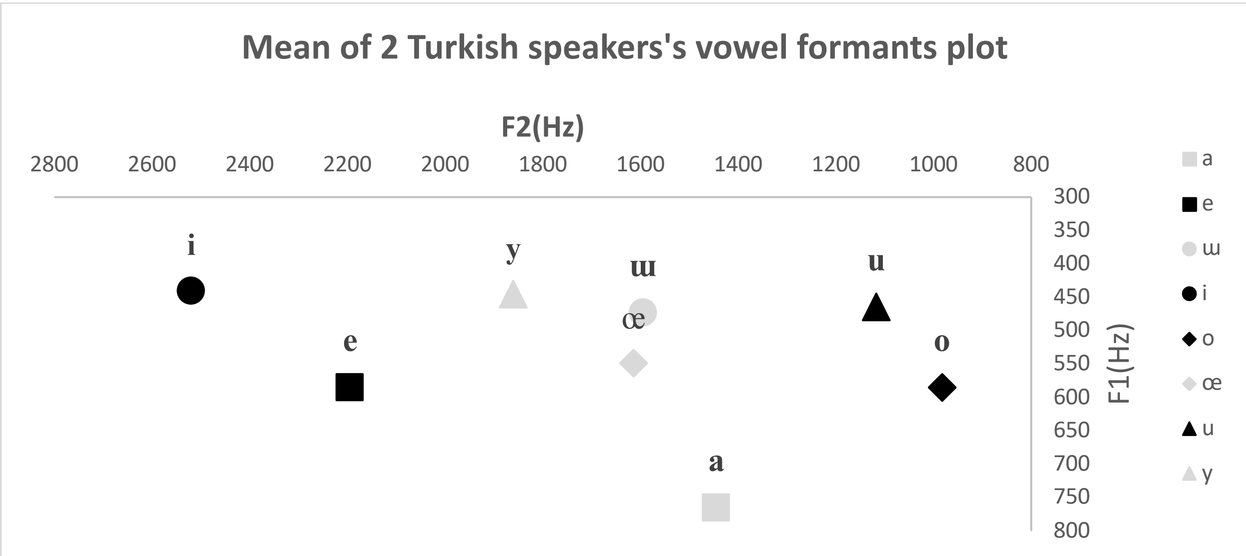


Figure 2-4: Mean of 2 Turkish female speakers' F1 and F2 values as used in this work

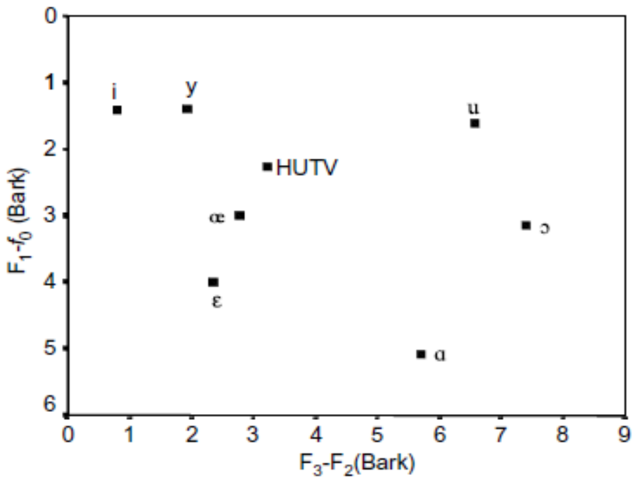


Figure 2-5: Mean positions of Turkish vowels (Kiliç & Giriç, 2003)

vowel	Mean (F2)	Mean(F1)	Word	Gloss
a	1445	766	Can	Soul
ɛ	2194	586	Çek	Cheque
ɯ	1594	473.5	Sır	Secret
i	2519.5	441	Fil	Elephant
o	983	586	fon	Phone
œ	1614	549	çöl	Desert
u	1117	464.5	sur	Wall
y	1860	446	kül	Whole

Table 2-3: Approximated F1 and F2 mean values of 2 female Turkish speakers' production

Regarding its consonant inventory, MST has 23 consonantal phonemes as shown in table 2-4. It is noteworthy to mention that Turkish orthography is phonetic in nature, i.e., most Turkish letters correspond to the IPA transcription system except for specific sounds such as '**ş**'/ʃ/, 'ç'/tʃ/, 'j' /ʒ/, 'c' /dʒ/ and the silent letter 'ğ' known as yumuşak gay[2] (soft g). The description of consonants used here largely comes from Kornfilt (1997).

[2] In Modern Standard Turkish /ğ/is used in Turkish orthography to symbolize what Turks call *yumushak gay* (literally meaning 'soft g'), a consonant that does not surface in the pronunciation of words but has the function of lengthening a preceding vowel (compensatory lengthening).

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive and Affricate	p b		t d		tʃ dʒ	c ɟ	k g	
Nasal	m		n					
Fricative		f v	s z		ʃ ʒ		ɣ	h
Tap				ɾ				
Approximant						j		
Lateral Approximant			ɫ		l			

Table 2-4: Turkish consonants phonemic inventory, (International Phonetic Association, 1999, p. 154)

2.1.4 Vowel Harmony

Vowel harmony (VH) is one of the characteristics of Turkish and Ural-Altaic languages. Under VH, vowels [+syllabic] in Turkish harmonize for backness and roundness to preceding vowels within morphemes and across morpheme boundaries. That is, if the first vowel in a word is front, then all the following vowels are front and vice-verse. Moreover, if the first vowel is rounded, then the following vowels are also rounded and vice-verse. This is expressed linearly by the two rules in (3) and (4).

(3) Back Harmony:

[+syllabic] ---> [αback]/ [+syllabic, αback] (C) __

(4) Round Harmony:

[+syllabic, +high] ---> [αround]/ [+syllabic, αround] (C) __

The words in (5) and (6) exemplify VH both in native MST and in ALT words within morphemes and across morpheme boundaries. Noteworthy to mention, native Turkish words yield to vowel harmony; however, the vowels of the suffixes yor, mtrak, ki, ken, gil and leyin do not assimilate to preceding vowels. Thus, the outcome of suffixation looks like vowel disharmony as in (b3) in (5) “gül+-yor”= “gülüyor” /gylyjor/ (smiling/laughing), (b1) in (6) “hisset-***iyor***”= “hisset*iyor*” /hissetijor/ ‘(s)he/it is feeling’ and “yeşil+-mtrak”=“yeşilimtrak” /jeşilimtrak/ (greenish).

(5) VH in Turkish words

Within morphemes	**Across morpheme boundaries**
(a1) yangın /a/ & /ı/ /jangɯn/ ‘fire’	(a1) kedi-*ler* (plural suffix) /kediler/‘cat’
(a2) sıska /ı/ & /a/ /sɯska/ ‘skinny’	(a2) araba-*lar* (plural suffix) /arablar/ ‘car'
(b1) çizmek /i/ & /e/ /tʃizmek/ ‘draw’	(b1) gel-*(i)yor* (present continuous) /gelijor/ ‘(s)he/it is coming’
(b2) çocuk /o/ & /u/ / tʃodʒuk/‘child/infant’	(b2) al-*(ı)yor* (present continuous) /alɯjor/ ‘(s)he/it is taking’
(c1) röntgen /ö/ & /e/ /rœntgen/ ‘x-ray’	(b3) gül-(ü)yor (present continuous) /gylyjor/ ‘(s)he/it is laughing’
(c2) yüksek /ü/ & /e/ /jyksek/ ‘high’	(b4) bul-*(u)yor* (present continuous) /bulujor/ ‘(s)he/it is dying’
	(c1) ye-*mek* (infinitive) /jemek/ ‘to eat’
	(c2) oku-*mak* (infinitive) /okumak/ ‘to read/study’

(6) VH in Arabic loans into Turkish

Within morphemes	**Across morpheme boundaries**
(a1) vücut /ü/ & /u/ /vydʒut/ ‘body/existence’	(a1) hile-*ler* (plural suffix) /hileler/ ‘trick’
(a2) hortum /o/ & /u/ /hotum/ ‘elephant’s trunk’	(a2) ahşap-*lar* (plural suffix) /ahʃap/ ‘wood’
(b1) akıl /a/ & /ı/ /akɯl/ ‘reason/wisdom’	(b1) hisset-*iyor* (present continuous) /hissetijor/ ‘(s)he/it is feeling’
(b2) fırsat /ı/ & /a/ /fɯrsat/ ‘chance’	(b2) edebiyat-*ı* (accusative suffix) /edebiyatɯ/ ‘literature’
(c1) hizmet /i/ & /e/ /hizmet/ ‘service’	(b3) fark-*lı* (Adjectival suffix) /farklɯ/ ‘different’
(c2) defin /e/ & /i/ /defin/ ‘burial’	(b4) iklim-*sel* (Adjectival suffix) /iklimsel/ ‘climatic’
	(c1) kabul-um (possessive suffix) /kabulum/ ‘acceptance’
	(c2) (c2) adur-*mak* /adurmak/ ‘to continue, keep doing what one is doing’

In the root words ‘*yangın’* and *‘sıska’* in 5(a1) and (a2), the back unrounded vowels 'ı' /ɯ/ and 'a' /a/ assimilate to the [+back], [-round] features of the preceding vowel, ‘a’ /a/ and 'ı' /ɯ/ respectively. The same applies to the Arabic loans *‘akıl’* and *‘fırsat’* in 6(b1) and (b2). In 5(b2) and 6(a2), the [+back], [+round] features of the last vowel /u/ in *‘çocuk’* and *‘hortum’* agree with the preceding vowel’s /o/ features of [+back] and [+round]. In the same manner, the rest of the words with internal VH in 5. and 6. can be described.

VH also applies across morpheme boundaries as in the words to the right in (5) & (6). For instance, the vowels /e/ and /a/ in the suffixes *–**ler*** and *–**lar*** agree for [-back] with the last vowel of the root in *‘kedi’* 5(a1) and *‘hile’* 6(a1), and *‘araba’* 5(a2) and *‘ahşap’* 6(a2) respectively. The suffix denoting the present continuous tense in Turkish (***yor***) as in 5(b1), (b2), (b3), (b4) and 6(b2) is one in which an epenthetic vowel harmonizes as i /i/ (iyor), ı /ɯ/ (ıyor), ü /y/ (üyor) or u /u/ (uyor) for the back and round feature of the last vowel of the root word. In (6), some Arabic loanwords exhibit vowel harmony only across morpheme boundaries; meaning the vowel of the suffix only harmonies with the last root vowel but not the remaining root vowels. An examples is (b2) “edebiyat-***ı”*** /edebiyatɯ/ ‘literature’.

Crucially, what these data in (5) and (6) show is that the front-back and rounded and unrounded agreement relations under VH in Turkish apply to both native Turkish words as well as ALT words. Vowel disharmony (VDH) where vowels do not have the same harmonic features is also witnessed in Turkish as has been widely cited in the literature (Clements & Sezer, 1982; Kirchner, 1993; Krämer, 1998; Van Der Hulst & Van De Weijer, 1991). The data in (7) and (8) elucidate VDH in Turkish and ALT root words and across morpheme boundaries in ALT only. All of the disharmonic Arabic loanwords within morphemes are taken from Clements & Sezer (1982); otherwise, the rest of the data (harmonic and disharmonic) is mine (from introspection).

(7) Vowel disharmony in Turkish native words

<u>**Within morphemes**</u>

(a1) anne /a/ & /e/ /anne/ ‘mother’
(a2) elma /e/ & /a/ /elma/ ‘apple’
(b1) hani /a/ & /i/ /hani/ ‘where is’
(b2) şişman /i/ & /a/ /ʃiʃman/ ‘fat’
(c1) kuzey /u/ & /e/ /kuzej/ ‘north’
(c2) onbir /o/ & /i/ /onbir/ ‘eleven’

<u>**Across morpheme boundaries**</u>

Not applicable

(8) Vowel disharmony in Arabic loanwords

<u>**Within morphemes**</u>

(a1) hesap /e/ & /a/ /hesap/ ‘account/bill’
(a2) haber /a/ & /e/ /haber/ ‘news’

(b1) vakit /a/ & /i/ /vakit/ ‘time’
(b2) kitap /i/ & /a/ /kitap/ ‘book’
(c1) munis /u/ & /i/ /munis/ ‘easy going’
(c2) suret /u/ & /e/ /suret/ ‘copy’

<u>**Across morpheme boundaries**</u>

(a1) harf-*ler* (plural suffix) /harfler/ ‘letters’
(a2) dikkat-*li* (adjectival suffix) /dikkatli/ ‘with precision’
(b1) mahsul-*ü* (accusative suffix) /mahsuly/ ‘produce’
(b2) idrak-*i* (accusative suffix) /idraki/ ‘perception’
(b3) hakikat-*ler* (plural suffix) /hakikatler/ ‘truths’
(b4) misal-*ler* (plural suffix) /misaller/ ‘examples’
(c1) sürat-*li* (adjectival suffix) /syratli/‘fast’
(c2) (c2) sabır-*etmek* (infinitive suffix) /sabɯretmek/ ‘to be patient’

What is interesting to note about the data in (7) and (8) is that some native Turkish words manifest VDH within morphemes in a similar fashion to ALT words and other borrowed words such as /politaka/ *politaka* ‘politics’, /limon/ *limon* ‘lemon’ and /otel/ *otel* ‘hotel’.

Generally speaking, Turkish is a systematic language which is governed by vowel harmony (VH). Most Turkish words yield to the rules of vowel harmony whether stem internally or across morpheme boundaries. Within the VH framework then, Turkish contrasts e-a, i-ı and ü-u as in the present continuous suffix choice (of (i)yor, (ı)yor, (ü)yor, (u)). This, in turn, means that Turkish has a front-back, rounded-unrounded and high-low contrasts as reflected in the rules of VH which are reproduced below for saliency.

a. If the 1st vowel of a word is back, then subsequent vowels are back; and if the 1st vowel is front, then following vowels are front. (VH for backness)
b. If the 1st vowel is unrounded, then subsequent vowels are unrounded. (VH for roundness)
c. If the 1st vowel is rounded, then subsequent vowels are either rounded and close or unrounded and open. (VH for roundness and height)

2.2 Arabic background

2.2.1 Genetic affiliation of Arabic and geographical location of where it is spoken

Arabic is known as a Semitic language; however, two conflicting views regarding its genetic affiliation exist (Faber, 1997, pp. 5, 6). One is that it belongs to the South West branch, a sister branch of both Ethiopian and Modern South Arabian Semitic as cited in Watson (2002, p. 5) who refers in a footnote to a third view by Zaborski (1994; 1997) attributing Arabic to Proto-Afroasiatic. Another prevailing model is that it belongs to the Central Semitic branch which is a sister branch of North-West Semitic, the same branch of languages such as Hebrew, Aramaic and Ugaritic (Hetzron, 1972) as cited in Watson (2002, p. 6).

Regardless of the classification, Arabic manifests some phonological, morphological and syntactic traits which differentiate it from non-Semitic languages. Phonologically, Arabic has a large consonantal inventory of 28 consonants and only three cardinal vowels a-i-u which are contrastive for length. In addition, it groups — as other Semitic languages do— guttural sounds as a natural class[3]. Morphologically, Arabic is famous for its productive triconsonantal root paradigm (root-and-pattern morphology) such as k-t-b 'write' in kataba 'he wrote', from which many more forms are derived. Syntactically, the original word order is verb-subject-object (VSO) in Arabic and Semitic languages in general. However, variations within the Semitic languages do exist including within Arabic itself where in addition to the VSO order, the subject-verb-object (SVO) order is widely used in many Arabic dialects (Watson, 2002).

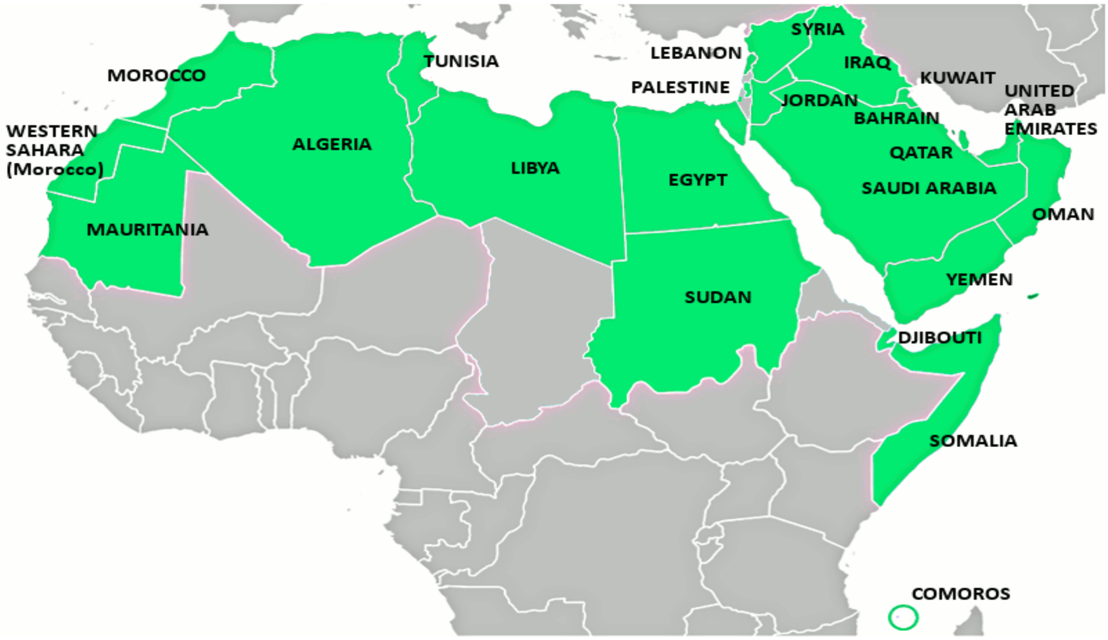


Figure 2-6: Map of the Arab world

Arabic today is spoken by approximately 300 million people (Owens, 2013, p. 5). Arabic is the **official language** of the Arab League countries, some 25 countries in the Middle East, stretching from West Asia and southern Iran to North and central Africa. This includes Oman, Yemen, United Arab Emirates, Qatar, Kuwait, Bahrain, Saudi Arabia, Jordan, Syria, Palestine, Lebanon, Egypt, Libya, Tunisia, Algeria, Morocco, Muritania, Somalia, Djibuti, Sudan, North Sudan and Comoro Islands (figure 2-6). It is also spoken by minority groups or as a **co-official** language as in Eriteria, Chad, Zanzibar, Western Sahara and Israel (Owens, 2013). In addition, Arabic is spoken by minority groups in south-western Iran, southern Turkey, western Africa and by immigrants around the world (Watson, 2002, p. 8)

[3] As discussed in section 2.1.3.1.

2.2.2 Arabic development and diglossia

Modern Standard Arabic (MSA) and Classical Arabic (CA) are the most widely used terms when referring to Arabic. Fischer (1969) maintains that Arabic in the eighth century became more standardized as a result of development in grammar, hence, the term Standard Arabic. MSA is the literary variant used in the media of the countries where Arabic is the official language, for example in TV, news broadcasts and press. It is contended that no Arabic speaker has MSA as their mother tongue any more (Watson, 2002). In addition to MSA (formal/high register), there exist different Arabic vernaculars or dialects (low register) spoken on day-to-day basis in homes and outside in social interactions. Examples of these *inter alia* include Egyptian Arabic, Syrian Arabic, Gulf Arabic, Sundanese Arabic, and Moroccan Arabic.

MSA is a descendent of Classical Arabic which was spoken by Arab tribes in the Arabian Peninsula including Hijaz, Najd and their bordering tribes. An example of the use of Classical Arabic is that of Ashshi'r Aljaahilii (pre-Islamic poetry) when poems where written in gold and the best of which were hung on the curtains of Ka'bah. Ar-Rajhi (1969) cites a saying by Ibn Abbas, who was a companion of the Prophet Mohammed stating that the Holy Quran, where Classical Arabic is codified and preserved, was revealed in *Seven Ahruf* (dialects). Some examples of these seven include four dialects spoken by the tribes of *Hawazin,* being Sa'ad Bin Bakr, Jasham Bin Bakr, Nasr Bin Mo'awiyah and Thaqeef (ibid).

This language situation where two or more varieties coexist within the same territory but with one being of a higher register than the other is called diglossia in sociolinguistics. According to Ferguson (1959, p. 336), "In addition to the primary dialects of the language ..., there is a very divergent, highly codified ... superposed variety, the vehicle of a large respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any sector of the community for ordinary conversation". That is, in Arabic MSA and CA would be considered higher varieties which are more formal and more eloquent than the many Arabic vernaculars.

2.2.3 Sound system of MSA

In contrast to MST, MSA has only three vowels; /a/, /i/ and /u/ which are contrastive for length (i.e., /aː/, /iː/ and /uː/) and two diphthongs: /aj/ and /aw/ as in /sajl/ 'torrent, swift and violent flood of water' and /qawm/ 'a group of people' respectively. Long vowels are represented in the written form as ?alif (أ), yaa? (ي) and waaw (و), whereas short vowels are only optionally represented (diacritics) as fatHah (َ), Dhammah (ُ) and kasrah (ِ); so for the long vowels there is a clear indication of what the vowel 'should' be, whereas for the short vowels these can be figured out from perception alone. Vowel length is shown by the near and/or minimal pairs in (8) and the IPA vowel chart of MSA as given in Figure 2-7.

(8) Vowel length contrast in Arabic

/a/	**/ʕadda/**	'he counted'	عَدَّ
/a:/	**/ʕa:da/**	'he returned'	عاد
/i/	**/ʕid/**	'promise (imperative)'	عِد
/i:/	**/ʕi:d/**	'Eid; celebration'	عيد
/u/	**/ʕud/**	'turn around (imperative)'	عُد
/u:/	**/ʕu:d/**	'oud, musical instrument' Or incense chips'	عود

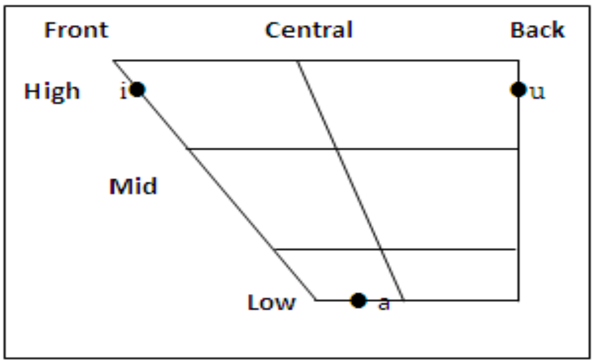


Figure 2-7: Arabic vowels in IPA (International Phonetic, 1999, p. 52)

On the other hand, pharyngealized/uvularized vowels in Arabic are not phonemic but rather allophonic. This is because they are predictable and depend on the presence of gutturals in words. Some examples include the words in (9) where pharyngealization spreads rightward from the triggering guttural consonant (in bold) to the target vowels (underlined).

(9)

[aˤ]	**dˤ**aˤraˤr	'harm'
[a:ˤ]	**sˤ**a:ˤd	'a letter in the Arabic alphabet ص'
[iˤ]	**tˤ**iˤfl	'baby'
[i:ˤ]	ba**sˤ**i:ˤr	'can see; not blind'
[uˤ]	**tˤ**uˤbu:ˤl	'drums'
[u:ˤ]	**tˤ**u:ˤl	'length'

In Arabic, there is a clear front/back difference between [a]~[aˤ] (i.e. in IPA this would in fact be [a]~[ɑ]); but that for i/u (i.e. [i]~[iˤ] and [u]~[uˤ]) the +/- emphatic acoustic distinction is somewhat smaller. This is illustrated in figure 2-8 below of the F1/F2 plot (of Arabic).

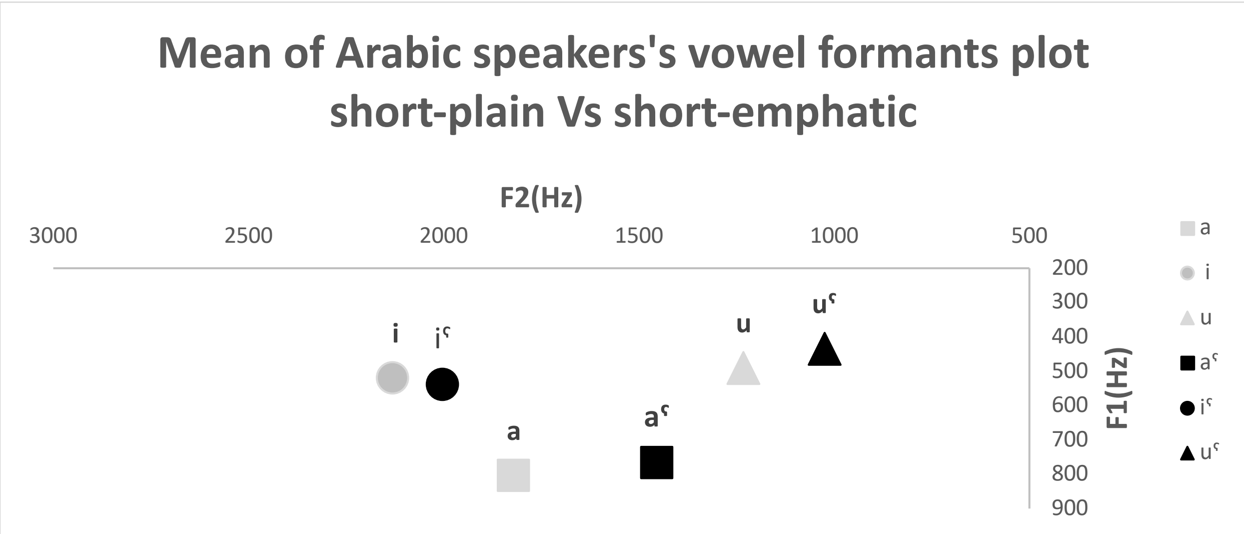


Figure 2-8: F1/F2 plot of short plain-emphatic Arabic vowels

As for the relevant feature specifications of Arabic vowels, these are given in table 2-5 below. Rounding can be predicted from the value of the [front] feature for all Arabic phonemes, thus the feature [round] is underspecified, or inactive, in the language. The pharyngealised 'a' [ɑ] is a counterexample to this (it is back but unrounded, in most dialects, though not all) hence using the word 'phoneme' in the sentence above.

Vowel	[front]	[high]	[round]
/a/	+	-	-
/i/	+	+	-
/u/	-	+	+
/a:/	+	-	-
/i:/	+	+	-
/u:/	-	+	+

Table 2-5: Feature specifications of Arabic

Most analyses now encode phonemic length prosodically, i.e. in representation of the syllable structure of the word (Odden, 2011). However, phonemic vowel length in Arabic can be analyzed, as in SPE, in terms of binary features, i.e. +/-long feature. In the current work, the +/- long features are used for Arabic as a 'shorthand' to express the phonemic length contrast without committing to a particular stance on its underlying representation.

Similar to the Turkish recorded data, 72 Arabic real monosyllabic words of the 12 Arabic vowel categories (3 short and plain, 3 short and emphatic, 3 long and plain, 3 long and emphatic X 6 words each) were recorded by a female Syrian speaker from Aleppo and were later plotted using PRAAT (Boersma & Weeink, 2009). The words are given in table 2-6 below. The choice of the speaker to be of Syrian origin stems from the following observation. The ratio of Arabic words of Syrian origin compared to those from other varieties in Tietze's (1958 and 1999) lists is higher being 72 words out of 216 compared to 2 Lebanese words, 1 Iraqi and 1 Egyptian. The remaining 140 words seem to have been borrowed from Classical Arabic or from other Arabic dialects which Tietze did not mention.

1	2	3	4	5	6
/rabb/ رَب God	/ʃarr/ شَر evil	/samm/ سَم poison	/fann/ فَن art	/fadʒdʒ/ فَج mountain highway	/fakk/ فَك jaw
/birr/ بِر righteousness	/dʒinn/ جِن jinn	/zirr/ زِر button	/dʒidd/ جِد find	/sirr/ سِر secret	/zidd/ زِد add
/bunn/ بُن Coffee (powder)	/ʔumm/ أم mother	/sull/ سُل tuberculosis	/dubb/ دُب bear	/wudd/ وُد love/like	/kull/ كُل eat/every/all
/ma:tt/ مَات died	/sa:dd/ سَاد prevailed	/da:ss/ دَاس stepped on	/ta:mm/ تَام complete	/za:rr/ زَار visited	/ba:bb/ بَاب door
/ti:nn/ تِين figs	/di:rr/ دِير abbey	/di:kk/ دِيك rooster	/fi:ll/ فِيل elephant	/r:ʃʃ/ رِيش feathers	/ri:mm/ رِيم proper name
/du:dd/ دُود worms	/nu:rr/ نُور light	/tu:tt/ تُوت blueberry	/fu:ll/ فُول fava beans	/du:rr/ دُور dwellings	/su:rr/ سُور wall
/dˁabb/ ضَب giant lizard (Mastigure)	/batˁtˁ/ بَط ducks	/sˁakk/ صَك document	/qadd/ قَد may	/xall/ خَل vinegar	/ɣall/ غَل shackle
/dˁid/ ضِد against	/ðˁill/ ظِل shadow	/qinn/ قِن Chicken coop	/qiff/ قِف stop	/xill/ خِل lover	/ɣill/ غِل venom
/sˁunn/ صُن protect	/dˁurr/ ضُر harm	/xuðð/ خُذ take	/tˁuff/ طُف rotate	/qull/ قُل say	/ɣull/ غُل does not spend
/ɣa:rr/ غَاب disappeared	/sˁa:rr/ صَار became	/ɣa:rr/ غَار cave	/fa:qq/ فَاق exceeded	/qa:ll/ قَال said (masculine)	/tˁa:rr/ طَار flew (masculine)
/tˁi:nn/ طِين mud	/sˁi:tt/ صِيت fame	/qi:ll/ قِيل was said	/si:xx/ سِيخ skewer	/ɣi:dd/ غِيد (soft) beauties	/bi:dˁdˁ/ بِيض White(plural)
/ku:xx/ كُوخ cottage	/tˁu:bb/ طُوب bricks	/su:qq/ سُوق market	/tˁu:ll/ طُول length	/ɣu:ll/ غُول ghoul	/tˁu:rr/ طُور phase/the mount

Table 2-6: Arabic stimulus material for plotting formants

The following observations can be made about MSA vowels based on the current work as illustrated in figure 2-9 and table 2-7.

1. Phonemically, MSA as used in the study has three short vowel phonemes /a/, /i/ and /u/ and three long phonemes /a:/, /i:/ and /u:/. It also has 6 phonetic short and long vowel categories, namely [aˁ], [iˁ], [uˁ]; [a:ˁ], [i:ˁ] and [u:ˁ] where ([ ]) denotes a phonetic category.
2. Short vowels except [uˁ] are more centralized than their long counterparts forming a triangle shape as in Classical Arabic (figure 2-9). Generally, short vowels tend to be lower than their long counterparts as follows except for [a] which is higher than [a:], and [uˁ] which is higher than both [u:ˁ] and [u]. Figure (2-9) shows values for a small sample of data only.

2.1. /i/ is found lower and more centralized than /i:/ with F1 values of 520(hz) to 458 (hz) respectively, similar to Syrian Arabic SA (Almbark, 2012) and the same for the phonetic categories of the short [iˁ] and its long counterpart [i:ˁ] with F1 values of 539(hz) to 491(hz) as in Table (2-7).

2.2. /a/ is higher than /a:/ with F1 values of 805: 878, in line with Almbark (2012) whereas the phonetic category of the long [a:ˁ] appears higher than its short counterpart [aˁ] with F1 values of 713(hz) to 767(hz) and the plain long counterpart /a:/ as in Table (2-7).

2.3. The short high back rounded vowel /u/ appears lower than /u:/ with F1 values of 490(hz) to 446(hz), similar to SA in Almbarak (2012), Cowell, (1964, p. 9), and Allatif's (2008) whereas the short emphatic variant [uˤ] was found higher than its long counterpart [u:ˤ] with F1 values of 433(hz): 475(hz) as in Table (2-7).

3. In terms of backness, the three vowel phonemes of Arabic /a/, /i/ and /u/ and their -/+ emphatic and -/+length counterparts were found different from each other in their F1 (height) and F2 (frontness/backness) values, similar to SA as in Cowell (1964, p. 9), Allatif's (2008) and Almbarak (2012).

3.1. The long vowel /a:/ has a backer vowel quality than its short plain counterpart /a/; F2 1820(hz) to 1652(hz). The two emphatic vowel variants [aˤ] and [a:ˤ] are both backer than /a/ and /a:/ with F2 values of 1454(hz) to 1189(hz) respectively (Almbark, 2008, p. 192; Khattab, Al-Tamimi, & Heselwood, 2006). Vowel /a:/ appears to be as a central vowel, being backer than /a/ but fronter than /aˤ/.

3.2. Vowel [u:ˤ] was the only long variant not found on the periphery of their short counterparts; otherwise the remaining long plain and emphatic /u/ variants (/a:/, /i:/, [a:ˤ], [i:ˤ] and /u:/) were plotted on the periphery. The short high back rounded vowel /u/ was found lower and more centralized than its longer counterpart. The long emphatic vowel [u:ˤ] was found fronter than [u:].

3.3. Both short plain and emphatic high front vowels /i/ and [iˤ] were found to be lower and more centralized than their long plain and long emphatic counterparts /i:/ and [i:ˤ] respectively.

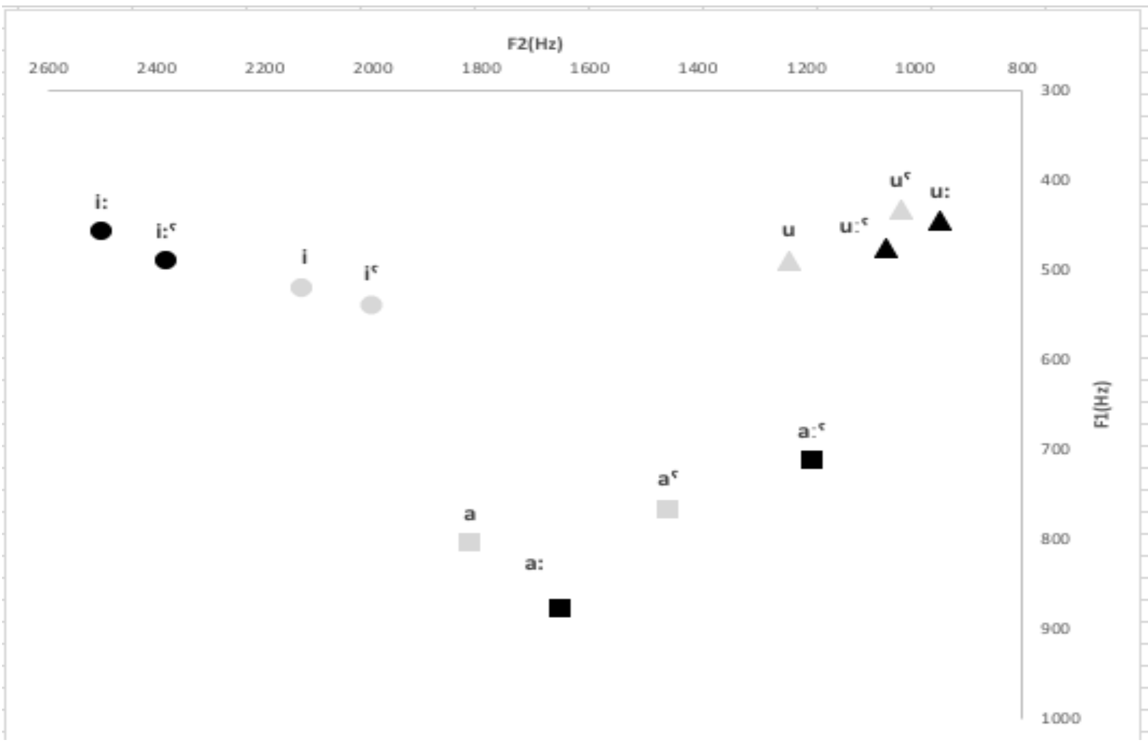


Figure 2-9: Mean of Arabic speaker's F1/F2 plot of plain and emphatic short and long Arabic vowels

vowel	AvgF2	AvgF1	$Time_1$	F1	F2	$Time_2$	F1	F2	$Time_3$	F1	F2	word
a	1820	805	0.2444	826.45	1738.1	0.2388	819.33	1708.2	0.214	792.66	1667.2	fakk
i	2130	520	0.2751	496.52	2195.9	0.2814	529.85	2124.7	0.3814	534.71	2070.2	zidd
u	1232	490	2.4376	491.82	1234.4	2.4414	489.97	1230.1	2.4531	489.42	1231.2	sull
aˤ	1454	767	1.8146	783.43	1391.6	1.8271	832.51	1360.4	1.8833	686.16	1609.8	qad
iˤ	2002	539	1.7813	537.61	2002.9	1.7681	534.31	1994.9	1.7624	545.28	2008.4	qiff
uˤ	1025	433	1.0124	456.26	1544.6	1.0186	332.63	1.0061	0.9999	511.9	1529.3	Tuff
a:	1652	878	3.5747	879.95	1643.3	3.5809	864.82	1640.8	3.5872	890.48	1672.2	saad
i:	2501	458	0.164	480.78	2463.6	0.1702	454.65	2489.2	0.1765	439.62	2549.6	riim
u:	953	446	1.1093	467.01	971.63	1.1343	436.88	921.33	1.153	434.55	965.16	duur
a:ˤ	1189	713	0.9736	712.81	1181.7	0.9798	713.77	1186.7	0.9861	711.74	1197.9	qaal
i:ˤ	2381	491	2.3812	485.25	2567.5	2.3875	478.05	2612.7	2.35	509.64	1963.1	qiil
u:ˤ	1053	475	0.366	468.36	1048.7	0.3847	458.45	1029.8	0.291	499.27	1081.8	Tuur

Table 2-7: Approximated F1 and F2 mean values of 1 Arabic speaker's (not normalized) production of Arabic short-long vowels in plain and emphatic environments

As for the consonant inventory of MSA, there are 28 consonants as reflected in table 2-8. Only guttural consonants and emphatics and their plain counterparts are described below due to their relevance to the current work.

	Place of articulation									
Manner of articulation	*Bilabial*	*Labio-dental*	*Dental*	*Alveolar*	*Post-alveolar*	*Palatal*	*Velar*	*Uvular*	*Pharyngeal*	*Glottal*
Plosives	b		t d tˤ dˤ				k	q		ʔ
Nasal	m			n						
Fricatives		f	θ ð ðˤ	s z sˤ	ʃ		x ɣ		ħ	h
Affricates					dʒ					
Trill				r						
Approximant						j	w		ʕ	
Lateral Approximant				l						

Table 2-8: Arabic Consonants inventory, adapted from the International Phonetic Association (1999, p. 49)

In MSA, *uvulars* /x/ (voiceless velar fricative; /**x**usuːf/ 'lunar eclipse'), /ɣ/ (voiced velar fricative; /**ɣ**ɑr/ 'cave'), /q/ (voiceless uvular stop; /**q**ɑmiːsˤ/ 'shirt'), *pharyngeals* /ħ/ (voiceless pharyngeal fricative; /**ħ**alaːl/ 'Halal'), /ʕ/ (voiced pharyngeal approximant /**ʕ**imlaːq/ 'giant') and *laryngeals* /ʔ/ (voiced glottal stop /**ʔ**sraːr/ 'secrets') and /h/ (voiceless glottal fricative /**h**arab/ 'escaped') pattern as a natural class[4] which in Semitic is known as the guttural class.[5] Four other consonants that have a secondary pharyngealization/uvularization articulation pattern with the true gutturals. These are the /sˤ/ (voiceless denti-alveolar fricative; /**sˤ**ɑbr/ 'patience'), /dˤ/ (voiced denti-alveolar stop; /**dˤ**ɑrbɑh/ 'stroke'), /tˤ/ (voiceless denti-alveolar; stop /**tˤ**iːb/ 'scent') and /ðˤ/ (voiceless dental fricative /**ðˤ**ɑlɑːm/ 'darkness'). The four emphatics /sˤ/, /dˤ/, /tˤ/ and /ðˤ/ are parallel to the plain /s/, /d/, /t/ and /ð/ respectively since pharyngealization/uvularization is a secondary feature.

Noteworthy is that some emphatics surface differently in some Arabic dialects. For instance, the /dˤ/ is realized as /ðˤ/ in most Gulf Arabic dialects (/dˤifdˤɑʕ/~[ðˤifðˤɑʕ]'frog')[6] but as /dˤ/ in the Bahaarna dialect in Bahrain (/dˤifdˤɑʕ/>[dˤifdˤɑʕ]'frog'), the /ðˤ/ either as /z/ in Egyptian, Libyan and Levantine dialects (/ðˤɑːlim/>[zˤɑːlim] 'oppressor') or /dˤ/ in the Bahaarna dialect (/ðˤɑːlim/~[dˤɑːlim] 'oppressor'). No Arabic dialect (to my knowledge) has retained the full set of MST emphatics. Some dialects also depharyngealize the emphatics, meaning that they are realized as plain coronals. For example, in Egyptian Arabic /sˤ/ surfaces as /s/ in some words such as /sʲaffaqa#lahu/>[saʔʔaf#lʊh] 'he clapped to/for him'.

[4] Crystal (2008, p. 323) states that "a set of segments is said to constitute a natural class if fewer phonetic features are needed to specify the set as a whole than to specify any one member of the set." In this sense, gutturals form a natural class since they lower surrounding vowels, tend not to co-occur within the same root, and tend not to occur in a syllable final position to mention but a few of their characteristics.

[5] To mention but a few works on the guttural class's naturalness cross linguistically are those of McCarthy (1994), Rose (1996) and Zawaydeh (1999). Some characteristics of the guttural class are 1) lowering surrounding vowels (Cowell, 1964; Herzallah, 1990; Rose, 1996), 2) historical mergers in the gutturals set (McCarthy, 1991) , 3) degemination (McCarthy, 1991) , 4) root consonant co-occurrence restrictions (Greenberg, 1950; McCarthy, 1991) , 5) cross-guttural vowel assimilation (McCarthy, 1991) and 6) avoidance of syllable final gutturals (McCarthy, 1994).

[6] The first form represents the underlying form of the word while the second indicates the surface form. Thus, the word frog in the Bahaarna dialect has the same form as the underlying form.

One final remark about the two MSA uvulars /x/~/χ/ and /ɣ/~/ʁ/ is that they sometimes surface as velar (/x/, /ɣ/) as in Palestinian Arabic and other Levantine dialects but as uvular (/χ/, /ʁ/) in some other dialects. McCarthy (1991) analyzes velar /x/ and /ɣ/ as being underlyingly uvular since they behave like uvulars, having a primary place Pharyngeal node which they lose during the derivation and surface as velars. I endorse McCarthy's view and follow his approach in this work.

2.2.4 Emphasis spread

In Arabic, a number of dialects including Palestinian Arabic, Abha dialect of Saudi Arabia, Cairene Arabic and Sanaani Arabic have been reported to exhibit emphasis spread or pharyngealization "tafxiim" where an underlying emphatic consonant spreads its features to neighbouring vowels (Davis, 1995; Herzallah, 1990; McCarthy, 1997; Shahin, 1997; Watson, 1999; Younes, 1991; Zawaydeh & de Jong, 2003).

Tafxiim was mentioned in the Middle Ages by Arabic grammarians and philologists who described emphasis to span over more than one segment (Barkat, 2006a, p. 669). Pharyngealized vowels (vowels under the effect of pharyngealization from a neighbouring emphatic consonants) have been described as being backed and lowered. From an articulatory point of view, emphatic consonants spread their backing effect up to three adjacent segments (ibid). Regarding the lowering effect, this is caused by the retraction of the tongue root (RTR) as a result of the co-articulatory effect of the constriction of the pharynx witnessed when emphatic consonants are produced in Arabic. (ibid)

Acoustically, pharyngealized vowels have been found to result in an increase on the F1 (open/close jaw; high/low) axis of the vowel diagram and a lowering of F2 (front/back position of the tongue) in a number of Arabic dialects (Al-Ani, 1970; Al-Ani & El-Dalee, 1983; Barkat-Defradas, Al-Tamimi, & Benkirane, 2003; Ghazali, 1983; Norlin, 1987; Younes, 1991; Zawaydeh, 1997). In other words, pharyngealized vowels are typically backer and lower than their counterparts (oral vowels), and the effect is most salient and consistent in the low vowel /a/.

The primary effect of emphasis is a change in quality of the vowels neighbouring emphatics. The short low vowel /a/ has been found to be acoustically more salient and to show more pharyngealization than /i/ and /u/ in Arabic (Albashir, 2008, pp. 66-71; Almbark, 2012; Gairdner, 1925). In the same line, front vowels (e.g., /i/) have been found to be more affected by pharyngealization than back ones such as /u/ (Albashir, 2008, pp. 66-71; Almbark, 2012, p. 65).

As for directionality and application of pharyngealization, it was found that leftward spreading of emphasis is more iterative (repetitive; unlimited) and greater than rightward spreading which is usually blocked by certain opaque segments (j, ʒ, ʃ, i), palatal vowels or consonants (Hellmuth, 2013). Furthermore, this spreading is greater from coronal emphatics than from pharyngeals. Some examples of emphasis spread are provided in (10) and (11) which come from a southern rural variety of Palestinian Arabic as cited in Davis (1995). In (10) and (11), capital letters stand for 'emphasis' where pharyngealization affects all consonants and vowels that are capitalized.

(10) Leftward unlimited pharyngealization (←) (Davis, 1995, p. 473-474)

Words displaying the leftward spread of emphasis

a. BALLAAṢ 'thief'
b. ħAðð̣ 'luck'
c. ʔABSAṬ 'simpler'
d. BAAṢ 'bus'
e. ʕAṬšaan 'thirsty'
f. MANAAFIð̣ 'ashtrays'
g. XAYYAAṬ 'tailor'
h. NAŠAAṬ 'energy'
i. TAMŠIIṬA 'hair styling'
j. MAJAṢṢAṢiš 'it didn't become solid'

(11) Rightward limited pharyngealization (→)

Words displaying the rightward spread of emphasis

a. ṢABAAḥ 'morning'
b. ʔAṬFAAL 'children'
c. ṬUUB-AK 'your blocks'
d. ṬWAAL 'long (pl.)'
e. Ṭiin-ak 'your mud'
f. ṢAyyaad 'hunter'
g. ʕAṬšaan 'thirsty'
h. ð̣Ajjaat type of noise (pl.)
i. ṢOOT-AK 'your voice'
j. ṢEEF-AK 'your sword'

As shown in (10) emphasis spread starts from the emphatic consonant and spreads leftwards to the beginning of the word without being blocked by any segment in between. On the other hand, rightward emphasis spread is not as consistent as the leftward one. For instance, in 11a-11d and 11i-11j emphasis spreads from the emphatic rightward without being limited. On the other hand, emphasis is blocked in 11e, 11f, 11g and 11h respectively by /i/, /j/, /ʃ/ and /y/.

2.2.5 'Imala

Another vowel harmony process affecting vowels is that of vowel fronting and raising known as 'imala in Arabic grammar books. 'imala was mentioned as a process that applied in Classical Arabic by Sibawayeh (c.760-796 ce) in his book *kitaab* and in the qiraa'aat of Quran, plural form of qiraa'ah in Arabic or recitation of Quran (Quranic variation). In some Arabic vernaculars, the vowels /a/ or /a:/ raise to /e, ɛ, ie, or æ/ and become more front within a word (medial position) or at the end of a feminine word (noun or adjective ending with -ah) in the presence of /i/ or /i:/ (Kaye, 1997; Levin, 1998; Owens, 2005; Torreblanca, 1994, p. 198). An example of this is the word /muʃkilah~ miʃikle/ in Syrian Arabic. Acoustically, 'imala causes F1 lowering and F2 raising as opposed to emphasis spread mentioned in 2.1.2.3 (Benkirane, 1981).

The presence of the umlauting /i or i:/ is not strictly a prerequisite for the application of 'imala since a number of Arabic spoken dialects have been reported to display 'imala in the absence of blocking pharyngealized consonants (Barkat, 2006b, p. 678). For instance, the examples in (12) come from Syrian Arabic where the instances to the left show vowel raising in contrast to the ones to the right which have gutturals (/h/, /ħ// and /rˤ/) immediately before the feminine suffix. The presence of the gutturals blocks vowel raising or in other words spreads emphasis effect to neighbouring vowels (uvularization/pharyngealization). The dialects showing imala in the absence of pharyngealized consonants can still show it in the presence (fikra~fikre 'idea', tazkira~tazkire 'ticket' in Syrian Arabic) or absence of i/i: (cf. 12a, 12b and 12c).

(12) Syrian Arabic feminine suffix -e/a (Cowell, 1964; Rose, 1996)

a. daraž-e /daraʒe/ 'step'	d. wa:ž**h**-a /waʒha/ 'display'
b. šerk-e /ʃerke/ 'society'	e. mni:**ḥ**-a /mni:ħa/ 'good'
c. madras-e /madrase/'school'	f. dagga:**R**-a /dagga:rˤa/'tanning'

Today 'imala is attested across Arabic dialects with variation in its degree, with some dialects ranging from medium i.e., /a/>[æ], [ɛ] or [e] to strong i.e., /a/> [e], [i] (Barkat, 2006b) or even [ie] as in naas>nies in Libyan Arabic (Owens, 2005). According to Barkat (1997) who studied four Arabic dialects, namely Lebanese, Syrian, Algerian and Moroccan Arabic, 'imala can be said to range in its degree between Western to Eastern dialects on a scale of non-existent (zero 'imala) to strong as shown in table 2-9. Maghrebi/Western dialects were found not to display 'imala whereas medium 'imala was found in most Syrian dialects and strong 'imala in the Lebanese dialect of Beirut and the dialect of Homs in Syria.

Type of dialect	Name of dialect	Cited location	Type of 'imala	Example
Western	Moroccan		Zero 'imala	NA
Western	Algerian		Zero 'imala	NA
Eastern	Syrian dialect	Horan /ħora:n/	Medium	[bjənsa:ha:~bense] 'he will forget her'
	Syrian dialect	Damascus	Medium	[samaka~samake] 'fish'
	Syrian dialect	Homs	Strong	[malika~mæliki] 'queen'
Eastern	Lebanese dialect	Oasis of Sukhne east of Palmyra	Strong	/ʃa:rib~ʃi:reb] 'moustache'
	Lebanese dialect	Beirut	Medium	

Table 2-9: 'imala across four Arabic dialects (Barkat, 1997)

Some sociolinguists also claim that ‘imala is influenced by some sociolinguistic factors such as gender, age, residence (urban vs. rural) among others. For instance, Barkat (2006b) classifies rural dialects in the Levant region as non-raising dialects as compared to the dominant (of higher status) urban raising dialect. Conversely, social domination in Egypt is for the non-raising dialect in Cairo (Woidich, 1994). Regarding the factors of gender and age, Kaye (1997) points out that women reflect more ‘imala than men whereas Al-Wer (1998) contends that vowel raising is a characteristic of young Syrian and Palestinian females in accordance with the dominant urban dialects in the Levant region.

Thus far, the previous sections offered background on the sound systems of Turkish and Arabic and front versus back vowel harmony in both languages. In the next subsection, I compare the ALT sound system to those of MST and MSA.

2.3 Arabic Loanwords into Turkish background

2.3.1 Sound system of ALT

The Arabic loanwords into Turkish (ALT) exhibit the same eight short vowels and twenty three consonants as in Turkish; however, ALT also displays long vowels (/iː/, /eː/, /uː/ and /aː/) in words borrowed from Arabic and Persian. The examples in (12) come from Göksel & Kerslake (2005).

(12) Long vowels in Arabic and Persian words into Turkish
/aː/ *matbaa* /matbaː/ 'press', *kira* /ciɾaː/ 'rent', *mavi* /maːvi/ 'blue', *arif* /aːɾif/ 'wise person'
/uː/ *mevzu* /mevzuː/ 'topic', *suret* /suːɾet/ 'copy', *buse* /buːse/ 'kiss', *Numan* /nuːman/ 'a name'
/iː/ *fiil* /fiːl/ 'verb', *ilan* /iːlan/ 'advertisement', *sine* /siːne/ 'bosom', *Didem* /diːdem/ 'a name'
/eː/ *teessüf* /teːssyf/ 'sorrow', *temin* /teːmin/ 'acquisition', *tesir* /teːsiɾ/ 'effect'

Long vowels borrowed into Turkish are either originally long such as *askerî* /asceɾiː/ 'military' and *ahlâk* /ahlaːk/ 'morals' (where the circumflex (^) on the vowel denotes a historical vowel) or became long through deleting /ʔ/, /ʕ/ or /ɣ/ and undergoing compensatory lengthening as in *maalesef* /maʕ ʔalʔasaf/>/maːlesef/ 'unfortunately', *şiir* /ʃiʕr/>/ʃiːɾ/ 'poetry' and *mağlup*

/maɣlu:b/>/maːlup/ 'defeated' (where ğ is a silent consonant that lengthens a preceding vowel). Nowadays, the use of the circumflex is vanishing from dictionaries, i.e., orthographically long vowels are written as short; however, auditorily they are still long. In addition, long vowels resulting from compensatory lengthening (compensating for the deletion of one of the three gutturals /ʔ/, /ʕ/ or /ɣ/) are reflected in the orthography as a sequence of two vowels. Compensatory lengthening takes place when one of the three gutturals /ʔ/, /ʕ/ or /ɣ/ occurs intervocalically, in a coda position followed by a tautosyllabic consonant or another consonant in the next syllable or across morpheme boundaries in the Arabic source words. Examples of this are /taʔasuf/>/teessyf/ 'sorrow', /fiʕl/>/fiil/ 'act' and /maɣa:ra/>/maara/ 'cave' (Göksel & Kerslake, 2005).

In the ALT, diphthongs are a combination of a vowel and the glide /j/ or /v/ which compares to the glide /w/ in original Arabic words. The examples in (13) below illustrate ALT diphthongal realizations.

(13) Diphthongs in ALT
/aj/> [aj], [ej] and [e]: /hajvan/ *hayvan* 'animal', /mejil/ *meyil* 'slope' and /dʒep/ *cep* 'pocket'.
/aw/>[av], [ev], [œ], [œv] and [y]: /tavsije/ *tavsiye* 'recommendation', /mevism/ *mevsim* 'season', /nœbet/ *nöbet* 'turn/shift', /tœvbet/ *töbet*'repenting' and /cyme/ *küme* 'heap'.

2.3.2 Historical mergers within the guttural class: sound change

Cross-linguistically, gutturals group as a natural class. This is supported by historical mergers/adaptation within this one set of sounds. For example, McCarthy (1991) reports the historical sound changes listed in (14).

(14) Historical neutralizations within the class of gutturals (McCarthy, 1991)
/ʁ/ ---> /ʕ/ (Hebrew, Aramaic, Maltese)
/χ/ ---> /ħ/ (Hebrew, Aramaic, Maltese)
/ħ/ ---> /h/ (Chad, Arabic, Socotri)
/ʕ/ ---> /ʔ/ (Chad, Yemenite, Anatolian Arabic, Socotri)

In ALT, the adaptation of these gutturals is illustrated in (**15**).

(15) Historical neutralization of Arabic gutturals into Turkish (Ar.= Arabic, Ott.= Ottoman language and MST= Modern Standard Turkish)

<u>Ar.</u> ---> <u>Ott.</u> --->	<u>MST</u>	
ʕ ---> ʔ --->	ʔ--->ø	(/sa:ʕah/ --->/saʔat/--->/saat/ 'hour/watch')
ʔ ---> ʔ --->	ø	(/maʔmu:r/ --->/meʔmûr/--->/memur/ 'official')
x ---> h ~	h	(/ta:ri:x/ ---> /târîh/--->/tarih 'history')
ɣ ---> g or ğ ~	g or ğ	(/ɣa:fil/ --->/gâfil/--->/gafil/ 'heedless' and /maɣfu:r/ --->/mâfûr/---> /mafur/ 'forgiven')
q[7] ---> k --->	k	(/baqqa:l/--->/bakkâl/--->/bakkal/'grocer')
ħ ---> h ~	h	(/ħa:l/--->/hâl/ --->/hal/ 'condition/state')
h ---> h ~	h	(/ʔaha:li:/--->/ʔahâlî/--->/ahali/ 'inhabitants')

When the Turks borrowed Arabic words, whether through Persian or directly from the different dialects they were exposed to, they reflected the guttural sounds in their orthography as evidenced in the Ottoman alphabet (Figure 2-10) which, just like Arabic and Persian, was written right-to-left and with Arabic characters. The Arabic guttural letters are highlighted in figure 2-10.

[7] Although in Ottoman Turkish, the uvular stop [q] was written as /q/, it was pronounced as a [k] and the velar stop /k/ was pronounced as a palatalized k, i.e. /c/. The orthographical q was dropped from Turkish in the 20th century.

پ	ب	آ	ا	ء
pe	be	medli elif	Elif	Hemze
ح	چ	ج	ث	ت
ha	çim	Cim	se	Te
ز	ر	ذ	د	خ
ze	re	Zel	dal	Hı
ض	ص	ش	س	ژ
dat	sad	Şin	sin	Je
ف	غ	ع	ظ	ط
fe	gayın	Ayın	zı	Tı
ل	ڭ	گ	ك	ق
lâm	sağır	Gef	kef	Kaf
لا	ه	و	ن	م
lâm elif	he	Vav	nun	Mim
	ی	ة		
	ye	Te		

Figure 2-10: The Ottoman Turkish Alphabet (Develi, 2004, p. 18)

Despite being written in Ottoman Turkish (Osmanli), the only guttural sounds pronounced by Osmanli users except for bilingual and trilingual Ottomans (who spoke Arabic and Persian in addition to Turkish) were /h/ and /ʔ/. Lewis (2000, p. 8) reports that the glottal stop had two functions; either to represent the original Arabic hamza /ʔ/ (*fi'l* /fiʔl/>/fiil/ 'act') or the Arabic voiced pharyngeal /ʕ/ (e.g., *şer'î* /ʃarʕi/>/ʃerʔi/ 'pertaining to the sacred law', Kur'an). He also states that the distinction between the two sounds was not maintained in intervocalic position (*müdafaa* /muda:faʕah/>/mydafa:/ 'defence', *teesüf* /taʔasuf/>/te:syf/ 'regret') and the distinction in spelling is not held in modern usage.

In the next part, a review of past studies on loan phonology is provided which highlights the current models and how they differ from each other.

2.4 Current models of loanword adaptation

2.4.1 Phonetics, Phonology or both?

Three theories have been proposed over the past two decades or so to explain the adaptation of loanwords cross linguistically. Proponents of the **Phonetics or Perception approach** (Silverman, 1992; Yip, 1993; Dupoux, Kakehi, Hirose, Pallier, & Mehler, 1999; Peperkamp & Dupoux, 2003; Kenstowicz, 2003 and Peperkamp, Vendelin, & Nakamura, 2008) contend that the initial stages of loanword adaptation take place in perception where the perceptual input (auditory or articulatory in nature) of the donor/source language is matched to parallel phonetic categories in the borrowing (recipient) grammar. The assumption according to this theory is that adaptation occurs during the perception and learning of the foreign word by naïve listeners. Conversely, advocates of the **Phonological model** (Jacobs & Gussenhoven, 2000; LaCharité & Paradis, 2005; Paradis, 1995) maintain that if the adaptation is performed by bilinguals who have the closest knowledge (percept) to native speakers, then the input is the underlying representation of the word in the source language from which they then create the surface form. If they use the phonology of the source language, the word then is pronounced similar to other source language words. However, if they use the phonology of the recipient language during the production of the word, then the word is adapted/matched to the recipient language's grammar and sounds closer to the recipient language. A medial approach is the **Phonetic-Phonological (hybrid) theory** adopted by Kenstowics and Suchato (2006), Smith (2006), Chang (2008) and Dolus (2013) among others. According to this model, the input to the recipient language is the source language output which can be either phonetic (perceptual cues) or phonological (feature combination) in essence and phonetic, phonological or grammar external factors such as orthography also determine the adaptation.

In the remainder of this opening section I sketch key features of early theories about the mechanisms underlying loanword phonology. Then, in separate subsections, I introduce each of the three contemporary models of loanword adaptation. Finally at the end of this section, I discuss other more general issues related to loan phonology and conclude by outlining the approach used in the current work.

In his seminal early work on English loanwords into Cantonese, Silverman (1992) proposes a multiple scansion (level) model where the adaptation process is divided into two separate levels; Perceptual Level and Operative level. The input to the Perceptual Level is a raw acoustic signal constrained by native segment and tonal constraints. According to this model, if a segment is not phonetically salient enough such as word-final obstruents (e.g. English 'warrant' /wɔ́.rənt/) and not part of the native language inventory, then it is deleted (Cantonese [wɔ.løn]). Otherwise, it is passed on to the second level. In the Operative Level, phonotactic constraints of the native language are applied on the segment coming from the Perceptual Level where ordered rules which are specific to loan phonology only are applied. For example, English 'bus' /bʌs/ with a word-final fricative /s/, being a salient segment, is perceived as [pa si] with the /s/ being retained and followed by an epenthetic /i/ under the phonotactic constraints of Cantonese.

Jacobs and Gussenhoven (2000) and Yip (1993) criticize Silverman's model (1992), especially the notion of loanwords-specific ordered rules. Yip (1993) still agrees with Silverman that loanword adaptation can be explained by two separate levels; Perceptual and Production Level. Yip advocates segments' perceptual saliency in the Perceptual Level; that less salient segments are not preserved. However, she dispenses with ordered rules. Instead she employs constraints ranking in Optimality Theory, and contends that loanwords can be accounted for by the same set of constraints used for the borrower's native language but using different rankings.

Recent research has abandoned the two separate perception and production levels. Instead recent theories of loanword phonology have taken one of three stances; i) that loanword adaptation happens in perception and can only be explained in terms of phonetic cues

(Perceptual/phonetics approach), ii) that the phonology of the source language only determines the adaptation based on proximity of distinctive features in the source and recipient language (Phonology approach) and iii) that both phonetic and phonological factors determine the borrowing (Phonetics-Phonology approach). These three approaches are elaborated below.

2.4.1.1 The Phonetic approach

An assumption underpinning the Phonetic Approach is that it considers perception as a stage of the adaptation process dependent on phonetic/perceptual (acoustic) similarity (Kenstowicz, 2007; Peperkamp & Dupoux, 2003; Silverman, 1992 and Yip, 1993) as perceived by naïve listeners (non-native speakers of the source language) who misperceive the incoming foreign word. LaCharité and Paradis (2005) who are not proponents of the Perceptual stance refer to this assumption as *Phonetic approximation* in their characterization of the perceptual stance to show that it does not work. This notion proceeds such that the adaptation of the foreign segments is based on how the outputs (phonetic surface forms) of the source language and recipient language are phonetically similar to each other. In the words of Peperkamp et al. (2008) "proximity in the sense of fine-grained articulatory gestures" while other supporters of the Phonetic model assume proximity in terms of acoustic features (e.g. Kenstowicz and Mou, 2009). This theory downplays the role of bilinguals and phonological factors in the adaptation. For instance, Silverman (1992, p.296) maintains that "Many Cantonese speakers who employ English loanwords possess a good command of both spoken and written English" yet they (according to Silverman) are constrained by their native language and are expected to "represent and produce the native segment which most closely approximates the input in articulatory and/or acoustic properties" (ibid, p.296). The Phonetic approach has been criticized for downplaying the role of bilinguals as the agents of borrowing; their role is especially evident in communities where bilinguals constitute the majority of the population such as Montreal where both Quebec French and English are used (LaCharité and Paradis, 2005).

Phonetic approximation was widely used in the loanword literature by the proponents of the Phonetic approach such as Silverman (1992), Yip (1993), Kenstowicz (2001) and Peperkamp and Dupoux (2002, 2003). They maintain that *phonetic approximation* can entail a number of predictions including *phoneme non-perception*, i.e. phoneme deletion and *incorrect phoneme categorization/mismatching*, i.e., adaptation (after Silverman, 1992). The supporters of the Phonetic approach attribute phoneme deletion and mismatching to perceptual 'deafness' (Peperkamp and Dupoux, 2003). That is, phonetic approximation entails mapping of non-native segments onto the phonetically closest categories in the native language. During this process, some segments might be deleted due to the listeners being 'deaf' to certain segments or structures which are not present in their native inventory or due to 'deforming' of sounds resulting in adaptation.

An example of adaptation by phonetic approximation is the adaptation of English VN rhymes into Mandarin Chinese (MC) reported by Hsieh, Kenstowicz and Mou (2009). In MC, nasal coda consonants are phonologically contrastive for their place of articulation between coronal vs. dorsal nasal consonants. In Hsieh, Kenstowicz and Mou (2009), the adaptation of a word-final nasal in an English word depends on the F2 [backness] value of the vowel in the English word (a salient phonetic cue) rather than the place of articulation of the nasal consonant (a contrastive phonological feature). Thus, English back vowels trigger adaptation of any nasal to /ŋ/ (e.g. *Congo* $[ɑŋ]_E$>gang.guo $[ɑŋ]_{MC}$)[8], front vowels yield /n/ (e.g. *clan* $[æn]_E$>ke.lan $[an]_{MC}$), and central vowels (e.g. *punch* $[ʌn]_E$> pan.qu $[an]_{MC}$ and *young* $[ʌŋ]_E$>yang $[aŋ]_{MC}$) trigger faithful mapping of the English nasal consonant. This example from MC clearly manifests that adaptation is largely dependent on the acoustic quality of the source language vowel which is more salient than the phonemic nasals in MC.

Adopting an extreme version of the Phonetic approach, Peperkamp and Dupoux (2003) argue that "all loanword adaptations are phonetically minimal transformations that apply during speech perception"(p. 342). This position is also held by Peperkamp et al. (2008) yet with two slight modifications. They view "all loanword adaptations that do not represent generalisations to a default pattern and that are not influenced by orthography to originate in perceptual assimilation" (p. 160). The 'default pattern' which they exclude refers to

[8] All data come from Hsieh, F., Kenstowicz, M., & Mou, X. (Eds.). (2009). Mandarin adaptations of coda nasals in English loanwords (Vol. 307): John Benjamins Publishing company.

overgeneralizations in the borrowing language, which means that their definition allows for the phonology of the native language to explain loanword adaptation at least in part. Likewise, they admit that orthography too can play a role in the nativization process. On the other hand, their definition excludes any influence by the phonology of the source language.

The case that Peperkamp et al (2008) study is the adaptation of English and French word final <n> by Japanese speakers who adapt nasal word-final <n> in English loanwords as a moraic nasal consonant (e.g. English /pe**n**/ 'pen'>[pe**n**]) but in French loans as an epenthetic vowel /ɯ/ (e.g. French /kan/ 'Cannes'> [kann**u**]). They conduct two experiments, one on monolinguals and then a second experiment on bilinguals. In the first one, the monolingual speakers listen to stimuli of the shape CVCVCVN produced by American and French speakers and are then instructed in an identification task to choose the closest non-word they hear. The researchers maintain that they controlled for the effects of phonology and orthography by i) recruiting naïve Japanese listeners with as minimal knowledge of English and French as possible and ii) not informing the participants about the nature of the stimuli. The recorded stimuli by the French speakers showed that both male and female releases from the nasal /n/ had formants whereas in English only female speakers' recordings exhibited release with formants but not the male speakers. This confirms that the difference between English and French words with word-final nasal /n/as perceived by Japanese speakers lies in perception.

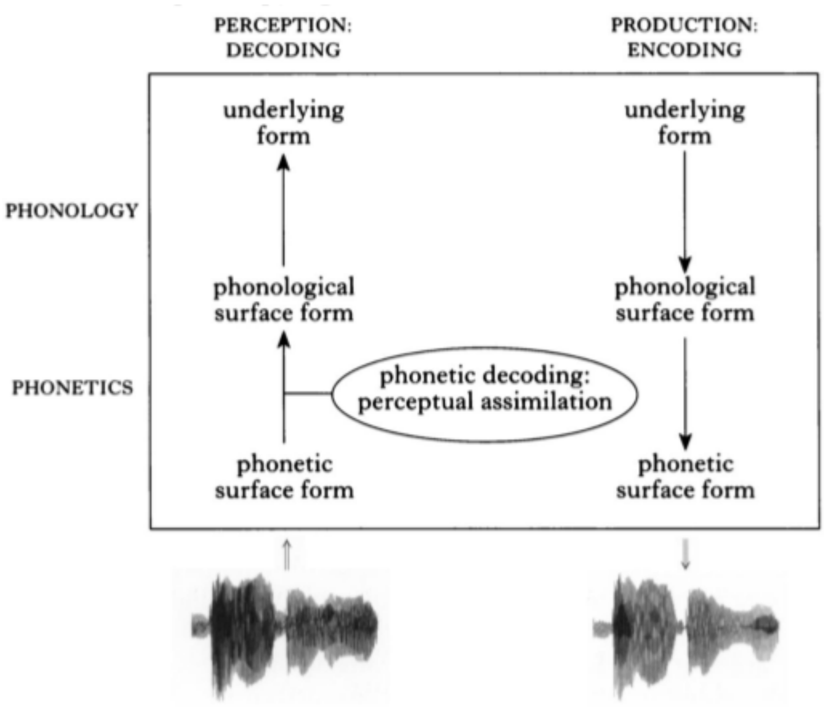
One could argue here that the authors downplayed any potential effect of the source language phonology and did not even entertain it in the design of the first experiment. That is, they could have included bilinguals in this first experiment already, to check whether they would yield different responses to those of the monolingual group.

Nonetheless, the authors gauged the effect of the bilingual group in the second experiment in order to confirm that perceptual assimilation was responsible for the epenthesis of the vowel after the nasal in French loanwords. They also recruited monolingual French participants as a control group. They used a discrimination task of the shape ABX where A and B are always different words while X could be identical to either A or B. This time the stimuli included tokens of the shapes CVCVCVn, CVCVCVnɯ and CVCVCVni (/i/ used as a control vowel) which were recorded by a French man and a woman who were instructed to read the words aloud, placing stress on the penultimate syllable (before /n/) in vowel-final words. The tokens were then modified such that the word endings [n], [nɯ] and [ni] were spliced onto the same CVCVCV base.

The hypothesis in this experiment is that Japanese speakers will have more difficulty in discriminating the words ending with /n/ and with /ɯ/ than those ending with /i/. The results confirm the hypothesis and reveal that it is more difficult for the Japanese speakers to perceive the discriminated words ending with n and those with /ɯ/. The authors interpret the results such that the release from words ending in n is closer to /ɯ/ than /i/ and as such the adaptation is born in perceptual assimilation.

One caveat in the design of the second experiment may lie in the authors' interpretation of bilingualism. Clearly they are adopting a definition of 'bilingual' which includes late bilinguals, whom they label as proficient; this is despite reporting that the bilingual participants, who are aged between 21 and 37, resided in Paris for an average of only three years during which they studied French in a language center or at a university. Furthermore, the participants scored their own language competence as a little over the average (6.3 in comprehension, 5.9 for production and 5.9 pronunciation) which suggests that the participants should perhaps be labeled as second language learners and not proficient bilinguals per se.

In sum, Peperkamp et al. (2008) argue that loanword adaptation takes place in perceptual assimilation by which both native words and loanwords are mapped to their phonetically closest counterparts during speech perception (1994); i.e., based on acoustic distance (Kuhl, 2000) or articulatory gestures (Best & Strange, 1992). Figure 2-11 depicts the place of perceptual assimilation in phonetic decoding, in the model developed by Peperkamp et al. (2008), as it applies to bilinguals and monolinguals in which "continuous universal phonetic representations into discrete language-specific ones" (p. 137) exist.



A speech-sound processing model. The processes indicated by single arrows are language specific; those marked by double arrows are universal.

Figure 2-11: Speech sound processing model by Peperkamp et al. (2008)

According to this model, a two-way distinction can be made; one between perception and production and the other between phonetics and phonology. Crucially, one of the assumptions of this model is the primacy of perception over production; the role of production is evident either after the operation of perceptual assimilation, in order to deal with adjustments, or when perceptual assimilation does not take place. In the latter case, loanwords which are of the exact shape in the source language are produced. Moreover, the input to perception as shown in the model is always of a phonetic nature which leaves no place for phonological or orthographic forms although Peperkamp address orthographic effects in other studies (Vendelin and Peperkamp, 2006). Finally, the role of phonology in the model is confined to supplying well-formed native (Phonological) categories to which the non-native forms are mapped.

Another group of studies in the Perceptual approach considers phonology as part of the native language perception by which loanwords are adapted (Boersma & Hamann, 2009). In their article entitled “Loanword adaptation as first-language phonological perception”, Boersma and Hamann (2009) argue that phonology is a process of the native language perception and is the locus of loanwords adaptation. They use a single model for both native language processing and loanwords adaptation as in figure 2-12 to explain vowel insertion in English loanwords into Korean but not in native words. Their analysis is couched within Optimality Theory (1993, 2004), in which the competing forces at work in comprehension and production are modelled in terms of constraints and their interactions. Their proposal is that the anomaly, of epenthesis in loanwords but not in structurally parallel native words, is solved by considering that adaptations can be accounted for by interaction of structural (STRUCT) and faithfulness (FAITH) constraints during the comprehension stage, i.e., in perception. In contrast there is no epenthesis in the native words as these are modelled through interaction of structural and faithfulness constraints during production.

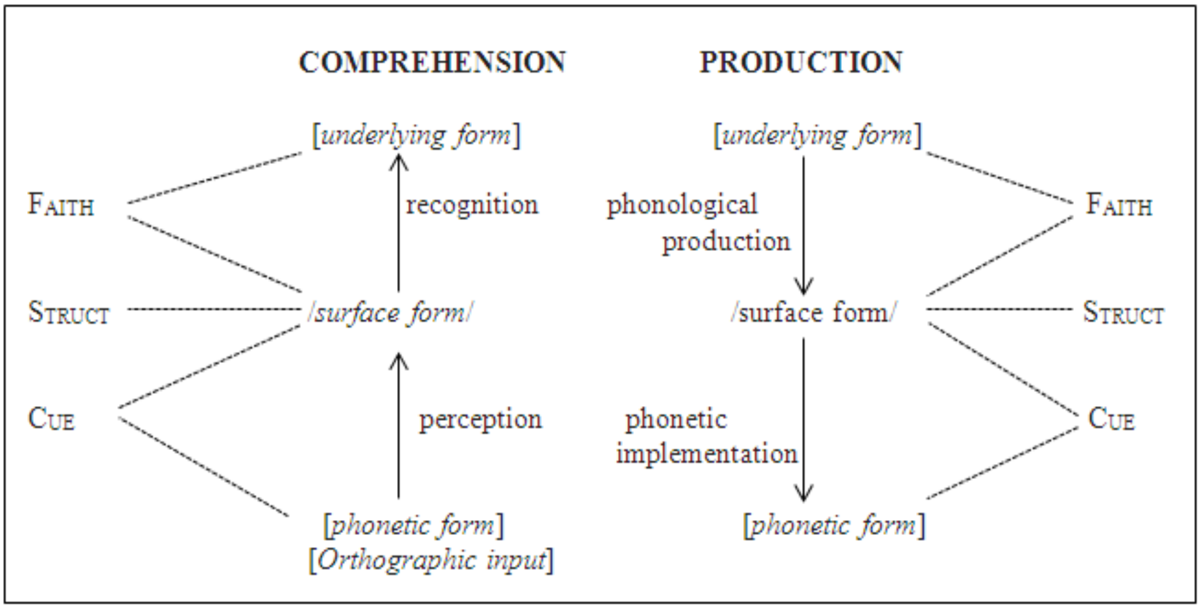


Figure 2-12: A single model for native language processing as well as loanwords adaptation

In the standard version of Optimality Theory (1993, 2004), the notion of input to output mapping correlates with that of underlying representation (UR) to surface representation (SR); i.e. input-output= UR-SR. Nevertheless, many works, like McCarthy (2011) among others, note that such a derivation is problematic for an analysis based on perceptually grounded faithfulness, since the underlying representation lacks perceptual/phonetic information that is important for comprehension/perception. Boersma and Hamann's (2009) model in figure 2-12 addresses this distinction by saying that the input during the comprehension stage is of a phonetic (auditory or articulatory) nature and supplies information to the surface structure, which is then relegated to the underlying representation. In the production stage , the input would be the underlying representation along with all the information stored from earlier iterations of the comprehension/perception stage; i.e., the input again has perceptual cues along with phonological information and possibly also some grammar-external information (orthographic). While the model of Boersma and Hamann (2009) looks comprehensive, it does not allow for any role of the phonology of the source language as a determining factor in loanwords adaptation. The role of the source language phonology is discussed in the next subsection as part of the Phonological approach.

2.4.1.2 The Phonological approach

Unlike the Phonetic approach, the Phonological approach entails that adaptation of borrowed words is based on phonological distance. That is, the source language word is mapped onto its phonologically closest equivalent in the borrowing language in terms of distinctive features (feature combinations) rather than phonetic one. Advocates of this approach include La Charité and Paradis (Paradis, 1996; Paradis and LaCharité, 1997 and LaCharité and Paradis, 2005) who base their claims on findings from 12 large corpora of English and French loanwords into different languages in the Project CoPho (2005).

In their 2005 article, La Charité and Paradis conclude that *phonetic approximation* accounts for a limited number of cases of the Project CoPho. Instead they establish that most of the cases can be explained by *category preservation* and *category proximity* as defined below.

> **Category preservation:**
>
> If a given L2 phonological category (i.e., feature combination) exists in L1, this L2 category will be preserved in L1 in spite of phonetic differences. (LaCharité and Paradis, 2005, p.226)
>
> **Category proximity:**
>
> If a given L2 phonological category (i.e., feature combination) does not exist in L1, this L2 category will be replaced by the closest phonological category in L1, even if the L1 inventory contains acoustically closer sounds. (LaCharité and Paradis, 2005, p.227)

An example of the Phonological approach cited in LaCharité and Paradis (2005) is the adaptation of English voiced stops in Spanish. Phonetically, English voiced stops and Spanish voiceless stops overlap in their correlates of Voice Onset Time (VOT) in the range 0-30 ms. This would predict that English voiced stops in loanwords would be perceived as voiceless by Spanish speakers learning English. However, La Charité and Paradis (2005) found from their CoPho data that i) none of their Mexican Spanish (MS) speakers in one corpus (MS1) devoiced the English stop (0/566), ii) only two speakers devoiced the stop in MS2 out of a total of 802 and iii) only three deleted the voiced stop. The data in Figure 2-13 come from LaCharité and Paradis (2005, p.252).

		English		Spanish	
/b/	bar	[b̲ɑɹ]	→	[b̲aɾ]	*[p̲aɾ]
	baseball	[b̲esb̲ɑl]	→	[b̲esb̲ɔl]	*[p̲esp̲ɔl]
/d/	dip	[d̲ɪp]	→	[d̲ip]	*[t̲ip]
	darling	[d̲ɑɹlɪŋ]	→	[d̲aɾlin]	*[t̲aɾlin]
/g/	golf	[g̲ɑlf]	→	[g̲ɔlf]	*[k̲ɔlf]
	gang	[g̲æŋ]	→	[g̲aɲ]	*[k̲aɲ]

Figure 2-13: Examples of unchanged voiced stops in English loanwords in Mexican Spanish (LaCharité and Paradis, 2005)

The data above (figure 2-13) clearly show that Mexican Spanish speakers in the CoPho project borrowed English voiced stops as voiced rather than voiceless in the loanwords, preserving the phonological feature [voice]. This sustains category preservation rather than category proximity.

2.4.1.3 The Phonetic-Phonological approach

As mentioned at the beginning of section 2.4.1., in addition to the Phonetic Approach and the Phonological Approach there exists a third *hybrid* approach of both phonetics and phonology in addition to other factors such as orthography. Researchers in this third approach include Smith (2006), Kenstowicz and Suchato (2006), Chang (2008) and Dolus (2013). According to this approach, a combination of both phonetic and phonological factors influence loanword adaptation side by side.

One case study of the *hybrid* approach is by Smith (2006) who argues that perception by itself is not sufficient to account for loanword adaptaion. She studies English loan doublets in Japanese where English words are borrowed in two different ways in Japanese, one showing epenthesis and the other deletion. She acknowledges the role of native language phonology in perceptual assimilation, and thus in loan phonology, but establishes that other factors including the phonology of the source language and orthography are also needed. In other words, a hybrid model of phonetics, phonology of both source language and native language as well as orthography can explain loanword adaptation. What is interesting about Smith's position is that it explicitly takes account of the phonology of the source language as well as that of the native language. This is expressed in her analysis in terms of Output-Output faithfulness (OO-FAITH) constraints between the source language and the native language.

A second example of a hybrid approach study is the one by Kenstowicz and Suchato (2006) where they investigate the adaptation of English loanwords into Thai from a corpus of 800 words and again analyze their data within Optimality Theory (1993, 2004). They reach four specific conclusions pertaining to the Thai data. One, they establish that auditory similarity but not phonological proximity may explain some cases where a mismatch occurs in the mapping of consonants (e.g., ʃ>ch and v>w). However, in other cases the mismatch can be explained articulatorily as in the mapping of the interdentals as dentals but not labio-dentals

Two, they deduce that when there is a phonetic distinction (e.g., voiceless-voiced stops) in the source language (English), the more perceptually salient category in the contrast becomes the norm and will be more likely to influence adaptation. As an example, word-initial voiceless English stops are systematically adapted as aspirated stops in Thai which suggests that adaptation is influenced by the phonetic details of the source language. However, there is no parallel pattern with English word-initial voiced stops.

Three, they make use of output-output alignment constraints (phonological constraints) to explain two competing repair strategies of truncation vs. epenthesis for illicit clusters in prosodic structures. These repair strategies are needed because, in Thai, the final syllable has to be heavy and is the site of the primary stress. Kenstowicz and Suchato also argue that these output-output constraints would account for cases when novel consonants are imported into Thai.

Finally, Kenstowicz and Suchato tested tone assignment in Thai via a phonetic experiment. The native Thai rule states that "syllables terminating in a sonorant take M [mid tone] and syllables terminating in an obstruent take H [high tone]. It was shown that the latter [H tone] rule is sensitive to a covert obstruent that is not realized in the loan" (Kenstowicz & Suchato, 2006, p. 27). The main conclusion of the article is that neither the Perceptual approach nor the Phonological approach alone can explain the different patterns in the corpus they examined. Rather a model that stipulates both phonetic and phonological constraints can account for them.

A third hybrid approach article is by Dohlus (2013) who investigates how German and French mid front rounded vowels (/œ/ and /ø/) are adapted differently into Japanese. Dohlus argues that German /œ/ and /ø/ are adapted as Japanese /e/, and that this adaptation is phonologically grounded; in contrast, French /œ/ and /ø/ are adapted as Japanese /u/, which is phonetically grounded. Dohlus contends that German and French borrowings entered Japanese differently leading to different adaptation forms. German words in Japanese date back to the end of the 19th century and were mainly borrowed as written forms (through scientific studies). On the other hand, French words infiltrated Japanese as everyday words (fashion, dancing, military, arts, French cuisine) in their oral word forms. Furthermore, Dohlus maintains that the difference in German and French spelling of the mid front rounded vowels may be the cause of the different adaptation of the phonemes into Japanese. The study by Dohlus establishes that this asymmetry, along with other examples of loanwords in Japanese, can be explained in both phonetic and phonological terms. Furthermore, she concludes that other factors are also at play such as orthography (faithful mapping of the source forms/phonemes), conventions (following grapheme-to-phoneme correspondence; thus causing phonological approximation) and knowledge of the source language (at a phonological level through classroom instruction).

One of the conclusions of Dohlus (2013) is that loan phonology is intrinsically phonetic in nature, as the source language phonemes are mapped onto their closest phonetic categories in the borrowing language. However, she adds that: "If there is a lack of oral input or the possibility of 'faithful perception' due to the presence of written form, then this triggers phonological approximation" (p.131). The role of orthographic factors is discussed next.

2.4.2 Role of orthography

So far the above review included previous studies that promote either the Phonological or Perceptual approaches or both. Other studies have shown that extragrammatical factors also, such as orthography, can play a role in loanwords adaptation. Some of these prominent studies include Detey & Nespoulous, 2008; Kaneko, 2006; Smith, 2006; Vendelin & Peperkamp, 2006.

Vendelin and Peperkamp (2006) maintain that the role of orthography has been downplayed in the literature due to the confused effect of source language sounds on native language perception (e.g. French learners perception of English sounds). This is especially true of the perception of adaptations emerging as a result of grapheme-to-phoneme correspondence, where the perception can be attributed to proximity either to 'phonetic or phonological minimality' (p. 1004). In this context, the authors distinguish between two possible effects on adaptations based on orthography; i) reading adaptation (reading source language words with native language pronunciation) and ii) grapheme-to-phoneme correspondence rules. However, they also argue that sometimes the adaptations can be identical in form, thus causing difficulty to tease apart the source of the adaptation.

Reading adaptation entails reading source language words with native language pronunciation, such as reading the adapted French word cul-de-sac /kytsak/ in English as /kʌldəsæk/, in accordance with the native language (English) phonology/grammar (Vendelin & Peperkamp, 2006). Grapheme-to-phoneme adaptation involves reading the source language word in accordance with how their native language has standardized the pronunciation of the source language graphemes. The grapheme-to-phoneme correspondence examples that Vendelin and Peperkamp (2006) cite are pronouncing <u> in *but* by French learners of English reading English as /œ/ not /ʌ/, and <oo> in book as /u/ not /ʊ/.

Thus, the authors tested the role of orthography in a production experiment that involved French-English bilinguals adapting on-line English nonce words of the monosyllabic shape CVC (/fVp/, /mvb/, /pvd/). They ran the experiment in two conditions: oral and oral-written (mixed) where half of the participants performed the first condition while the other undertook the other condition and vice versa. In the oral condition, participants first listened once to American English non-words of the shape CVC, then they were presented with a French carrier sentence in which they had to insert the non-word. In the oral-written condition, the participants first saw the written English nonce word on the screen once, then after 700 ms they listened to the English non-word and then saw the French carrier sentence in which they were instructed to insert the word they borrowed.

The hypothesis the authors tested was that grapheme-to-phoneme correspondence was the expected strategy the participants would use, with any other strategy as unexpected. The results of this experiment showed that in both conditions, the grapheme-to-phoneme strategy occurred more than reading adaptation. In addition, grapheme-to-phoneme strategy occurred more in the mixed oral-written condition than in the audio only condition. The authors conclude that adaptations yielded by bilingual speakers are influenced by orthography.

Thus, the findings of the previous studies on loanwords adaptation show that perception, production and orthography all contribute to our understanding of loanword adaptation. As Calabrese puts it in the introduction of the book Loan Phonology:

> a bilingual borrower first produces the word in L2 and then uses the surface representation as input to the nativization process, which is phonological. If this is correct, the perceptual stance and phonological stance models no longer need be contrasted,

> and could be largely unified: the input to nativization is always phonetic, the word as it is “heard”. The treatment, on the other hand, is always phonological and it can occur either during perception or during production. (Calabrese and Wetzels, 2009, p.9)

In the next subsection, we draw out one final key question related to loanword phonology, regarding the assumed agents in the borrowing process.

2.4.3 The agents of loanword adaptation: who does the borrowing?

One further important question in the literature regarding loanword adaptation, which still stirs controversy, is the following. Who performs the adaptation of loanwords?

A widely held assumption about who introduces loanwords is that the active borrowers are bilingual speakers who have access to both the source and target language. Advocates of the Phonological model who hold this view include LaCharité & Paradis, 2005, LaCharité & Paradis, 2000; Paradis & LaCharité, 1997, 2001a; Paradis & Prunet, 2000 among others. This postulation can be traced to as early as 1880 by Hermann Paul who maintained that

> “all borrowing by one language from another is predicated on some minimum of bilingual mastery of the two languages. For any large-scale borrowing a considerable group of bilinguals has to be assumed. Also, the more bilinguals there are in a community, the more borrowing will occur. The analysis of borrowing must therefore begin with analysis of the behaviour of bilingual speakers.” (cited in Haugen, 1950, p. 210)

Hence according to this view, bilinguals are responsible for introducing new words into their communities and then through communication with monolinguals, the new word is nativized and spread (Paradis & LaCharité, 2001b). Along these lines, regarding the mode of adaptation, it has been contended (Heffernan, 2007) that if the bilinguals are of high proficiency, they would tend to input phonological representations rather than phonetic cues. In this respect, bilingual speakers make what LaCharité and Paradis (2005) label as “*intentional phonetic approximation*” where they import a new word into the recipient language with its proper source language shape, meaning they are using the surface form of the word rather than its underlying representation. However, if the borrowing is performed by monolingual speakers who are not familiar with the source language grammar and structure, then the resulting words would be adaptations. LaCharité and Paradis (2005, p.231) refer to this as *"naïve phonetic approximation"* .

A competing view on the nature of the input, however, is that bilinguals access the underlying representation (UR) of the source language and then derive surface forms from it. The content of the UR is phonemic/phonological (cf. distinctive features) but not phonetic/allophonic (since allophones are surface representations) in nature, and repair strategies are later applied to the UR to render well-formed structures in the recipient language (LaCharité & Paradis, 2005). One example of this is the adaptation of stops in English loanwords into Mandarin Chinese (MC) as discussed in Paradis and Tremblay (2009). Phonetically, English has aspirated and non-aspirated stops; MC has phonemic/contrastive aspirated stops which would be predicted to facilitate phoneme categorization from English to MC. However, MC speakers categorize aspirated and unaspirated voiceless English stops as aspirated stops in MC whereas they map aspirated and unaspirated voiced English stops as unaspirated. This suggests that despite the phonemic distinction of aspiration in MC, speakers are not influenced by the phonetic features in English; this supports the Phonological model over the perceptual one.

On the other end of the spectrum, it is argued in psycholinguistic studies (Peperkamp et al., 2008, p. 341) that monolingual speakers might not necessarily have heard the input word (when it was first borrowed) yet they use it, which means that for them the underlying and surface representations of the source word are the same. Hence, according to Peperkamp (2004, p. 345), "there is no reason to keep the corresponding forms in the source language as the underlying forms in the lexicon of the borrowing language". That is, reference in loanword adaptation should not be made to phonemes (corresponding forms and underlying representations). Such loanwords should be analysed, in Peperkamp's view, in the same manner as native words, without having to assume any loan-specific constraints that are different than those for the native words.

Finally, a notion embraced by the proponents of the Phonological approach and the hybrid approach of phonetics and phonology, especially within constraint-based frameworks (Jacobs & Gussenhoven, 2000; Kenstowicz, 2005; Paradis, 1995; Yip, 1993), is that loanword adaptation yields unique insights into the grammar/phonology of the native language which it would not be possible to see solely through analysis of the native language. For example, the correct ranking of faithfulness constraints in the native language is more easily understood by exploring the rankings used for the loanwords. It is suggested that loanwords function as a wug test (Kang, 2011), which abstracts away from the form/realization of the word itself and reveals the adaptation strategies employed by learners/speakers which, in turn, reflects their native language phonology. This supports the use of nonce words in the design of perceptual studies, for example.

2.5 Summary

To sum up, this chapter provided theoretical background on the languages involved in the Arabic loanword adaptation in Turkish. This included the historical development of Turkish, Arabic and Arabic loanwords in Turkish; the genetic affiliation of Turkish and Arabic and where they are spoken in addition to background on the front-back vowel harmony in the two languages. This was followed by a review of past studies of loan phonology that helped in shaping the research questions and the methodology of the perceptual study used in this thesis.

2.5.1 Approach used in the current work

The current work has benefited from the body of literature on the topic of loanword phonology in a number of ways, starting with formalizing the research questions and methodology and ending with the approach used in the discussion of the results. This is elaborated as follows.

RQ1 on how current day Turkish speakers categorize the three Arabic vowels in emphatic/non-emphatic and short/long settings in nonce words in the PAT chapter (chapter 4) was based on the studies on perceptual assimilation, for example Peperkamp et al. (2008). The selection of the participants in the perceptual study of this thesis to include monolingual, bilingual speakers and Quranic speakers of Turkish (TQ) was specifically inspired by Peperkamp et al. (2008). As noted in section 2.4.1.1., in their first experiment, the authors selected only monolingual speakers in order to control for the effect of phonology. However, in the design of the experiment presented here in chapter 4, monolingual, bilingual and also Quranic speakers of Turkish (TQ) were recruited so as to check whether the phonology of the source language determines the expected categories. The addition of the TQ group also serves to test whether the degree of bilingualism influences the match of listeners' categorizations to the patterns observed in the corpus (that is, in the Turkish lexicon).

RQ2 tests whether speakers of Turkish would display categorizations in a perception study using real words which reflect both the patterns observed in the corpus and in the PAT data (which used nonce words, as discussed in the previous paragraph). RQ2 was motivated by the experimental design adopted in Vendelin and Peperkamp (2006) who used as stimuli a mix of non-words, low frequency English words borrowed in French and French words. RQ3 on the role of orthography in influencing the quality of the vowel neighbouring guttural consonants was also inspired by the experiment design of Vendelin and Peperkamp (2006). In their design, Vendelin and Peperkamp present their stimuli in two conditions: audio and audio-written; in this work a third condition of written-only is added.

Last but not the least, the debate over what constitutes the input to the adaptation process, and in particular whether it is phonetic or phonological (surface or underlying representation), sparked the idea of using perceptual studies designed to model the Ottomans' grammar (to which we do not have direct access anymore). Thus, we would treat Arabic words as inputs and not underlying forms, and the ALT (in the Turkish lexicon) as the output forms, and model the input-output correspondence experimentally. The choice of modern Turkish in this modelling scheme is due to the similarities between the two languages phonetically and phonologically (Turkish being the closest to Osmanlica in grammar).

The expectation espoused at the start of this work, and subsequently adopted in full in the discussion chapter (chapter 7), is that loanword adaptation is a complex process which relies on both phonetics and phonology side by side, in addition to other factors such as orthography, native language phonology and morphology. This approach was inspired by the works of LaCharité and Paradis (2008) and Dolus (2013).

In the next chapter, the ALT corpus data are presented along with the results of a stratification task that was carried out to determine the etymology of the corpus words.

3 Arabic loanwords into Turkish corpus data

3.1 Introduction

In this chapter the Arabic loanwords into Turkish corpus data are presented. The earliest word in the corpus dates back to the 11th century (1070) whereas the latest words can be traced back to the 1930, two years before establishing the Turkish Language Society (TDK). The chapter is divided into four main sections. Section 3.2. describes the research methodology used in collecting the data of the current work. Section 3.3. presents the main phonological patterns found in the corpus data including adaptation of long vowels into short ones and the residual effects of guttural sounds. Section 3.4. reports on the results of a stratification task to determine the etymology of the corpus words. Section 3.5. concludes with the main findings of the chapter and presents the rationale for the perceptual study discussed in chapters four, five and six.

3.2 Methodology

The data used in the current research comprises a corpus of 1,118 Arabic written loanwords in Turkish collected by the researcher who is a native speaker of Arabic and a student of Turkish over a span of two years. These are taken from Turkish sources including textbooks, TV shows, songs, hardcover and on-line dictionaries (Arabic-Turkish, Turkish-Arabic, Turkish-Turkish and Turkish-English). One important dictionary to be pointed out is the one in Turkish (Turkish-Turkish) and published by the Türk Dil Kurumu TDK [The Turkish Language Association] (2005), the same body responsible for eliminating a large number of Arabic and Persian words from Turkish. After collecting the data, the researcher entered them into excel spread sheets and divided them into columns including the Turkish spelling which is largely phonetic in nature (i.e., orthographic letters match IPA sounds), IPA transcription of both Turkish and Arabic pronunciation, Arabic and English glosses, etymological remarks and sources of each word. An illustration of the corpus used is provided in figure 3-1 below.

	A	B	C	D	E	F	H
1	Arabic loanword in Turkish	meaning	Phonemic transcription of the source and target words	ic equiv	Etymology	remark	source
2	āzā	members of a club	/ʔaʕdˤaːʔ/ ---> /âzâ/	أعضاء	1330	Aşık Paşa, Garib-name, 1330	Dağrcık sözlük
3	âdet	customs; habit	/ʕaːdah/ ---> /âdet/	عادة	1354	Mesud b. Ahmed, Süheyl ü Nevbahar, 1354; (	ww.turkishlanguage.co.uk/freqvocab.htm
4	aba	cloak or coat made	/ʕabaːʔah/ ---> /aba/	عباءة	1341	anon., Tezkiret-ül Evliya, 1341	TDK CD
5	abes	uselessness; absu	/ʕabaθ/ ---> /abes/	عبث	<1500	Kıpçak Türkçesi Sözlüğü, <1500	Dağrcık sözlük
6	acaba	I wonder	/ʕadʒaban/ ---> /adʒaba/	عجبا	1391	Seyf-i Sarayî, Gülistan tercümesi, 1391	ww.turkishlanguage.co.uk/freqvocab.htm
7	acayip	strange things; wo	/ʕadʒaːʔib/ ---> /adʒajip/	عجائب	<1377	Erzurumlu Darir, Kıssa-i Yusuf tercümesi, <13	TDK CD
8	acele	hurry; hastily	/ʕadʒalah/ ---> /adʒele/	عجلة	1360	Danişmend-Name, 1360	ww.turkishlanguage.co.uk/freqvocab.htm
9	Acem	Persian, pertaining	/ʕadʒam/ ---> /adʒem/	عجم	<1300	Mukaddimetü'l-Edeb, <1300	TDK CD
10	acil	urgent	/ʕaːdʒil/ ---> /adʒil/	عاجل	1680	Meninski, Thesaurus, 1680	Dağrcık sözlük
11	acilen	urgently	/ʕaːdʒilan/ ---> /adʒilen/	عاجلا	1680	related to acil; Meninski, Thesaurus, 1680	Dağrcık sözlük
12	aciz	inability	/ʕadʒz/ ---> /adʒiz/	عجز	1391	Seyf-i Sarayî, Gülistan tercümesi, 1391; no co	Dağrcık sözlük
13	adale	muscle	/ʕadˤalah/ ---> /adale/	عضلة	1680	Meninski, Thesaurus, 1680	Dağrcık sözlük
14	adalet	justice	/ʕadaːlah/ ---> /adalet/	عدالة	<1600	Selanikli Mustafa Âli, <1600	Dağrcık sözlük
15	adet	number	/ʕadad/ ---> /adet/	عدد	<1300		Dağrcık sözlük
16	adi	common; vulgar	/ʕaːdiː/ ---> /adi/	عادي	1680	Meninski, Thesaurus, 1680	معجم الألفاظ العربية في اللغة التركية
17	adil	just	/ʕaːdil/ ---> /adil/	عادل	<1300	Mukaddimetü'l-Edeb, <1300	Dağrcık sözlük
18	af	pardon; forgivene	/ʕafw/ ---> /af/	عفو	<1300	Mukaddimetü'l-Edeb, <1300	Dağrcık sözlük
19	afet	calamity, disaster,	/ʔaːfah/ ---> /afet/	آفة	1330	Aşık Paşa, Garib-name, 1330; check with Sam	Türkçe-İngilizce Dizin (2)
20	afiyet	health	/ʕaːfjah/ ---> /afijet/	عافية	<1300		www.turkishlanguage.co.uk/freqvocab.htm

Figure 3-1: A print screen image of the Arabic loanwords into Turkish corpus

Next, the researcher manually mapped the vowels of the Arabic loans to their Arabic counterparts in Arabic words and then verified the mapping statistically using the filter function in Excel to compute the number of words exhibiting each vowel mapping. This process led the researcher to the patterns that manifest shortening of long vowels and the residual effects of gutturals neighbouring vowels that shall be discussed in section 3.3.

Then, two speakers of Turkish checked whether all the words in the corpus are still in use in Modern Standard Turkish (MST). This step enabled the researcher to substitute old words with new ones and, in turn, led to remapping the vowels to eliminate old words. This was followed by stratifying the words by the years in which they were in use. Three references were consulted, namely Tietze's etymological volumes (2002a, 2002b) , Hasan Eren's book (1999) and Nişanyan's online etymological dictionary (2007).

In what follows, the actual vowel patterns are described. All data come from current work unless otherwise stated.

3.3 Arabic loanwords in Turkish corpus data

The Arabic loanwords into Turkish (ALT) data exhibit two types of patterns, namely shortening of the three long Arabic vowels and residual effects of pharyngealization and uvularization on neighbouring vowels. First, the three long Arabic vowels /a:/, /i:/ and /u:/ are rendered short regardless of the surrounding vowels or consonants in the Turkish word or their Arabic cognate. Three words exemplifying this are /ʕaːr/>/ar̥/ 'shame', /kafiːl/>/cefil/ 'guarantor' and /maħsˤuːl/>/mahsul/ 'crops' respectively.

In addition, short vowels neighbouring guttural sounds reflect effects of pharyngealization and uvularization harmony similar to emphasis spread found in many Arabic dialects. In MST, these three vowel mappings — on the surface — seem to follow the rules of Turkish vowel harmony (VH) where a vowel harmonizes for backness and roundedness to the vowel in the preceding syllable internally or across morpheme boundaries. However, VH is not sensitive to the presence of guttural sounds; i.e., pharyngeals, emphatics and uvulars simply because MST does not have gutturals in its inventory. This thesis proposes that the perceived vowel patterns in the ALT are traces or residues of guttural sounds in the original Arabic cognates. These effects or vowel mappings along with the shortening of long vowels in the ALT are further explained in the next subsection.

3.3.1 Mapping of long vowels

3.3.1.1 /a:/

In the ALT corpus, it is found that /a:/ is adapted as a short vowel in 481 cases where /a:/ is realized as /a/ 479 times whereas it is mapped to /e/ in only 2 words. The words in table 3-1 below illustrate the adaptation of /a:/ as /a/ which clearly shows that the quality of the resulting vowel is not affected by the presence or absence of any triggering segments be they vowels, or guttural or non-guttural consonants.

1. /a:/> /a/

Guttural context in source word	No guttural context in source word
a. /ʕaða:b/>/azap/ 'punishment'	b. /dawa:m/>/devam/ 'continuation'
c. /ħa:l/>/hal/ 'condition'	d. /niha:jah/>/nihajet/ 'end'
e. /baqqa:l/>/bakkal/ 'grocer'	f. /maka:n/>/mecan/'location'
g. /xa:ʔin/>/hain/ 'traitor'	h. /rafa:h/>/refah/ 'prosperity'
i. /ʔiʃɣa:l>/iʃgal/ 'distraction'	j. /dʒalla:d/>/dʒellat/ 'executioner'
k. /sˤa:ħib/>/sahip/ 'possessor'	l. /ʔinsa:n/>/insan/ 'human being'
m. /ʔimdˤa:ʔ/>/imza/ 'signature'	n. /tada:wi:/>/tedavi/ 'treatment'
o. /ʔibtˤa:l/>/iptal/ 'cancellation'	p. /ta:bu:t/>/tabut/ 'coffin'
q. /ðˤa:lim/>/zalim/ 'oppressor'	r. /ða:tan/>/zaten/ 'in fact'

Table 3-1: /a:/ to /a/mapping in the ALT corpus data

As mentioned above, very few examples deviate from the main pattern as only in 2 words the vowel /a:/ is borrowed as /e/ as in 2. below.

2. /a:/> /e/

i. /ma da:m/>/madem/ 'since'
ii. /ʔilza:m/>/elzem/ 'most necessary'

3.3.1.2 /i:/

Words with original[9] long /i/ are also generally adapted as a short vowel of the same quality /i/. In the corpus, this amounts to 249 cases as exemplified in 1. (table 3-2) below while in 3 words the long /i/ is borrowed as the back vowel /ɯ/ and in 1 word as /e/.

1. /i:/> /i/

Guttural context in source word	No guttural context in source word
a. /masa:ʕi:/>/mesai/ 'efforts'	b. /tasli:m/>/teslim/ 'delivering'
c. /tardʒi:ħ/>/terdʒih/ 'preference'	d. /ʃahi:d/>/ʃehit/ 'martyr'
e. /daqi:qah/>/dakika/ 'minute'	f. /taʃki:l/>/teʃcil/ 'formation'
g. /faxri:/>/fahri/ 'honorary'	h. /kafi:l/>/cefil/ 'guarantor'
i. /tabli:ɣ/>/teblii/ 'notification'	j. /jami:n/>/jemin/ 'oath'
k. /nasˤi:b/>/nasip/ 'fate'	l. /taʔsi:s/>/tesis/ 'establishing'
m. /fadˤi:lah/>/fazilet/ 'virtue'	n. /hadi:jjah/>/hedije/ 'present'
o. /ʃari:tˤ/>/ʃerit/ 'ribbon'	p. /nati:dʒah/>/netidʒe/ 'result'
q. /waðˤi:fah/>/vazife/ 'duty'	r. /laði:ð/>/leziz/ 'tasty'

Table 3-2: i:>i mapping in the ALT corpus data

[9] The use of the word original refers to length in the cognate Arabic word.

As can be seen from the words in 1., the vowel selection is not determined by the nature of neighbouring vowels or consonants (gutturals or non-gutturals). Nevertheless, four words only as in 2. below reflect possibly the influence of the emphatic consonants, /q/ or the pharyngeal **/ħ/** as in /sˤ/ ‘ħa**sˤ**iːr’, /q/ ‘**q**iːmah’ and /dˤ/ ‘raː**dˤ**iː’ and ‘**ħ**asˤiːr’. Moreover, one loanword has a shortened /e/, namely ‘eyvallah’ (meaning if you say so) in iv.

2. /iː/> /ɯ/, /e/

i. /**ħasˤ**iːr/>/hasɯr/ ‘hasır: reed mat’
ii. /**q**iːmah/>/kɯjmet/ ‘kıymet: value’
iii. /raː**dˤ**iː/>/ɾâzɯ/ ‘razı: willing’
iv. /ʔiːh wallah/>/ejvallah/ ‘so be it, if you say so’

3.3.1.3 /u:/

Long /u/ is systematically adapted as a short vowel in 71 cases of which it is borrowed as /u/ 69 times, /o/ and /y/ only one time each. As with the vast majority of /a:/ and /i:/, the resulting vowel’s selection is not sensitive to the presence or absence of surrounding vowels or consonants (guttural or non-guttural in the Arabic cognate words) as in 1. (table 3-3) below.

1. /uː/>/u/

Guttural context in source word	**No guttural context in source word**
a. /maʃruːʕ/>/meʃru/ ‘legal’	b. /hudʒuːm/ >/hydʒum/ ‘assault’
c. /marħuːm/>/meɾhum/ ‘deceased’	d. /maʃhuːr/>/meʃhuɾ̥/ ‘famous’
e. /jaːquːt/>/jakut/ ‘ruby’	f. /kaːbuːs/> /câbus/ ‘nightmare’
g. /maxluːq/> /mahluk/ ‘creature’	h. /masʔuːl/> /mesul/ ‘official’
i. /maʃɣuːl/>/meʃgul/ ‘busy’	j. /sukuːn/>/sycun/ ‘ tranquility’
k. /ʔusˤuːl/>/usul/ ‘fundamentals’	l. /dʒaːsuːs/>/dʒasus/ ‘spy’
m. /dˤaruːriː/> /zaɾuɾi/ ‘essential’	n. /dajjuːθ/>/dejjus/ ‘pander’
o. /rutˤuːbah/>/ɾutubet/ ‘humidity’	p. /maktuːb/>/mectup/ ‘letter’
q. /maðˤluːm/>/mazlum/ ‘oppressed’	r. /maʔðuːniːjah/>/mezunijet/ ‘permission’

Table 3-3: u:>u mapping in the ALT corpus data

Nevertheless, the words in 2., i.e. *bornoz* and *hükümet* manifest two separate patterns. In the word *bornoz*, the /u:/ is adapted as /o/ which might have been generalized from Persian words that entered Ottoman language in the same era (<1300) such as the Persian word *dost* from the Persian cognate /du:st/ 'friend' which had an older /o:/ (Stein, 2006). Another possibility for the word bornoz is that it might have been borrowed directly from Moroccan Arabic reflecting the vowel /o/ used in the vernacular.

2. /u:/> /o/, /y/

i. /burnu:sˤ/>/bornoz/ 'hooded gown worn in Morocco'

ii. /ħuku:mah/>/hycymet/ government'

Thus far, the adaptations of the three Arabic long vowels have been described and are statistically summarized in table 3-4 before describing the adaptations of the Arabic short vowels in section 3.3.2.

Arabic	i:	i:	i:	a:	a:	a:	u:	u:	u:
Turkish	i	ɯ	e	a	i	e	u	o	y
Frequency	249	3	1	483	0	2	69	1	1

Table 3-4: Number count of long vowels shortening in the ALT corpus data

3.3.2 Mapping of short vowels

In this subsection, the residual effects of gutturals on neighbouring vowels are described as part of the adaptation of all the short vowels found in the corpus.

3.3.2.1 /a/

1. /a/ > /a/, /e/

In the corpus of the current work, the Arabic short vowel /a/ is adapted either as the back vowel /a/ or the front vowel /e/ among few other categories. It is borrowed as the vowel /a/ 352 times when surrounded by any one of the nine guttural consonants [pharyngeals, uvulars and emphatics] in the Arabic source word whereas as the vowel /e/ 485 times elsewhere. These observations conform to those found and described by Schaade (1927). The two patterns are exemplified in table 3-5 below word internally and across morpheme boundaries.

/a/>/a/	/a/> /e/
a. /**ɣa**flah/>/**ga**flet/ 'unawareness'	b. /ʔawwalan/>/evvela/ 'first'
c. /**ʕa**baθ/>/**a**bes/ 'uselessness'	d. /badan/>/beden/ 'body; trunk'
e. /na**ðˤa**r/>/nazar/ 'look; glance'	f. /dawlah/>/devlet/ 'state; nation'
g. /**xa**bar/>/**ha**ber/ 'news'	h. /falsafah>/felsefe/ 'philosophy'
i. /**ħa**kam/>/**ha**kem/ 'referee'	j. /dʒalsah/>/dʒelse/ 'law hearing'
k. /**tˤa**bi:b/>/tabip/ 'doctor'	l. /laððah/>/lezzet/ 'pleasure; enjoyment'
m. /ta**sˤ**wi:r/>/tasviɾ̥/ 'depiction'	n. /rasim/>/resim/ 'picture'
o. /ma**q**a:m/>/makam/ 'position; office'	p. /**ha**di:jjah/>/**he**dije/ 'present'
q. /na**bdˤ**/>/nabɯz/ 'pulse'	r. /kafan/>/kefen/ ' sheet to wrap the dead'

Table 3-5: a>a and a>e mappings stem internally (data from current research)

From table 3-5, one can infer that the gutturals subset of the original Arabic words surrounding a cognate /a/ which is borrowed as the back vowel **/a/** in Turkish comprises the **pharyngeals /ʕ/** and**/ħ/**, the **uvulars /x/, /ɣ/**and **/q/**, the **emphatics /tˤ/, /dˤ/, /ðˤ/** and **/sˤ**/ with the exclusion of the two **laryngeal** sounds **/ʔ/** and **/h/** since they do not participate in this effect as illustrated in b. and p. in table 3-5.

The adaptation of /a/ as /e/ in the ALT resembles vowel raising (***Imala***) in the Levant dialects of Arabic except that Imala is not an active process in MST. The examples in (**1**), reproduced from section 2.2.5. are from Syrian Arabic where the words to the left exemplify vowel raising in contrast the ones to the right with the gutturals /h/, /ħ/ and /rˤ/ immediately before the feminine ending. The presence of the gutturals spreads emphasis effect to the neighbouring vowels (uvularization).

(**1**) Syrian Arabic feminine suffix -e/a (Cowell, 1964; Rose, 1996)

a. daraž-e 'step'	d. wa:z**h**-a 'display'
b. šerk-e 'society'	e. mni:**ḥ**-a 'good'
c. madras-e 'school'	f. dagga:**R**-a 'tanning'

Unlike Arabic, Turkish does not have a gender system. However, many Arabic loans were borrowed with their feminine endings from Arabic. In these words, as illustrated in table 3-

6, the feminine suffix is realized as /-e/ after non-guttural sounds and /-a/ after gutturals in the original Arabic words (Perry, 1984).[10]

Mapping of /a/ to /e/	**Mapping of /a/ to /a/**
a. /dʒal**s**-**a**h/>/dʒels**e**/ 'session; hearing'[11]	b. /bilxa:**sˤsˤ**-**a**h/>/bilhass**a**/ 'especially'
c. /saʕa:**d**-**a**t/>/saad**e**t/ 'happiness'	d. (No word with **dˤ** before the feminine ending)
e. (No word with **t** before the feminine ending)	f. /nuq**tˤ**-**a**h/>/nokt**a**/ 'full stop'
g. /nub**ð**-**a**h/>/nebz**e**/ 'a tiny bit'	h. /ħa:fi**ðˤ**-**a**h/>/hafɯz**a**/ 'memory'
i. /taʕrifah/>/tarife/ 'tariff'	j. /nus**x**-**a**h/>/nysh**a**/ 'copy'
k. /laʕ**n**-**a**h/>/lan**e**t/ 'curse'	l. (No word with **ɣ** before the feminine ending)
m. /ħara**k**-**a**h/>/hareke**t**/ 'movement; act'	n. /ʕala:**q**-**a**h/>/alak**a**/ 'relationship'
o. /tawsˤi:**j**-**a**h/>/tavsij**e**/ 'recommendation'	p. /ra:**ħ**-**a**h/>/rah**a**t/ 'peace; comfort'
q. /ʃir**k**-**a**h/>/ʃirc**e**t/ 'company'	r. /fa:dʒi**ʕ**-**a**h/>/fadʒi**a**/ 'calamity'
s. /ʃub**h**-**a**h/>/ʃyph**e**/ 'suspicion'	t. /ʕab**a:ʔ**-**a**h/>/ab**a**/ 'coat made of cloth'

Table 3-6: a>e and a>a across morpheme boundaries with feminine ending -e/-a

Despite the regularity of the two patterns mentioned above, some exceptions are found. (**2**) below illustrates /a/>/e/ mapping where /a/>/a/ mapping is expected. We find effects of emphasis spread in the Arabic word as the presence of emphatics including uvular /q/ reflects effects of unlimited leftward uvularization/emphasis spread (regressive assimilation) in 2A. and limited rightward uvularization (progressive assimilation) as in 2B.

(2) Exceptions to the /a/>/a/ mapping

A. Leftward assimilation

i. /ba.**sˤ**i:.rah/>/ba.si.ɾet/ 'basiret' (foresight)
ii. /wa.**sˤ**i:.jah/>/va.si.jet/ 'vasiyet' (will, statement of a dying person)
iii. /ra.fi:**q**/>/ɾa.fik/ 'rafik' (companion)
iv. /man.**ðˤ**ar/>/man.za.ɾa/ (inserted back vowel) 'manzara' (view)
v. /fa.**dˤ**i:.lah/>/fa.zi.let/ 'fazilet' (virtue)
vi. /ra.ma.**dˤ**a:n/>/ɾa.ma.zan/ 'Ramazan' (The Holy month of Ramadan)
vii. /ta.laf.fu**ðˤ**/>/te.laf.fuz/ 'telaffuz' (pronunciation)
viii. /mu.sa:.ba.**q**ah/>/my.sa.ba.ka/ 'müsabaka' (competition)
ix. /na.**ðˤ**ar/>/na.zaɾ̥/ 'nazar' (look, glance)
x. /wa.**tˤ**an/>/va.tan/ 'vatan' (motherland)

[10] Perry refers to 1500 loanwords in the feminine ending in Persian, Turkish and other non-Arabic languages, of which 1150 items in Turkish, 580 ending in -at (contextual form) and 570 in –a (pausal form).

B. Rightward assimilation

i. /fa**d**ˤ.lan/>/faz.la/ 'fazla' (too much, too many)
ii. /**d**ˤa.rar/>/za.ɾar̥/ 'zarar' (damage, injury)

C. Blocking of rightward assimilation

i. /**d**ˤar.bah/>/daɾ.be/ 'darbe' (hit) *darba
ii. /**ð**ˤa.far/>/za.fer̥/ 'zafer' (victory) *zafar
iii. /**q**a.lam/>/ka.lem/ 'kalem' (pen, pencil) *kalam
iv. /**s**ˤa.daf/>/se.def/ 'sedef' (pearls shells) *sadaf/sadef

D. Blocking of leftward assimilation

i. /ʔa**t**ˤ.raːf/>/et.ɾaf/ 'etraf' (sides) *atraf

The observations seen in 2A and 2B conform to the trend of emphasis spread in a number of Arabic dialects most of which display unlimited leftward assimilation and limited rightward assimilation blocked by some opaque segments, namely [ʃ, ʒ, j and i] (Davis, 1995; Herzallah, 1990; McCarthy, 1997; Shahin, 1997; Watson, 1999; Younes, 1991; Zawaydeh & de Jong, 2003). Although the effect of leftward emphasis spread is generally more iterative than the rightward spread, sometimes it is blocked by certain segments such as /l/ as in /**ta**.laf.fuðˤ/>/**te**.laf.fuz/ which is treated as a trigger of palatalization in Turkish Linguistics literature and a blocking segment in Arabic linguistics as mentioned above (Davis, 1995; Stein, 2006; A Tietze, 1992). In the same vein, emphasis spread displays directionality, meaning that it can proceed leftwards or rightwards, regardless of the position of the guttural sound in the cognate Arabic word as shown in tables 3-7 and table 3-8 below.

Guttural in onset position	Guttural in coda position
a. /**ħa**rf/~/h**a**rf/ 'letter (of the alphabet)'	f. /ʃ**art**ˤ/~/ʃ**a**rt/ 'condition; provision'
b. /**qa**lb/~ /k**a**lp/ 'heart'	g. /f**arq**/~/f**a**rk/ 'difference'
c. /**ɣa**m/~ /**ga**m/ 'grief; sorrow; worry'	h. No example found in data
d. /**ð**ˤ**a**n/~/z**a**n/ 'suspicion; doubt'	i. /f**ard**ˤ/~/f**a**rz/ 'religious duty'
e. /**ð**ˤ**a**rf/~/z**a**rf/ 'envelope'	j. /ʔ**ard**ˤ/~/**a**rz/ 'earth'

Table 3-7: directionality of emphasis spread in monosyllabic ALT words of the pattern a>a

Guttural in onset position	Guttural in coda position
a. /m**a**.**ħ**al/~/m**a**.h**a**l/ 'place; spot'	f. No example found in data
b. /nus.**x**a**h**/~/nys.h**a**/ 'copy'	g. No example found in data
c. /la.**q**a**b**/~/la.k**a**p/ 'nickname'	h. /mu.w**a**f.fa**q**/~/mu.v**a**f.f**a**k/ 'successful'
d. /dʒum.**ʕa**h/~/dʒu.m**a**/ 'Friday'	i. /d**aʕ**.wa:/~/d**a**.va/ 'lawsuit'
e. /wa.**t**ˤ**a**n/>/va.t**a**n/ 'motherland'	j. /mu.sˤal.l**a**tˤ/~/mu.sal.l**a**t/ 'who pesters'

Table 3-8: Disyllabic and polysyllabic words of the shape a>a

Other adaptations of the short vowel /a/ exist albeit in small numbers as shown in (3). The numbers to the right of the mapping indicate the number of examples detected.

(3) Other /a/ mappings

1. /a/> /i/ (5)
   - i. /manaːrah/>/minaɾ̥e/ 'minaret'
   - ii. /ʕarafah/>/aɾife/ 'the day before a religious holiday'
   - iii. /radʒaːʔ/>ɾidʒa/ 'request'
   - iv. /samiːd/>/simit/ 'savory roll covered with sesame seeds'
   - v. /ðarwah/>/ziɾve/ 'summit'

2. /a/> /ɯ/ (3)
   - i. /qaːlab/> /kalɯp/ 'mold'
   - ii. /mantˤiqah/>/mɯntɯka/ 'location'
   - iii. /tˤalsˤam/>/tɯlsɯm/ 'talisman, charm'

3. /a/> /u/ (3)
   - i. /tˤabl/>/davul/ 'drum'
   - ii. /maħabbah/>/muhabbet/ 'muhabbet' (affection, love)
   - iii. /matˤbax/>/mutfak/ 'mutfak' (kitchen)

4. /a/>/y/ (2)
   - i. /raʃwah/>/ɾyʃvet/ 'bribe'
   - ii. /zaraːfah/>/zyɾafa/ 'giraffe'

3.3.2.2 / i /

1. /i/> /i/, /ɯ/

In this pattern, the short vowel **/i/** is adapted as [ɯ] in the environment of either emphatics (/**tˤ**/, /**dˤ**/, /**ðˤ**/, /**sˤ**/) or the uvular /**q**/ as shown in table 3-9. On the other hand, in the environment of the uvular /**x**/, /**ɣ**/, pharyngeal and plain consonants, the **/i/** is borrowed as /**i**/. In the ALT corpus 30 words exhibit the /i/>/ɯ/ compared to 254 cases of /i/>/i/ mapping.

Mapping of /i/ to /ɯ/	Mapping of /i/ to /i/
a. /**qi**smah/>/k**ɯ**smet/ 'fate; destiny'	b. /b**i**na:**ʔ**/>/b**i**na/ 'building'
c. /riz**q**/>/ɾ**ɯ**zk/ '(one's) daily bread; food'	d. /ta:b**iʕ**/>/tab**i**/ 'dependent'
e. /ħa:f**iðˤ**ah/>/haf**ɯ**za/ 'memory'	f. /**xi**la:fah/>/h**i**lafet/ 'the Caliphate'
g. /**sˤi**fah/>/s**ɯ**fat/ 'adjective; role'	h. /ʃa:**hi**d/>/ʃah**i**t/ 'eyewitness'
i. /**tˤi**b/>/t**ɯ**p/ 'Medicine'	j. /**ħi**ka:jah/> /h**i**kaje/ 'story'
k. /r**idˤ**a/>/ɾ**ɯ**za/ 'consent'	l. /**ʔi**sm/> /**i**sim/ 'name'
m. /ħa:**dˤi**r/>/haz**ɯ**ɾ/ 'prepared'	n. /s**i**la:**ħ**/> /s**i**lah/ 'weapon'

Table 3-9: : i>i and i>ı in the ALT corpus data

In only two words in the ALT words with the uvular /**x**/, /**ɣ**/, the vowel /i/ was adapted as /ɯ/. These are /**ɣi**ða:ʔ/~/g**ɯ**da/ 'nourishment' and /**xi**nzi:r/~/h**ɯ**nzɯr/ 'nasty; mean; swine'. However, these words were found to be obsolete and not used in MST anymore; thus, they were removed from the corpus.

The mapping of /i/ to /ɯ/ is not conditioned by the position of the emphatics or q in the words as its effect is bidirectional. Tables 3-10 and 3-11 exhibit directionality of emphasis spread effect in ALT words whose cognate words have either emphatic consonants or q.

Emphatic or q in onset position	Emphatic or q in coda position
a. /**ðˤi**l/~/z**ɯ**l/ 'shadow'	d. /ħi**rsˤ**/~/h**ɯ**rs/ 'greed; being a miser'
b. /**tˤi**b/~/t**ɯ**p/ 'Medicine'	e. /ʕi**rq**/~/**ɯ**rk/ 'race'
c. /**ðˤi**d/~/z**ɯ**t 'opposite; contrary'	f. /r**iz**q/~/r**ɯ**zk/ '(one's) daily bread; food'

Table 3-10: i>ı mapping in monosyllabic words

Emphatic or q in onset position	**Emphatics or q in coda position**
a. /**s**ˤi.fah/ **>** /s**ɯ**.fat/ 'adjective'	e. /ha:.f**i**ð**ˤ**/>/ha.f**ɯ**z/ 'one who has memorized the Koran'
b. /**t**ˤib.bi:/ **>** /t**ɯ**b.bi/'medical'	f. /mu.w**a**f.fa**q**/~/mu.v**a**f.f**a**k/ 'successful'
c. /**q**is.mah/**>** /k**ɯ**s.met/'fate'	g. /la:.ʔ**iq**/>/la:.j**ɯ**k/ 'suitable, appropriate, proper'
d. **/q**ib.lah**/ > /**k**ɯ**b.le**/** 'direction to Mecca'	h. /sa:.b**iq**/>/sa.b**ɯ**k/ 'former; previous'

Table 3-11: i>ı mapping in disyllabic and polysyllabic words

In addition, the quality of the epenthetic vowel in the ALT words is found to be linked with the presence of emphatics and the uvular /q/ in the cognate Arabic word. For instance, the vowel /i/ is realized in /s**i**ħr/>/s**i**h**i**r/ 'magic', /**ʔi**sm/> /**i**sim/ 'name' among others while /ɯ/ is inserted in words such as **/sˤ**inf/>/s**ɯ**n**ɯ**f/ 'class; /**sˤi**fr/~/s**ɯ**.f**ɯ**r/ 'zero' and **/qi**sm/>/k**ɯ**s**ɯ**m/ 'section' among others.

Despite the productivity of the /i/ to /ɯ/ mapping, some exceptions were found including the words given below (22 words) where an emphatic or a q consonant is present in the Arabic source word. It seems that /i/ is affected by the presence of the laryngeal /ʔ/ consonant as a prefix or the co-occurrence of the q and the emphatic consonants in the same word. The words that follow in 2. represent other adaptations of /i/ found in the corpus.

Exceptions and disharmonic forms:

i. /fa:ʔidˤ/>/faiz/
ii. /ʔimdˤa:ʔ/>/imza/
iii. /ʔiqtisˤa:di:/>/iktisadi/
iv. /ʔiqtisˤa:d/>/iktisat/
v. /ʔinsˤa:f/>/insaf/
vi. /ʔisˤa:bah/>/isabet/
vii. /mutasˤawwif/>/mutasavvif/
viii. /niðˤa:m/>/nizam/
ix. /niðˤa:mi:/>/nizami/
x. /diqqah/>/dikkat/
xi. /ʔiqa:mah/>/ikamet/
xii. /ʔiqtida:r/>/iktidaɾ̥/
xiii. /ʔiqtisˤa:di:/>/iktisadi/
xiv. /ʔiqtisˤa:d/>/iktisat/
xv. /ʔintiqa:m/>/intikam/
xvi. /ʔistiqra:r/>/istikraɾ̥/

xvii. /ʔittifaːq/>/ittifak/
xviii. /qaːbiliːjah/>/kabilijet/
xix. /qaːfilah/>/kafile/
xx. /miqdaːr/>/mictar̥/
xxi. /muqtadir/>/muktedir̥/
xxii. /mustaqil/>/mystacil/

2. Other mappings /i/ > /e/, /a/, /y/:

A. /i/>/e/
i. /fidaːʔ/ > /feda/
ii. /ʔinqaːð/ > /enkaz/
iii. /xaːdimah/ > /hademe/
iv. /ħisaːb/ > /hesap/
v. /miʃʕalah/ > /meʃale/
vi. /riqaːbah/ > /ɾekabet/
vii. /tˤaːlibah/ > /talebe/
viii. /tˤaːlib/ > /talep/
ix. /tikraːr/ > /tecɾaɾ/
x. /wiraːθah/ > /veɾaset/
xi. /sˤinaːʕah/ > /zannet/

B. /i/ > /a/
i. /xilxaːl/ > /halhal/
ii. /ʕiʃq/ > /aʃk/
iii. /bikr/ > /baciɾe/
iv. /niʕnaːʕ/ > /nane/
v. /tˤiraːz/ > /taɾz/

C. /i/ > /y/
i. /mumkin/ > /mymcyn/
ii. /musrif/ >/mysɾif/

3.3.2.3 / u /

1. /u/> /u/, /y/

In this pattern, the vowel /u/ is realized as /u/ if the original Arabic word has a guttural consonant; otherwise, it is mapped as /y/. In the corpus, 61 words are found of the pattern /u/>/u/ compared to 82 words of /u/ to /y/. Table 3-12 below provides some examples of both patterns.

Mapping of /u/ to /u/	Mapping of /u/ to /y/
a. /buxa:r/>/buhar/ 'steam'	b. /muttafiq/>/myttefik/ 'in agreement'
c. /qubbah/>/kubbe/ 'dome'	d. /muddah/>/myddet/ 'period; duration'
e. /ðˤulm/>/zulym/ 'injustice'	f. /taʕahhud/>/taahhyd/ 'commitment'
g. /ħuqu:q/>/hukuk/ 'rights'	h. /dʒumlah/>/dʒymle/ 'sentence'
i. /ɣuru:r>/gurur/ 'pride'	j. /tasˤa:duf/>/tesadyf/ 'chance event'
k. /sultˤa:n/>/ sultan / 'sultan'	l. /taðˤa:hura:t/>/tezahyrat/ 'demonstrations'
m. /ʃuʕbah/>/ʃu.be/ 'branch office'	n. /ʃubhah/>/ʃyphe/ 'doubt; uncertainty'

Table 3-12: u>u and u>ü mappings in the ALT corpus data

In the corpus, no monosyllabic words manifesting the two patterns /u/>/u/ and /u/>/y/ were found except for the word ħur>**hür** 'free' where the /u/ is realized as /y/ despite the presence of the pharyngeal /ħ/ in the cognate word. Nonetheless, many disyllabic and polysyllabic words displaying the patterns were found which are given in table 3-13. Some exceptions to these patterns are also provided in 2. followed by other patterns of the vowel /u/.

Guttural in onset position in Arabic word	Guttural in coda position in Arabic word
a. /bu.xa:r/~/bu.har/ 'steam'	h. /mux.ta.lif/~/muh.te.lif/ 'diverse'
b. /ħum.ma/~/hu.ma/ 'fever'	i. /muħ.ta:dʒ/~/muh.tatʃ/ 'needy; poor'
c. /ru.tˤu:.bah/~/ru.tu.bet/ 'humidity'	i. /mutˤ.laq/~/mut.lak/ 'absolute'
d./mu.ʕa.ma.lah/~/mu.a.me.le/ 'treatment'	k. /muʕ.ta.dil/ ~/mu.te.dil/ 'moderate; mild'
e. /ɣu.ru:r/~/gu.rur/ 'pride'	/mu.waf.faq/~/mu.vaf.fak/ 'successful'
f. /qud.rah/~/kud.ret/ 'power, might'	/ʔu.fuq/~/u.fuk/ 'horizon'
g. /ðˤulm/~/zu.lym/ 'injustice'	/ta.laf.fuðˤ/~/te.laf.fuz/ 'pronunciation'
/tu.ħaf/~/tu.haf/ 'odd'	/ʃuʕ.bah/~/ʃu.be/ 'branch office; section'
/sul.tˤa:n/~/sul.tan/ 'sultan'	/muq.ta.dir/~/muk.te.dir/ 'capable'

Table 3-13: u>u and u>ü in disyllabic and polysyllabic words (initial position)

In table 3-13, some of the onsets are in the same syllable as the target vowels while others are not. Nevertheless, a lot of variation was found. In addition, no consistent pattern was detected since the token numbers of the gutturals being in either the onset or coda positions of the syllable in the corpus were not significant to make a generalization.

2. Exceptions:

i. /lutˤfan/>/lytfen/
ii. /muħtamalan/>/myhtemelen/
iii. /ħudʒrah/>/hydʒɾe/
iv. /ħukm/>/ħucym/
v. /ħukuːmah/>/hycymet/
vi. /ħur/>/hyɾ̥/
vii. /ħurriːjah/>/hyɾɾijet/
viii. /ħuzn/>/hyzyn/
ix. /musaːmaħah/>/mysamaha/
x. /mustariːħ/>/mysteɾih/
xi. /taħammul/>/tahammyl/
xii. /muħaːfaðˤah/>/muhafaza/
xiii. /muħaːsabah/>/muhasebe/
xiv. /muħtaːdʒ/>/muhtatʃ/
xv. /muħtamal/>/muhtemel/
xvi. /tuħaf/>/tuhaf/
xvii. /tasˤaːduf/>/tesadyf/
xviii. /taðˤaːhuraːt/>/tezahyɾat/
xix. /mulaːqaːh/ > /mylakat/
xx. /musaːbaqah/> /mysabaka/
xxi. /mustaqil/>/mystacil/
xxii. /muttafiq/> /myttefic/

3. Other mappings: /u/ > /ö/, /ɯ/

i. /ʕurf/>/örf/ ‘custom’
ii. /qurtˤaːsˤiːjah/>/kɯrtasije/ 'stationary'
iii. /furn/>/fɯrɯn/ ‘oven’

The patterns described above are substantiated by mapping the Arabic vowels to their ALT correspondents. The correspondence of the three Arabic vowels between Arabic and the ALT is represented in table 3-14 as follows[11].

Arabic	a	a	a	a	a	a	i	i	i	i	i	u	u	u	u	u	u	u
Turkish	a	e	i	ɯ	y	u	i	ɯ	e	a	y	u	y	o	e	i	ɯ	œ
Frequency	351	485	5	3	2	1	254	30	12	5	2	61	82	3	1	2	1	2

Table 3-14: Arabic-Turkish vowel correspondence from the corpus used

As can be seen from table 3-14, the vowel /e/ as a mapping to ALT from /a/ in Arabic has a higher frequency (485 tokens) in comparison to /a/$_A$>/a/$_T$ (351) rendering it as the default vowel. Similarly, there is a vast difference in distribution between the vowel /i/$_A$>/i/$_T$ (254) and /i/$_A$>/ɯ/$_T$ (with 30 only) whereas the difference in distribution between the vowel /u/$_A$>/y/$_T$ (82) and /u/$_A$>/u/$_T$ (61) is marginal. In the next section, the ALT corpus is stratified

3.4 Stratification task

This chapter has provided theoretical descriptions of the effects of gutturals on neighbouring vowels in Arabic loanwords in Turkish during the time of the Ottomans. Some assumptions were made based on the works by Tietze (1992) and Stein (2006) since the patterns and findings identified bear resemblance to the latter's. One of these is that the patterns pertaining to the residual effects of gutturals on neighbouring vowels date back to the second stage of Arabic loanwords adaptation in Turkish. Tietze (1992) assumes that the second stage might have begun early but was never completed causing a gap between the language registers in Osmanlica; i.e., high standard Osmanlica employed by educated people and the Turkish colloquial used by the masses.

In this thesis, a full stratification of the corpus words was not pursued since the words compiled, which are still in use in MST, come from different centuries. Instead, the researcher first mapped all the loanwords to their Arabic cognate counterparts as well as cited the dates

[11] Other vowel correspondences have not been discussed as primary alternations since their tokens are small compared to the ones boldfaced above. These are /a/~/ɯ/ surrounding residue of emphatics and uvulars, /i/~/ɯ/ surrounding residue of pharyngeals, uvulars and emphatics and /u/~/o/ in the proximity of residue of pharyngeals, uvulars and emphatics.

of when each word was in use. The oldest word (davul 'drum') was found to date back to 1070; i.e., to the 11th century when the Turks embraced Islam whereas the latest words (faaliyet 'activity' and nihai 'final') were in use in 1930, two years before the establishment of the Turkish Society Association (TDK) which eliminated thousands of Arabic and Persian words from Turkish. Table 3-15 below provides a summary of the corpus monosyllabic and polysyllabic words' counts used in the period before the 1300 and 1930.

<1300-1930		
Type of word	**Count**	**Examples**
disharmonic words	450	idrar, meşhur, hafif, dakika, ticaret, itaat, sükun
harmonic words	605	maaş, buhar, ıslah, batıl, ümle, mescit, bedevil
monosyllabic words	63	farz, örf, zevc, hür, hırs, zan, has, din, tam, nur
Total	1118	

Table 3-15: a summary table of the counts of harmonic and disharmonic words

The reason for dividing the words into harmonic and disharmonic is that many loanwords were subjected to the rules of Turkish vowel harmony during the second stage of adaptation. However, many words kept their older form which proves that the second stage was never quite completed as Tietze attests (1992). Table 3-16 below provides the specific time periods of the words used in the ALT corpus data along with a count of these words and some examples.

Time period	Count	Examples
1070	1	davul
<1300	301	aziz, batıl, mübarek, ayıp, ar, nikâh, nefis, tevekkül
<1353	2	akıbet, evlat
<1377	17	bornoz, ihtiyaç, ikram, ruh, vatan, basit, vatani, hüzün
1300	20	kalıp, fıstık, nokta, tavus, zeytin, maksat, meşhur, kira
1330	108	alet, darbe, ebedi, fail, farz, kabız, ceset, nasip, külliyen
1341	54	defin, itimat, kafile, hamam, bakkal, zelzele, mendil, ihtiyati
1354	33	cilt, fakir, kubbe, malum, kıyas, feda, şükür, sel, sevap, sokak
1360	39	ahbap, hatta, fert, hücum, makbul, isyan, hilafet, diyar, defter
1377	3	maaş, veli, azim

1378	54	gasp, haşin, ecza, hatip, hür, kısım, vesile, külfet, fidye, mesela
1391	62	devir, dakika, feza, has, heykel, itiraf, muhtemel, mutfak, teftiş
<1400	35	asker, hançer, kaide, lakin, mahal, lisan, hicri, hemze, meşgul, nesil
<1410	1	ceviz
<1481	1	Cep
1409	2	imdat, mükemmel
1410	1	Fırın
1420	9	fıtık, ishal, sara, kâbus, zamk, sumak, cüzzam, basur
1423	2	refah, tecavüz
1432	6	hassas, mahzen, hassasiyet, makale, nabız
1437	11	devre, evrak, imsak, saha, tesis, teşkil, ticaret, zambak, ticari
1449	4	gıpta, hamil, kutup, şube
1451	11	arz, ayet, daire, define, ders, dikkat, esna, esnaf, keyfiyet, keyif
1465	8	besmele, hortum, ihtimam, iştirak, nezle, teslim, tılsım, tılsım
1477	2	telaffuz, zürafa
1481	13	kanun, muavin, muvaffak, muvaffak, muvaffak, müzakere,
1482	11	tasvir, tedrici, tehdit, terkip, teyit, ecnebi, istiklal, mühendis, sene
1492	6	istiklal, mühendis, sene, sıhhat, veliaht
<1500	42	iffet, beraat, emel, cila, hecin, hudut, inkılap, isabet, kefil, kürevi
1501	17	hademe, edat, ifade, ihbar, ırk, makas, mastar, simsar, ziraat, tetkik
1520	1	kahpe
1533	1	Sıfır
1545	2	müsrif, zarf
1546	1	hattat
1549	1	üslup
1557	17	firar, valide, hadise, harita, hatıra, hükümet, ibraz, icraat, sükun
<1600	11	ahali, arazi, bakire, beyaz, celse, ciddiyet, kahve, maharet, sanayi
1645	1	Itaat
1647	18	cani, devam, haşarat, haşere, hukuk, iade, iftihar, ihtiram, kabiliyet
<1680	2	ahşap , papağan
1680	162	aci, cadde, cesur, ebat, cenin, hürriyet, ibra, bünye, akraba, aynı
1797	1	küme
1854	1	belediye
1870	1	Şifre
1876	15	bamya, itfaiye, müessese, itina, şelale, ikamet, fahri, istimlak, tarife
1896	2	bilhassa, ilelebet

1900	10	battaniye
1930	2	faaliyet, nihai

Table 3-16: stratification of the ALT corpus data according to the time of use

The results of the stratification task in table 3-16 can be summarized as follows.

1. Harmonic words exist as early as 1070 and <1300 indicating that these words were borrowed during the second stage of adaptation. Some examples are shown in table 3-17 such as tavul '/tˤabl/>/tavul/ drum' (1070), ayıp /ʕajb/>/ajɯp/ 'shame', batıl /ba:tˤil/>/batɯl/ 'untrue' (<1300), bornoz /bornu:sˤ/>/bornoz/ 'hooded gown worn in Morrocco' (<1377) and vatan /watˤan/>/vatan/ 'homeland' (<1377) among many others.

2. The perception of emphatics and gutturals in general in the source words as signals of disharmony is attested in early words too. Some examples provided in () include aziz (**ʕ** before /a/) /ʕazi:z/>/aziz/ 'dear' (<1300), kalıp (**q**) /**q**a:lab/>/kalɯp/ 'heart' (1300), gasp (**ɣ**) /ɣasˤb/>/gasp/ 'law wrongful seizure' (1378), haşin /xaʃin/>/haʃin/ (**x**) 'tough' (1378), ba**s**it (**sˤ**) /basˤi:tˤ/>/basit/ 'simple' (<1377) among others.

3. Long vowels were adapted as short vowels in early periods whether in the presence or absence of signal words in the cognate Arabic word as in aziz (**ʕ** before /a/) /ʕazi:z/>/aziz/ (<1300), **a**kıbet /ʕa:qibah/>/akɯbet/ 'outome' (**ʕ** before /a/) '' (1377), ba**si**t (sˤ) /basˤi:tˤ/>/basit/ 'simple' (<1377), ebedi /ʔabadi:/>/ebedi/ 'eternal' (1330); hücum /hudʒu:m/>/hydʒum/ 'assault' (1360) and meş**gu**l (ɣ) /maʃ**ɣ**u:l/>/meʃgul/ 'busy' (<1400).

4. The results in 1. through 3. above match those found in Tietze (1992) and Stein (2006).

3.5 Summary

To recap, this chapter demonstrates that the Arabic loanwords in modern Standard Turkish show residual effects of gutturals on neighbouring vowels. Although these patterns were previously noted in Schaade (1927), Tietze (1992) and Stein (2006), the current thesis confirms the patterns on the basis of a new, updated corpus of loanwords extracted from the current MST lexicon. The current research not only replicates the careful work of identifying the mappings in the corpus but also goes a step further in identifying the degree of variation in those mappings.

The residual effects of gutturals on vowels reveal information about the grammar that the speakers of Osmanlica had for Arabic loanwords. This grammar is depicted in figure 3-2 in both plain and velar environments. The subscripted "$_T$" denotes Turkish as the native language of the Ottomans, and subscripted "$_A$" denotes Arabic; the dˤ, q and ʕ subscripts represent the various non-plain environments in Arabic: emphatics, uvulars (q, x and ɣ) and pharyngeals (ʕ and ħ) respectively. A short plain $[a]_A$ vowel is mapped to the vowel $/e/_T$ whereas its long variant $[a:]_A$ is mapped to the vowel category $/a/_T$. In the environment of pharyngeals, uvulars (q, x and ɣ) and emphatics all $/a/_A$ variants are mapped to $/a/_T$. The mapping of the vowel $/i/_A$ is more restricted as the vowel /i/ surrounding emphatics ($[i]_{d^ʕ}$) and the uvular /q/ ($/i/_q$) but not /x/ and /ɣ/ or pharyngeals /ʕ/ and /ħ/ is assimilated as $/ɯ/_T$. In all other environments, $/i/_A$ is realized as a vowel of the same quality, i.e., $/i/_T$ including plain, emphatic and pharyngeal contexts. As for the vowel /u/, the plain short variant $[u]_A$ is assimilated as $/y/_T$ but as $/u/_T$ elsewhere; i.e., short and long $/u/_A$ variants surrounding emphatics, all uvulars and pharyngeals.

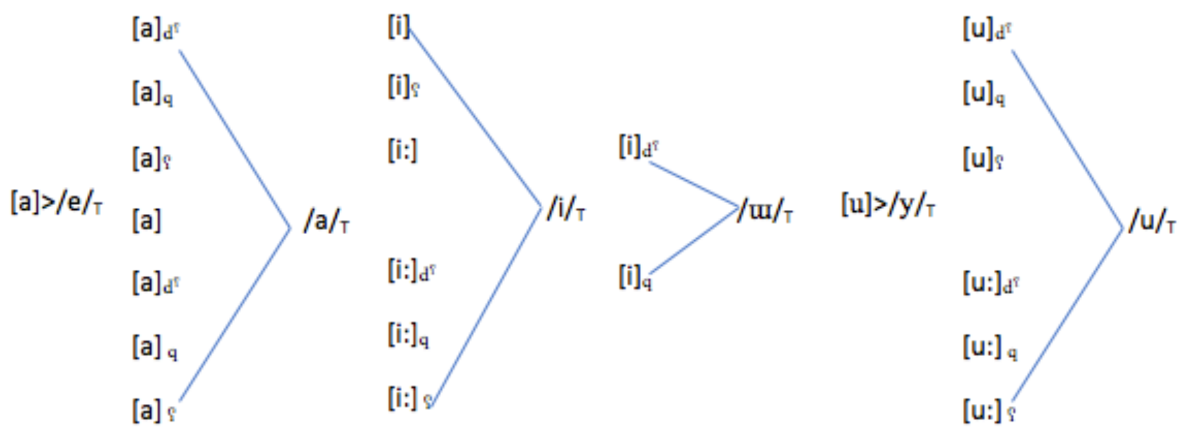


Figure 3-2: Corpus vowel mappings

Table 3-17 illustrates the mappings with words from the ALT corpus data, with short and long vowels in plain and guttural contexts. In the table, the term 'emphatic' is used to represent the whole guttural class and is not just limited to the four Arabic emphatic sounds (/sˤ/, /dˤ/, /tˤ/ and /ðˤ/). Furthermore, the vowels /aˤ/, /iˤ/ and /uˤ/ symbolize pharyngealized/uvularized vowels in the environment of all gutturals.

Context	*Vowel length*	a	i	u
Plain	*short*	i. a ---> e [dars~derṣ] 'lesson'	ii. i --->i [dʒin~dʒin] 'genie'	iii. u ---> y [ʃukr~ʃycyɾ] 'thankfulness'
	long	iv. a: ---> a [ʕan~an] 'instant'	v. i: --->i [fi:l~fil] 'elephant'	vi. u: ---> u [nuːr~nuɾ̥] 'light'
Emphatic	*short*	vii. ɑ/ aˤ ---> a [ðˤarf~zarf] 'envelope'	viii. iˤ ---> ɯ [tˤiˤp~tɯp] 'medicine'	ix. uˤ ---> u [tuħaf~tu̲haf] 'strange'
	long	x. a:ˤ ---> a [ʕaːr~aɾ̥] 'shame'	xi. i:ˤ ---> i [basˤi:ˤrah~basiret] 'foresight'	xii. u:ˤ ---> u [maðˤluːm~mazlum] 'wronged'

Table 3-17: Turkish adaptation of 12 Arabic vowels (allophones) in integrated loanwords

Table 3-17 and figure 3-2 show that Arabic long vowels were adapted as their short Turkish counterparts regardless of the presence or absence of gutturals in the Arabic word. That is, Arabic [a:], [a:ˤ], [i:], [i:ˤ], [u:] and [u:ˤ] are all adapted as the Turkish /a/$_T$, /i/$_T$ and /u/$_T$ (examples iv, x, v, xi, vi and xii respectively in table 3-17). The mapping of all Arabic long vowels to their counterpart Turkish vowels appears to be phonological in nature, and not sensitive to the phonetic detail of pharyngealized allophones of long vowels in the source word.

The adaptation of the Arabic short vowels is not as consistent as that of the long ones. On the one hand, the three Arabic short vowels [aˤ], [i] and [uˤ] are realized as their short Turkish counterparts, namely /a/$_T$, /i/$_T$ and /u/$_T$ (examples vii, ii and ix consecutively in table 3-17). For these three short vowels, then, phonological vowel quality appears also to be preserved.

On the other hand, the remaining three short Arabic vowels [a]$_A$, [iˤ]$_A$ and [u]$_A$ are adapted as the Turkish vowels /e/$_T$, /ɯ/$_T$ and /y/$_T$, i.e., where the phonological vowel quality of the Turkish vowel seems to be different from that of the Arabic vowel. The presence or absence of gutturals in the Arabic word thus seems to affect the vowel quality of the counterpart loanword Turkish vowel. For this second group of vowels then, we might want to argue that the loanword mapping is phonetic in nature, as it appears to be sensitive to the phonetic detail of the pharyngealized allophones of short vowels.

In fact, if we look at the acoustic phonetic realization of vowels in present day Arabic and Turkish, as illustrated in Figure 3-3 below, it is possible to argue that some of the mappings we have here tentatively characterized as phonological, could equally well be interpreted as phonetic.

Looking at figure 3-3, all of the Arabic long vowels, both plain (shown in green) and pharyngealized (shown in purple), are positioned in the vowel space *closer* to their Turkish counterpart vowel than to any other Turkish vowel, in terms of height (F1) and backness (F2). Thus, the adaptation of long Arabic vowels as their Turkish short counterparts could be argued to be both phonologically and phonetically motivated.

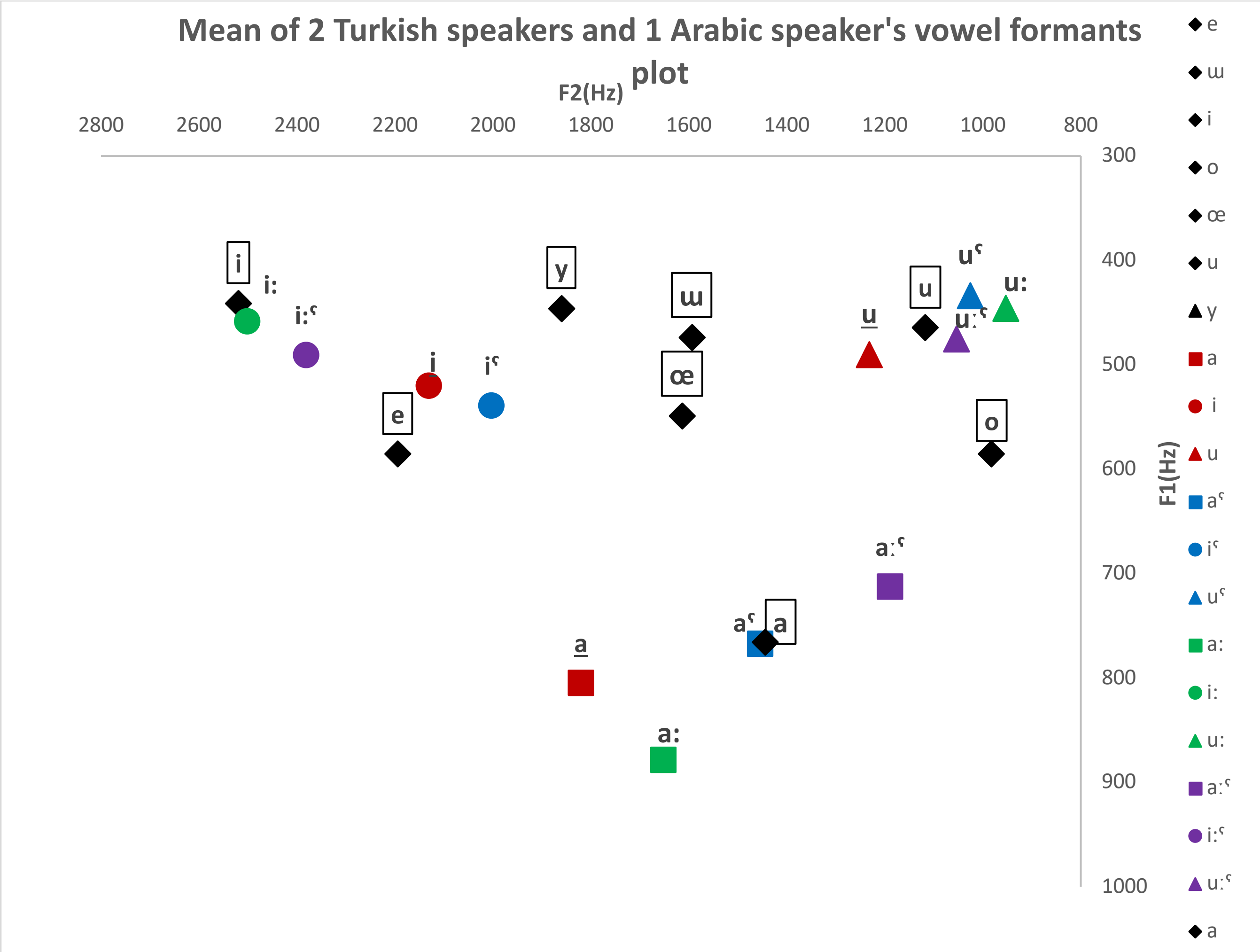


Figure 3-3: Vowel chart of Arabic and Turkish

Red = plain short Arabic vowels, **green**= plain long Arabic vowels, **blue**= emphatic short Arabic vowels, **purple**= emphatic long Arabic vowels and **black** diamond=Turkish vowels; circles = [i], squares = [a], triangles = [u]

Looking at the short vowels, however, the evidence for phonetic motivation of the corpus mapping is rather mixed. In the first group of short vowels, the short vowel $[a^{ʕ}]_A$ (in blue) is acoustically very similar to $/a/_T$ in terms of F1 and F2, and, similarly, the short vowel $[u^{ʕ}]_A$ (in blue) is acoustically close to its counterpart Turkish vowel $/u/_T$. However, short vowel $[i]_A$ (in red) is in fact acoustically closer to $/e/_T$, even though it is mapped to $/i/_T$ in the corpus.

For the second group of vowels, the short vowel $[a]_A$ (in red) is more fronted than long $[a]_A$ (in green), which might explain its mapping to $/e/_T$ even though it is acoustically somewhat closer to $/a/_T$ in terms of height (F1). The short vowel $[i^{ʕ}]_A$ (in blue) is similar in height (F1) to its counterpart vowel $/ɯ/_T$, but not in F2 (front/backness); it is in fact acoustically closer to

$/e/_T$, even though it is not mapped to $/e/_T$ in the corpus data. Finally, the short vowel $[u]_A$ (in red) is acoustically close to $/u/_T$ both F1 and F2 even though it is mapped to $/y/_T$ in the corpus data.

In sum, the adaptation of all of the long vowels and some of the short Arabic vowels can be explained by reference to the acoustic properties of Arabic and Turkish vowels, lending support to the Perceptual model of loanword adaptation. However, the role of the source language phonology is equally clear in the adaptation of the long vowels, and of the three short vowels where vowel quality is preserved ($[a^ʕ]_A$>$/a/_T$, $[i]_A$>$/i/_T$ and $[u^ʕ]_A$>$/u/_T$). This mixed picture suggests that we need a model which comprises both phonetics and phonology, and probably other factors such as orthography. The remainder of the thesis adopts a medial stance of loan phonology therefore, namely that most of the loanwords adaptations are phonetically grounded but with some effects of the source language phonology and orthography.

Thus far, the research has adopted a qualitative approach in stratifying the Arabic borrowed words which were adapted during two historical waves in addition to describing the resulting vowel mappings of these words. In the next chapters, specifically chapter four, five and six, a quantitative approach is used, in a series of perceptual studies conducted to test whether modern day Turkish speakers exhibit similar patterns to those seen in the corpus data.

Although the Osmanlica language is not used today, modern Turkish still displays patterns of the residual effects of gutturals neighbouring short vowels in Osmanlica. Thus, we can simulate how the speakers of Osmanlica perceived the Arabic vowels and borrowed them by analyzing how modern Turkish speakers assimilate Arabic vowels.

The rationale for the perceptual studies is based on the principle of uniformitarianism, which stipulates that the events, sound changes or processes used in a language that occurred in the past are the same as those applying nowadays (Murray, 2015). As such it is possible to simulate these sound changes in linguistic laboratories. Consequently, the perceptual study aims to simulate/model the grammar that the Ottomans employed when Arabic loanwords were borrowed in Turkish. This, in turn, would help in understanding characteristics related to language users' proficiency and bilingualism and, by extension, would help in establishing which of the various current loanword models can best account for the corpus patterns; the perceptual model (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992), phonological model (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001a, 2008; Peperkamp et al., 2008; Silverman, 1992) or a hybrid model of both perception and phonology (Kenstowicz & Suchato, 2006; Smith, 2006; Chang, 2003; Dolus, 2013).

The following chapters thus explore a number of research questions, the main ones of which are as follows. First, in chapter 4, how do Turkish language speakers categorize Arabic vowels into different Turkish categories? Second, in chapter 5, would speakers of Turkish language generalize the patterns of the effects of guttural consonants on neighbouring vowels to actual non-borrowed Arabic words? Third, in chapter 6, does the orthographic knowledge of Arabic language play a role in determining the quality of vowels neighbouring gutturals?

4 The Perceptual Study

Preamble:

The perceptual study tests the perception of the three Arabic short and long vowels in plain and emphatic contexts by representations of three groups of participants, namely Turkish only speaking participants with no/minimal knowledge of Arabic (**T**)[12]; Turkish and Arabic bilingual participants, who speak an Arabic dialect in Turkey in addition to Turkish (**TA**), and Turkish speaking participants with some knowledge of Arabic mainly from recitation of the Qur'an (**TQ**).

It comprises two main experiments, a Perceptual Assimilation Task (PAT) (Gilichinskaya & Strange, 2010; Strange et al., 1998) run as a listening task where respondents are asked to choose from a list of vowels and a Simulated Borrowing experiment (SB) conducted in three conditions: audio only, writing only, and a mixed condition of both audio and writing. The PAT experiment addresses the questions of how speakers map the vowels of another language to the nearest phonetic categories of their own language and whether knowledge of another language has an effect on the perception of the source vowels. The SB experiment tries to answer two main questions: 1) whether speakers of Turkish would generalize the residual effects of emphatics/gutturals on neighbouring vowels to real non-borrowed words and to nonsense words and 2) whether orthographic knowledge of Arabic affects perception. The two experiments combined ultimately try to gauge whether loanword adaptation takes place in *perception* (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992), *phonology* (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001a, 2008; Peperkamp et al., 2008; Silverman, 1992) or through some combination of both (Kenstowicz & Suchato, 2006; Smith, 2006; Chang, 2008; Dolus, 2013.

---

[12] Turkish only speaking participants with no/minimal knowledge of Arabic are referred to as such as they are already aware of the Arabic loanwords in Turkish. In addition, the parents of some of these speakers might be able to read/recite Quranic Arabic.

The Perceptual Assimilation Task (PAT)

The perceptual study was run online in two phases: one by the researcher herself using Qualtrics which provides online marketing survey tools over a period of 3 weeks, and another by the Qualtrics team. During the 1st phase, 54 participants attempted the survey; however, only 26 completed it. The other 28 respondents were screened out either for skipping some questions, not answering all different questions in the different blocks, or for supplying *gibberish* answers.[13] During the 2nd phase, 520 participants took the survey. Of these, 228 completed the survey while others were screened out for the same previously mentioned reasons.

In the Perceptual Assimilation Task (PAT), those who completed the survey were categorized as follows: Turkish (T), Turkish-Arabic (TA), and Turkish participants with Quranic Arabic knowledge (TQ). All three groups listened to recordings of all Arabic vowels including emphatic/guttural allophones produced by a native speaker Arabic, in monosyllabic nonsense words read in MSA/Classical Arabic. In addition, they were presented with some real Turkish vowels as distractors as a test of engagement with the test.

In the PAT experiment, the listeners were not told that the source vowels were Arabic and were asked 'what vowel did you hear?' on the computer screen and then had to choose from the set of all 8 Turkish vowels.

The PAT experiment addresses the question of how speakers map the vowels of another language to the nearest phonetic categories of their own language. This main research question is further divided into two sub-questions based on two contexts, namely vowel length and emphatic versus plain environment as set out below in (1).

(1) RQ1: How close is the perception of the listeners to the observed mappings in the qualitative corpus?

- How do Turkish participants perceptually assimilate the Arabic long vs. short vowels

[13] Some participants either did not consent to the survey's terms or had some technical issues with the audio files and hence their data were not incorporated. Gibberish here refers to supplying responses such as 'bbb' and 'zzzz', etc.

to the Turkish short vowels?
- How do Turkish participants perceptually assimilate the Arabic plain vs. emphatic vowels?

Another question probes whether knowledge of the phonology of the source language has an effect on the perception of the source vowels as stated in (2).

(2) RQ2: Does knowledge of Arabic (phonology) have any effect on perception?

The following sections are organized as follows. Section 4.1. presents the research hypotheses and predictions for the PAT patterns. Section 4.2. lays out the methodology followed including the stimuli, participants, procedure and rationale, data analysis and the results. Section 4.3. outlines the discussion of the data analysis. Section 4.4 concludes with a summary of the chapter.

4.1 *Hypotheses*

The main hypothesis is that the patterns of assimilation will match the mappings among vowels observed qualitatively in the research loanword corpus. This is the core of the evidence needed to support the ‘loanword adaptation as perception’ argument, i.e. that a mismatch between the source and target words occurs during the perception of foreign words (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992). The predicted assimilation patterns based on the corpus of 1118 words (introduced in chapter 3) are summarized in table 4-1.

Context	*Vowel length*	a	i	u
Plain	*short*	i. a ---> e [dars~deɾs] ‘lesson’	ii. i --->i [dʒin~dʒin] ‘genie’	iii. u ---> y [ʃukr~ʃycyɾ] ‘thankfulness’
	long	iv. a: ---> a [ʕan~an] ‘instant’	v. i: --->i [fi:l~fil] ‘elephant’	vi. u: ---> u [nu:r~nuɾ̥] ‘light’
Emphatic	*short*	vii. ɑ/ aˤ ---> a [ðˤarf~zaɾf] ‘envelope’	viii. iˤ ---> ɯ [tˤiˤp~tɯp] ‘medicine’	ix. uˤ ---> u [tuħaf~tu̠haf] ‘strange’
	long	x. a:ˤ ---> a [ʕa:r~aɾ̥] ‘shame’	xi. i:ˤ ---> i [basˤi:ˤrah~basiɾet] ‘foresight’	xii. u:ˤ ---> u [maðˤlu:m~mazlum] ‘wronged’

Table 4-1: Turkish adaptation of 12 Arabic vowels (allophones) in integrated loanwords

Table 4-1 shows that in the surrounding of all gutturals, i.e., emphatics, uvulars (/x/, /ɣ/ and /q/) and pharyngeals, $/a:/_A$, $[a:^ʕ]_A$ and $[a^ʕ]_A$ are assimilated as $/a/_T$ as in **iv**, **vii** and **x** but as $/e/_T$ surrounding plain consonants as in **i**. As for the Arabic vowel /i/, in the neighbourhood of q and emphatics only, it is mapped as $/ɯ/_T$ as in **viii** [$i^ʕ$>ɯ] but as $/i/_T$ elsewhere, /i>i/, /i:>i/ and [$i:^ʕ$>i] as in **ii**, **v** and **xi** in plain, pharyngeal and the two uvulars /x/ and /ɣ/'s surrounding. As for the vowel $/u/_A$, it is realized as $/u/_T$ in the environment of all gutturals but as $/y/_T$ elsewhere, i.e., plain setting.

4.2 Methodology

4.2.1 Stimuli

Stimuli were recorded during two sessions in the data lab in the Department of Language and Linguistic Science at the University of York. Both recordings were carried out using a Neumann U87i microphone, a TAC Scorpion mixing desk, M-Audio 24/96 Audio card and Adobe Audition CS5.5 on Windows 7 Pro x64 on PC software with 44.1 khz 16 bit sampling rate.

In the first session, a set of 12 Arabic monosyllabic nonsense words of the 12 Arabic vowel allophones (3 short and plain, 3 short and emphatic, 3 long and plain, 3 long and emphatic) were recorded by a native Arabic speaker of Syrian origin from Aleppo[14]. The 12 Arabic nonsense words were used as stimuli in the PAT experiment. A list of these words is illustrated in table 4-2 below. Then another set of words — real and nonsense — were recorded by the same speaker in monosyllabic and polysyllabic words for use in the Simulated Borrowing experiment and to plot the vowel formants of Arabic for the acoustic bases used in the discussion section.

In the second session, a native Turkish speaker from Ankara and another from Gaziantep were recorded reading a set of 48 words of the 8 Turkish vowels x 6 words per vowel in real monosyllabic words. A list of these words is given in table 4-3 below. These Turkish

[14] The choice of the speaker is explained in chapter 2.

recordings were used only to plot the vowel acoustics of Turkish as explained in chapter 2. In addition, 4 words with cardinal vowels close to the Turkish vowels /y, ɯ, œ and eː/ were used as distractors and recorded by the Turkish speaker from Gaziantep. All words were repeated 3 times per allophone totaling to 36 tokens for the Arabic words and 12 tokens for the distractors. The order of presentation of Arabic and Turkish words was also randomized and the listeners were informed that they might hear some words repeated more than once.

All the speakers of Arabic and Turkish were asked to read the lists of words row by row and to maintain an even pitch (tone) throughout.

S.no	***Arabic Vowel Category (stimulus vowel)***	***Stimulus nonsense word***	***Arabic nonsense word***	***Similar to target word***	***Target Arabic Real word***	***English glossary***
1	/a/	/had/ had	هَد	/hadd/	هدّ	He destroyed
2	/i/	/hɪd/ hid	هِد	/hɪd/	هِد	Destroy! (imperative)
3	/u/	/hʊd/ hud	هُد	/ʃʊd/	شُدّ	Pull! (imperative)
4	/a:/	/ha:d/ haad	هَاد	/dʒa:dd/	جادّ	Serious
5	/i:/	/hi:d/ hiid	هِيد	/dʒi:d/	جيد	Neck
6	/u:/	/hu:d/ huud	هُود	/hu:d/	هود	A prophet's name/ a proper name
7	/aˤ/	/haˤdˤ/ haDD	هَض	/haˤdˤ/	هَضّ	A fracture
8	/iˤ/	/hiˤdˤ/ hiDD	هِض	/hiˤdˤ/	هِضّ	Break (imperative)
9	/uˤ/	/huˤdˤ/ huDD	هُض	/huˤdˤuˤm/	هُضُم	Any substance that helps digest food
10	/a:ˤ/	/ha:ˤdˤ/ haaDD	هَاض	/ha:dˤdˤ/	هَاض	Broke and fractured
11	/i:ˤ/	/hi:ˤdˤ/ hiiDD	هِيض	/jahi:dˤ/	يَهِيض	He is breaking/fracturing
12	/u:ˤ/	/hu:ˤdˤ/ huuDD	هُوض	/nuˤhu:ˤdˤ/	نُهُوض	Rising

Table 4-2: The Arabic stimulus material for the PAT experiment: 12 hVd nonsense words with the Arabic plain and emphatic vowels long and short[15]

[15] In the table above, DD is used to denote an emphatic. This shorthand is used since the software R does not allow usage of IPA fonts.

S.no	Turkish word 1	Turkish word 2	Turkish word 3	Turkish word 4	Turkish word 5	Turkish word 6
1	had /had/ 'limit'	Can /dʒan/ 'soul'	dar /dar/ 'tight'	bal /bal/ 'honey'	kap /kap/ 'cover'	bağ /baa/ 'orchard'
2	hiç /hitʃ/ 'never'	bir /bir/ 'one'	diş /diʃ/ 'tooth'	fil /fil/ 'elephant'	bin /bin/ '1000'	Kim /kim/ 'who'
3	huy /huj/'nature'	çul /tʃul/ 'clothes'	muz /muz/ 'banana'	dul /dul/ 'widow'	sur /sur/ 'fence'	kum /kum/ 'sand'
4	hem /hem/ 'Both, and'	geç /getʃ/ 'late'	çek /tʃek/ 'check'	şef /ʃef/ 'chief'	ben /ben/ 'I'	her /her/ 'each'
5	çöl /kœl/ 'desert'	kör /kœr/ 'blind'	göz /gœz/ 'eye'	yön /jœn/'side'	köy /kœj/ 'village'	gök /gœk/ 'sky'
6	hol /hol/ 'hall'	doz /doz/ 'dose'	toy /toj/ 'immature'	mor /mor/ 'purple'	boş /boʃ/ 'empty'	fon /fon/ 'fund'
7	hız /hɯz/ 'speed'	kın /kɯn/ 'sheath'	yıl /jɯl/ 'year'	sır /sɯr/ 'secret'	dış /dɯʃ/ 'outside'	tıp /tɯp/ 'medicine'
8	hür /hyr/ 'free'	düş /dyʃ/'dream'	yüz /jyz/ '100'	kül /kyl/ 'ash'	gün /gyn/ 'day'	süt /syt/ 'milk'

Table 4-3: The Turkish stimulus material for Turkish acoustics

4.2.2 Participants

Two hundred and twenty eight (228) participants born in different parts of Turkey took part in the Perceptual Assimilation Task to represent three groups: 41 monolingual Turkish (T) speakers, 44 bilingual Turkish-Arabic (TA) speakers, and 143 Turkish speakers with some Arabic knowledge through Quran recitation (TQ). The choice of the three groups stems from the fact that the Ottomans, the original borrowers of Arabic words, spoke both Ottoman Turkish and Arabic (as well as Persian). The selection of the participants was done painstakingly. The T group (41 participants) included only monolingual Turkish speakers, meaning that Turkish speakers who spoke other languages such as German, Hebrew, Spanish, Kurdish, Armenian, Tatar, English or any other language were excluded from the survey. The language criterion was controlled for in the study via questions about the participants 1st, 2nd and 3rd language, their parents', which language the participants speak at home and whether they know Arabic from reciting the Quran (Appendix 4-1: language questionnaire). Thus, the T group are considered as 'naïve listeners', with no Arabic knowledge other than being able to recognize some of the borrowed Arabic words in Turkish.

The TA group of participants (44 participants) are fluent in both Turkish and Arabic and learned both since childhood. In the survey, these are the ones who indicated that they learned Arabic since childhood, whose either or both parents were bilingual speakers and who spoke Arabic in addition to Turkish at home. Second language learners of Arabic were not considered in the TA group. Most TA participants were born in provinces where Arabic is spoken as a dialect in Turkey such as in Hatay (Antakya, Iskendrun), Adana, Mersin, Gaziantep in addition to other cities where Turkish is more dominant such as Ankara, Istanbul, Izmir, düzce, Malatya, Sivas, Giresun and Denizli.

As for the TQ group, these are the majority of the participants (143). They are the ones who indicated (in the survey) that they knew Arabic through reading/reciting the Qur'an, used Arabic for a number of years and rated their proficiency skills in Arabic as sufficient enough to recognize Arabic characters despite not being day-to-day users of Arabic. Many of them were proficient second language users since they learned it prescriptively in Quranic schools since childhood -at the age of 7/8 and as part of high school and university training.

All participants were asked to rate their Turkish skills and the TA and TQ groups were also asked to rate their Arabic skills on a scale of 0 to 10 with 0 being poor and 10 being good. The participants come from different educational backgrounds ranging from high school (HS), university graduates (C), Master's holders (M), PhD holders (PhD) and others (O) including Vocational and Technical training. All participants have had a formal education during school for at least 12 years. The table below (table 4-4) summarizes the participants' sociolinguistic and language background.

Participant	Age	Gender	Highest level of education	Turkish Reading	Turkish comprehension	Turkish Speaking	Turkish Writing	Average	Arabic Reading	Arabic comprehension	Arabic Speaking	Arabic Writing	Average
T	17-56	22 m: 20 f	12 HS, 50 C, 2 M, 2 PhD, 1 other	9.8	9.7	9.8	9.8	9.8	N/A	N/A	N/A	N/A	N/A
TA	18-54	31 m:14 f	11 HS, 50 Co, 7 M, 1 PhD, 1 other	9.3	9.5	9.4	9.4	9.4	6.1	5.9	5.7	3.6	5.3
TQ	17-62	98 m: 43 f	33 HS, 172 C, 57 M, 0 PhD, 3 Other	9.5	9.5	9.4	9.3	9.4	6.5	3.3	2.8	3.7	4.1
Average	27	-	-	9.5	9.5	9.5	9.4	9.5	6.3	4.6	4.3	3.7	4.7

Table 4-4: Summary of the language questionnaire results of the Turkish participants

4.2.3 Rationale and procedure

The rationale behind the PAT is to quantitatively test whether the mappings made by the participants are based on auditory perception, phonological knowledge of Arabic or even an interaction of both. The results would help formalize which model to adopt in analyzing the larger corpus of 1118 Arabic loanwords in Turkish.

The PAT experiment was run online through Qualtrics portal. Before performing the listening experiment, the participants were informed in the information sheet that by participating in this study, they were helping and contributing to knowledge since group answers would be analyzed to see patterns and verify hypotheses on the sound system of Turkish. Then, they were asked to tick in the consent form before proceeding to the PAT experiment (Appendix 4-2). After the experiment, the respondents were asked to fill in the language and background questionnaire (appendix 4-1) before proceeding to the next experiment; i.e., Simulated Borrowing experiment. Noteworthy to mention is that instructions for the experiment were translated into Turkish — by a certified Turkish translator.

The procedure of the on-line listening experiment involved the participants listening to an Arabic vowel variant contextualized in a monosyllabic nonsense word of the shape **hVd** read by an Arabic speaker who is a trained phonetician, and then selecting the Turkish vowel closest to it. Only 12 Arabic words with the 12 Arabic vowel allophones in plain vs emphatic and short vs long environments were used; however, the participants were not told that the source vowels included Arabic vowels. In addition, 4 words with cardinal vowels close to the Turkish vowels /y, ɯ, œ and eː/ were used as distractors. All words were repeated 3 times per allophone totaling to 36 tokens for the Arabic words and 12 tokens for the distractors. The order of presentation of Arabic and Turkish words was also randomized. The listeners were informed that they might hear some words repeated more than once. Repeating the words three times was used as a goodness of fit scale with a score of 1 out of 3 interpreted as being not confident (poor), 2 confident (good) and 3 very confident (very good). Figure **4-1** below shows a screen shot of the PAT as presented to the participants.

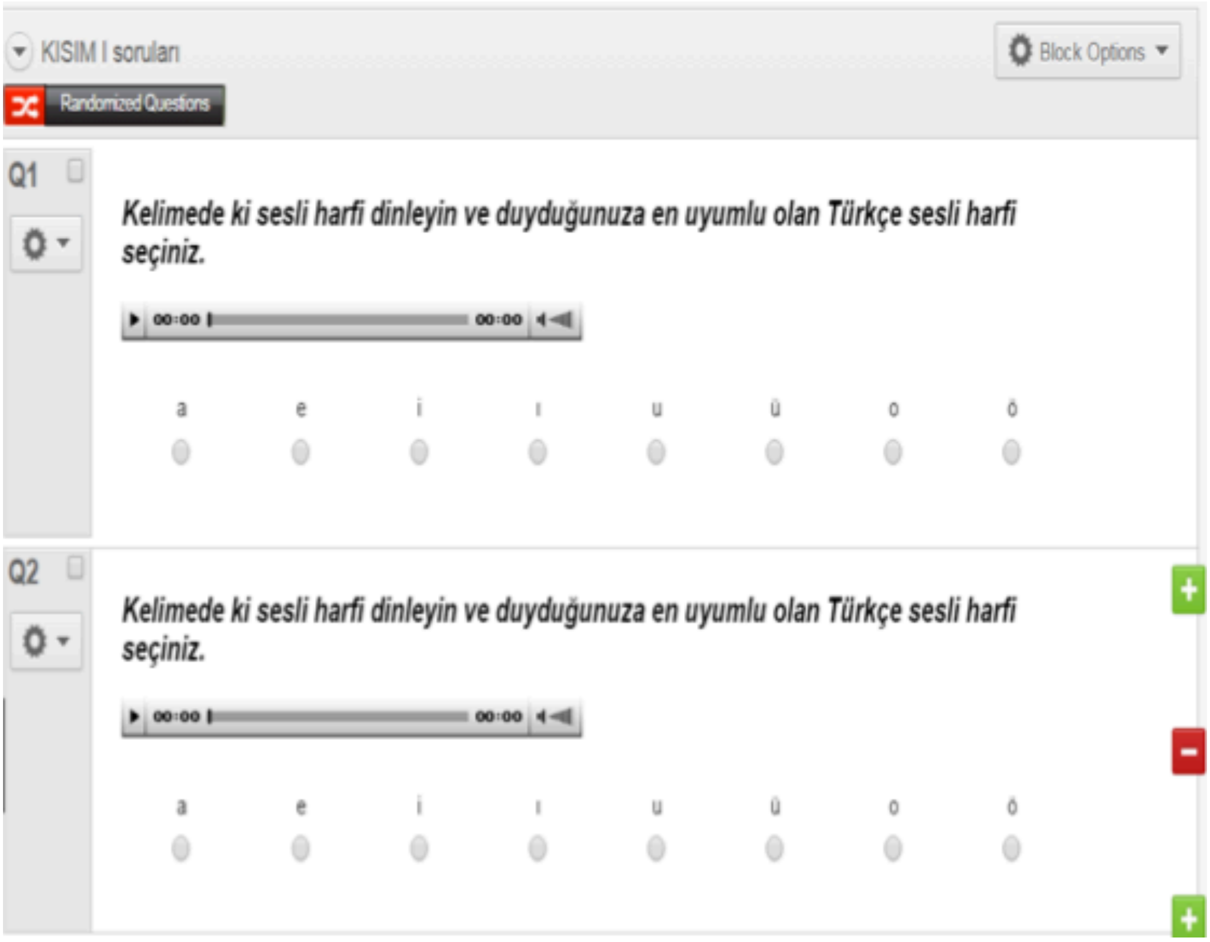


Figure 4-1: A screen image of the PAT experiment

The instructions in Turkish translates as 'in the words you hear, choose the closest vowel that most resembles Turkish vowels.'

<u>Procedure used in the Logistic regression and model selection</u>

In the logistic regression part (4.3.2.) in the data analysis section (4.3.), the analysis is carried out in two stages. First, the data sample size including the number of observations, the objective of the logistic regression and/hypotheses underpinning the analysis are presented. Second, a model selection protocol adapted from Zuur et al (2009), Winter (2013), Baayen (2008) and Barr et al (2013) is followed. Three main steps are involved in the protocol; i) determining the fixed and random effects structures, ii) creating the maximal or beyond optimal model and running regression models, and iii) validating the model of the best fit.
This design-driven approach is transparent in being easy to follow and being capable of addressing the complex nature of the data under study whereas each of the existing approaches in the literature (Zuur et al, 2009; Winter, 2013 and Baayen, 2008) can address particular parts of the analysis but not all.

4.3 Data Analysis

4.3.1 Raw data

This part presents the trends most salient in the PAT raw data. This is mainly done through inspection of confusion matrices, tables and barplots in Microsoft Excel and R software (Team, 2015). First, a summary table of the raw data of each of the listener groups is given in table 4-5 using the pivot table feature in excel. In the table, stimulus words (***stimulus***) appear in the first column, listener group (***Listgp***) in the first row, and response vowels (***RV***) in the second row. The numbers below each of the response vowels represent actual categorization tokens. When the observed vowel is the same as the predicted, this is considered a match and is shaded in green. On the other hand, when the most frequent response vowel is not the same as the predicted vowel, this is a mismatch which is shaded in red.

Listgp	T								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	**130**	**99**	**3**		**3**		**10**	**1**	**246**
had	20	98	1		2		1	1	123
haDD	110	1	2		1		9		123
a:	120	106	6	3	2	1	8		246
haad	10	106	3	1	2	1			123
haaDD	110		3	2			8		123
i	**4**	**81**	**124**	**15**	**7**	**6**		**9**	**246**
hid	1	2	110	5	3	2			123
hiDD	3	79	14	10	4	4		9	123
i:	3	3	229	6	3		2		246
hiid	1	1	115	3	2		1		123
hiiDD	2	2	114	3	1		1		123
u	**3**	**2**	**6**	**7**	**218**	**8**	**2**		**246**
hud	2	1	5	4	102	8	1		123
huDD	1	1	1	3	116		1		123
u:	4		2	5	229	4	2		246
huud	1		2	2	115	2	1		123
huuDD	3			3	114	2	1		123
Total	**264**	**291**	**370**	**36**	**462**	**19**	**24**	**10**	**1476**

Listgp	TA								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	**152**	**88**	**4**	**3**	**1**	**4**	**7**	**5**	**264**
had	35	88	1	3		2		3	132
haDD	117		3		1	2	7	2	132
a:	132	112	7	2	3	1	7		264
haad	14	110	5	2		1			132
haaDD	118	2	2		3		7		132
i	**6**	**72**	**129**	**30**	**7**	**1**	**2**	**17**	**264**
hid	1	7	111	10	3				132
hiDD	5	65	18	20	4	1	2	17	132
i:	2	4	236	7	5	6	2	2	264
hiid	2	2	119	2	2	4		1	132
hiiDD		2	117	5	3	2	2	1	132
u	**2**	**5**	**8**	**7**	**221**	**20**	**1**		**264**
hud	2	1	5	5	101	18			132
huDD		4	3	2	120	2	1		132
u:	4	4	6	2	232	6	8	2	264
huud	3		2	1	119	4	1	2	132
huuDD	1	4	4	1	113	2	7		132
Total	**298**	**285**	**390**	**51**	**469**	**38**	**27**	**26**	**1584**

Listgp	TQ								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	**441**	**345**	**11**	**12**	**17**	**12**	**15**	**5**	**858**
had	63	335	7	8	9	3	3	1	429
haDD	378	10	4	4	8	9	12	4	429
a:	408	372	11	16	9	9	19	14	858
haad	35	363	5	6	6	3	3	8	429
haaDD	373	9	6	10	3	6	16	6	429
i	**21**	**252**	**394**	**99**	**16**	**23**	**13**	**40**	**858**
hid	6	15	364	26	3	9	5	1	429
hiDD	15	237	30	73	13	14	8	39	429
i:	13	20	760	24	20	13	4	4	858
hiid	7	8	386	9	8	6	2	3	429
hiiDD	6	12	374	15	12	7	2	1	429
u	**13**	**12**	**18**	**43**	**697**	**53**	**12**	**10**	**858**
hud	9	5	11	29	322	45	3	5	429
huDD	4	7	7	14	375	8	9	5	429
u:	12	14	18	16	745	14	31	8	858
huud	3	6	13	9	377	12	5	4	429
huuDD	9	8	5	7	368	2	26	4	429
Total	**908**	**1015**	**1212**	**210**	**1504**	**124**	**94**	**81**	**5148**

Table 4-5: Summary confusion matrix of the PAT results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)
Red shade indicates 'most frequent' response vowel≠predicted vowel (mismatch)
SV= stimulus vowel; RV= response vowel; Listgp= listener group
T= monolingual listeners; TA= bilingual listeners; TQ= Turkish Quranic listeners
In the table above, DD is used to denote an emphatic. This shorthand is used since the software R does not allow usage of IPA fonts.

Table 4-5 above demonstrates that the three listener groups (T, TA and TQ) exhibit similar perceptual patterns, reflected by the green and red shaded cells which are the same for each group. Moreover, the only three vowels which are mismatched are /a:/>/e/ (mismatched 91%), /iˤ/>/e/ (85%) and /u/>/y/ (90%) as shown by table 4-6.

		match %	
St.vowel	**Stimulus**	**Match**	**Mismatch**
/a/	had	76.169591	23.83040936
/a:/	**haad**	8.625731	**91.37426901**
/aˤ/	haDD	88.450292	11.5497076
/a:ˤ/	haaDD	87.865497	12.13450292
/i/	hid	85.526316	14.47368421
/i:/	hiid	90.643275	9.356725146
/iˤ/	**hiDD**	15.05848	**84.94152047**
/i:ˤ/	hiiDD	88.450292	11.5497076
/u/	**hud**	10.380117	**89.61988304**
/u:/	huud	89.327485	10.67251462
/uˤ/	huDD	89.327485	10.67251462
/u:ˤ/	huuDD	86.988304	13.01169591

Table 4-6: crosstabs of match per stimulus vowel and match

Furthermore, match results in percentile across the three listener groups are given in table 4-7 and plotted in figure 4-2.

	Match		Mismatch	
Listgp	%	Count	%	count
T	69.783	1030	30.216	446
TA	67.803	1074	32.196	510
TQ	67.6573	3483	32.342	1665
average	68.414	1862.333	31.5855	873.666

Table 4-7: crosstabs of match results across Listgp

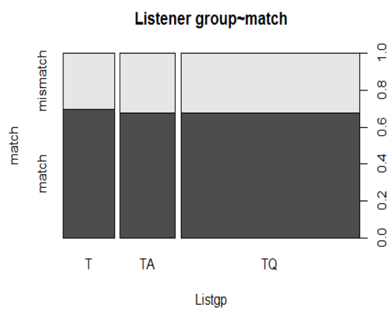


Figure 4-2: barplot of match~Listgp

Thus far, we can draw the following observations based on tables 4-5, 4-6, 4-7 and figure 4-2.

1. The percentage of match across the three listener groups is high at 68.4%.
2. Overall, the three groups (T, TA and TQ) manifest the same patterns for all the vowels. However, the T group slightly demonstrates more match at 69.7% than the two other groups at 67.8% for the TA group and 67.6% for the TQ groups. The last two groups more

or less reflect similar mapping patterns which indicates that knowledge of Arabic does not play a role in the categorization.

3. The average of match is higher among long vowels; [a:ˤ]>/a/, /i:/>/i/, [i:ˤ]>/i/, /u:/>/u/ and [u:ˤ]>/u/. In addition, some short vowels are matched to their predicted equivalents in Turkish; /a/>/e/, [aˤ]>a, /i/>/i/, [uˤ]>/u/.
4. Mismatch occurs among short vowels, namely [iˤ]>/e/, /u/>/y/. Moreover, the /a:/ is the only long vowel which is mismatched in the PAT experiment; /a:/>/e/.

In addition, table 4-8 and figures 4-3 and 4-4 show the percentage and count of match results across and within the three listener groups in the PAT task split by vowel quality[16]. The following observations can be made.

		Match		Mismatch	
Vowel	**Listgp**	**%**	**count**	**%**	**Count**
a	T	66.666	328	33.333	164
	TA	63.825	337	36.174	191
	TQ	65.326	1121	34.673	595
	average	65.272	595.333	34.727	316.666
i	T	70.934	349	29.065	143
	TA	69.507	367	30.492	161
	TQ	69.755	1197	30.244	519
	average	70.065	637.666	29.934	274.333
u	T	71.747	353	28.252	139
	TA	70.075	370	29.924	158
	TQ	67.89	1165	32.109	551
	average	69.904	629.333	30.095	282.666

Table 4-8: crosstabs of vowel.quality+Listgp~match

1. Across the three vowels, the categorization of /i/ and /u/ yielded more matching than /a/ at 70%, 69.9% and 65% respectively; these match percentages are still considerably high.
2. Within each of the three vowels, the T group displayed higher percentage of match than the TA and TQ groups while more variation is found in the TQ group for each of the three vowels (figure 4-3).
3. For the vowel /a/, the T group achieved higher match results at 67%, followed by the TQ at 65% and the TA at 64%.

[16] Vowel quality is used here and in the logistic regression analysis to refer to the three Arabic vowel phonemes /a/, /i/ and /u/ and does not relate to the spectral correlates (F1 and F2) of the vowels as is often used in acoustics.

4. For the vowel /i/, again the T group triggered higher match rates at 71% followed by the TQ and TA whose results were quite similar at 69.7% and 69.5%.
5. *For the vowel /u/, again the T group yielded higher match results at 72% followed by the TA group this time at 70% and the TQ at 68*

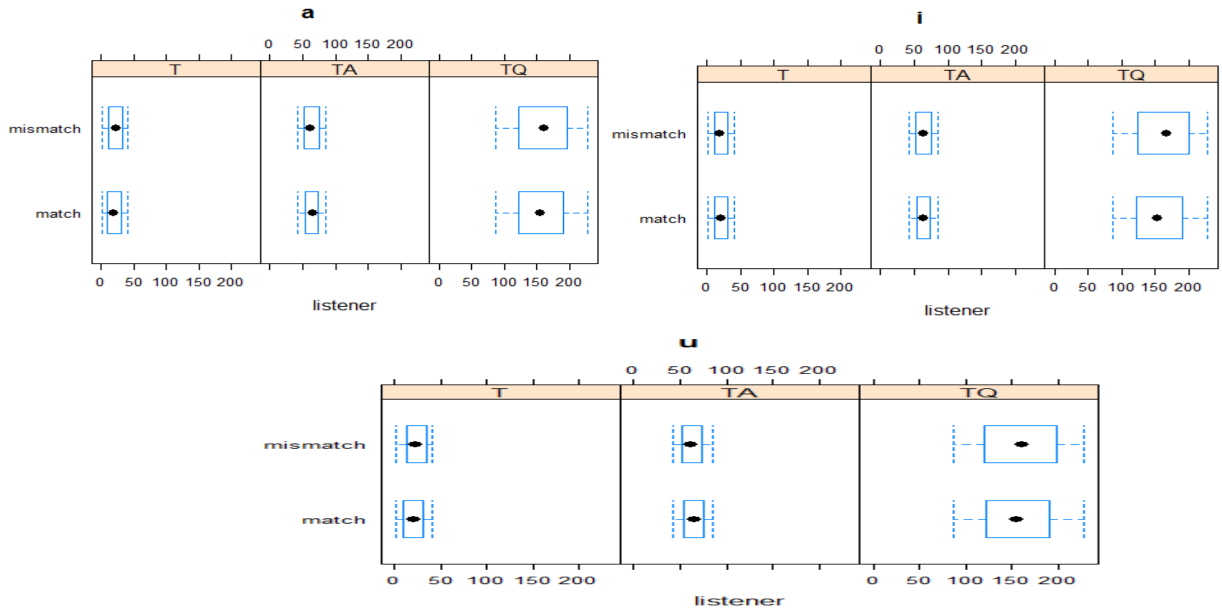


Figure 4-3: Bwplot of match~listener

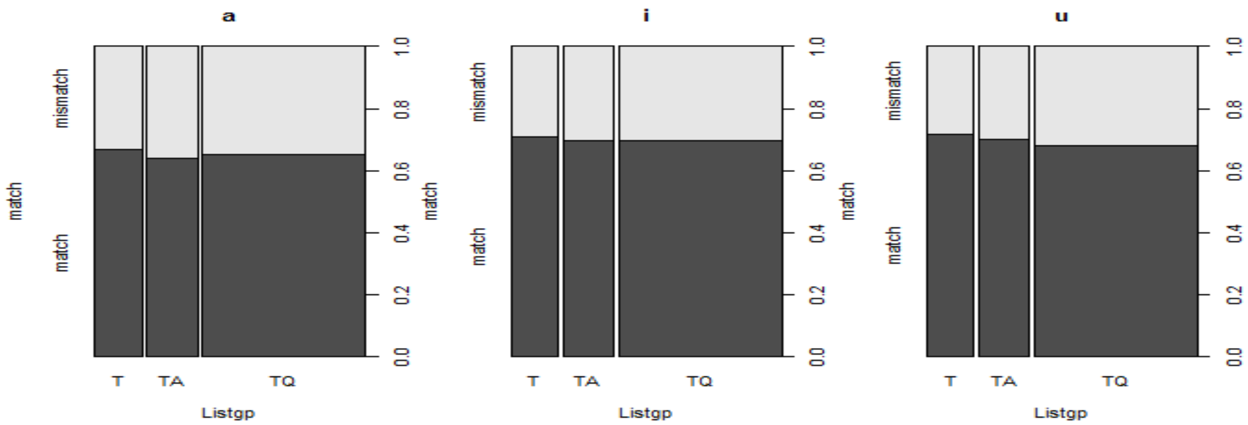


Figure 4-4: barplot of vowel.quality~match

In addition to the previous tables, the patterns found in the PAT are furnished in tables (4-9 to 4-11) below where vowels contrast for length, indicated with (:); and context (-/+emphasis) where emphasis is taken here to represent all gutturals, and the vowels surround one of the nine Arabic guttural consonants [ʕ, ħ, x, ɣ, q, s^ʕ, d^ʕ, t^ʕ and ð^ʕ]. The symbol used for emphasis is a superscripted ^{ʕ}, so [a^ʕ], for example, symbolizes the Arabic short guttural low back vowel. Turkish vowels are all phonemes shown with a subscripted $_T$ and between // in contrast to the Arabic allophones which are shown between []. Moreover, two types of mappings are reflected; near categorical and variable. Near categorical mappings are cases where most percentages are near to 100% (80-100%). These are represented with solid arrows in tables 4-9 through 4-11. On the other hand, variable mappings are percentages ranging between 50-80% and are indicated with dashed arrows.

had [a]--> (98/123) /e/$_T$ (79.67%) haad [a:]→(106/123) /e/$_T$ (86.17%)	Had [a^ʕ]→(110/123) /a/ (89.43%) haaDD [a:^{ʕ}]→(110/123) /a/ (89.43%)	Hud [u]→(102/123) /u/ (82.92%) hu:d [u:]→(115/123) /u/ (93.49%) huDD [u^ʕ]→(116/123) /u/ (94.3%) hu:DD [u:^{ʕ}]→(114/123) /u/ (92%)	hid [i]→(110/123) /i/ (89.43%) hi:d [i:]→(115/123) /i/ (93.4%) hi:DD [i:^{ʕ}]→(114/123) /i/ (92%)	hiDD [i^ʕ]--> (79/123) /e/ (64.22%)

Table 4-9: T group perceptual PAT maps

had [a]--> (88/132) /e/$_T$ (66.66%) haad [a:]→(110/132) /e/$_T$ (83.33%)	Had [a^ʕ]→(117/132) /a/ (88.63%) haaDD [a:^{ʕ}]→(110/123) /a/ (89.43%)	Hud [u]-->(101/132) /u/ (76.51%) hu:d [u:]→(119/132) /u/ (90.15%) huDD [u^ʕ]→(120/132) /u/ (90.90%) hu:DD [u:^{ʕ}]→ (113/132)/u/ (85.60%)	hid [i]→(111/132) /i/ (84.09%) hi:d [i:]→(119/132) /i/ (90.15%) hi:DD [i:^{ʕ}]→(117/132) /i/ (88.63%)	hiDD [i^ʕ]--> (65/132) /e/ (49.24%)

Table 4-10: TA group perceptual PAT maps

had [a]--> (335/429) /e/$_T$ (78.08%) haad [a:]→(363/429) /e/$_T$	Had [a^ʕ]→(378/429) /a/ (88.11%) haaDD	Hud [u]-->(322/429) /u/ (75.05%) hu:d [u:]→(377/429) /u/	Hid [i]→(364/429) /i/ (84.84%) hi:d [i:]→(386/429) /i/	hiDD [i^ʕ]-->(237/429) /e/ (55.24%)

(84.61%)	[a:ˤ]→(373/429) /a/ (86.94%)	(87.87%) [uˤ]→(375/429) /u/ (87.41%) hu:DD [u:ˤ]-->(368/429) /u/ (85.78%)	(89.97%) hi:DD [i:ˤ]→(374/429) /i/ (87.17%)

Table 4-11: TQ group perceptual PAT maps

4.3.2 Logistic regression of PAT data

Since the experiment design is complex where the dependent variable, i.e., match is categorical and binary, and the effects involved are a mix of fixed and random, the data were analyzed using logistic regression, specifically Generalized Logistic Mixed[17] effects Modelling (GLMM) using the software R (Team, 2015). In the PATdata set, the sample consisted of 8208 observations of 228 listeners distributed among (**T**: 41, **TA**: 44 and **TQ**: 143). This means that each listener yielded 36 responses (8202/228= 36) with the T group yielding 1476 (41X36), TA rendering 1584 (44X36) and TQ 5148 (143X36). The stimulus related effects included stimulus, st.vowel, vowel quality, context and length. For each of the three vowel qualities, there were 2736 observations (8208/3). Each stimulus vowel (12 vowels) was heard 684 times (8208/12), with each phoneme being played 1368 times (684X2). The 8208 observations were divided between plain and uvularized (4104 each) and long and short vowels (4104) each. The total number of match is 5587 compared to 2621 mismatch. The listeners ranged in age between 17 and 62.

The dataset used (PATdata) consist of The listener-related variables include listener group (***Listgp: T, TA*** *and* ***TQ***) and ***age*** as a control variable. On the other hand, the stimulus related variables include stimulus consonant[18] context (***context: emphatic, plain, pharyngeal and q***), stimulus vowel length (***length: long and short vowels***) and stimulus **vowel.quality** (a, i or u). The two variables ***stimulus*** and ***listener***[19] are treated as random effects. The measures of association/descriptives of these variables are given in table 4-12 below.

[17] The use of the mixed modelling is due to the fact that some variables are fixed while others are random.
[18] Consonant refers to the type of consonant that precedes the stimulus vowel.
[19] The variable listener is taken here as a random effect since the data was sampled randomly and the variable itself has +100 levels. As for the variable stimulus, despite the fact it has only 12 levels and that it was not

Listgp	Listener	vowel.quality	st.vowel	stimulus
T :1476	T1 : 36	a:2736	a :1368	haad : 684
TA:1584	T10 : 36	i:2736	aa:1368	haaDD : 684
TQ:5148	T11 : 36	u:2736	i :1368	had : 684
	T12 : 36		ii:1368	haDD : 684
	T13 : 36		u :1368	hid : 684
	T14 : 36		uu:1368	hiDD : 684
	(Other):7992			(Other):4104
context	**Length**	**Match**	**age**	
plain :4104	long :4104	match :5587	Min. :17.00	
uvularized:4104	short:4104	mismatch:2621	1st Qu.:24.00	
			Median :29.00	
			Mean :31.02	
			3rd Qu.:37.00	
			Max. :62.00	
			NA's :36	

Table 4-12: Measure of association of the PATdata dataset

4.3.2.1 Objective and hypotheses

The main question pursued here is to explain the match response computed by comparing observed versus expected values given the different listener and stimulus-related predictors. Furthermore, is the match dependent on predictors with significant p-values and coefficient estimates and/or interactions among these predictors?

The two hypotheses derived from RQ1 in 1. include the experimental and null hypotheses as given below.

- **H1:** The variables ***length***, ***context***, ***vowel.quality*** and ***Knowledge of Arabic*** (as represented in ***Listgp***: T (monolingual Turkish speakers) and TA and TQ (being groups with access to Arabic)) have effects on the DV match.
- **H0** The variables ***length***, ***context***, ***vowel.quality*** and ***Listgp*** do not have effects on the DV match.

sampled randomly, R displayed many warning messages when the variable was used as a fixed effect. In addition, it is different than the usual sociolinguistic variables such as age since it is specific to the current research data.

In R modelling, the family used is binomial (logit) and the proportions for the number of the given observations are weighted using this link function. Moreover, since the data is not normally distributed as the dependent variable (match) is binary (match/mismatch), the mixed model is fit by maximum likelihood criterion (Laplace Approximation as there are fewer than five random effects).

The basic R model used to fit the current data is produced below as base.Listgp where ***match*** is a binary categorical variable dependent on the fixed effect **Listgp**, being the variable of main experimental interest, which is tied to the random effects ***stimulus*** and ***Listeners***. The generalized linear mixed-effects family used here for the dataset PAT is "***binomial***" with the function link ***logit***. The basic or reduced model is given below along with a summary table (table 4-13) of the fixed effects fit measures.

base.Listgp<- glmer(match~ Listgp +(Listgp|stimulus) + (1|listener), data = PAT, family = "binomial")

Fixed effects:	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.1100	0.6188	-1.794	0.0728 .
ListgpTA	0.1905	0.1747	1.090	0.2756
ListgpTQ	0.2149	0.1590	1.351	0.1766
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table 4-13: Fixed effects table of base.Listgp model

As can be seen, the intercept is significant at a p-value of 0.07, a value equal to that of the T group since the T level is the base/reference level of the variable Listgp which is embedded in the intercept. We have already seen from the raw data analysis in section 4.3.1 that the T group achieved higher match scores compared to the other two groups, hence, the significant p-value in the model above.

4.3.2.2 Protocol for model selection

Step i: Building the structures of random and fixed effects

The first step of the protocol as mentioned in section 4.2.3 involves building both the fixed and random effects structures. In our context, the variables of main interest include ***Listgp***, ***length***, ***context*** and ***vowel.quality*** whereas ***age*** is treated as a control variable. These effects correlations with the match variable (DV) are plotted in figure 4-5.

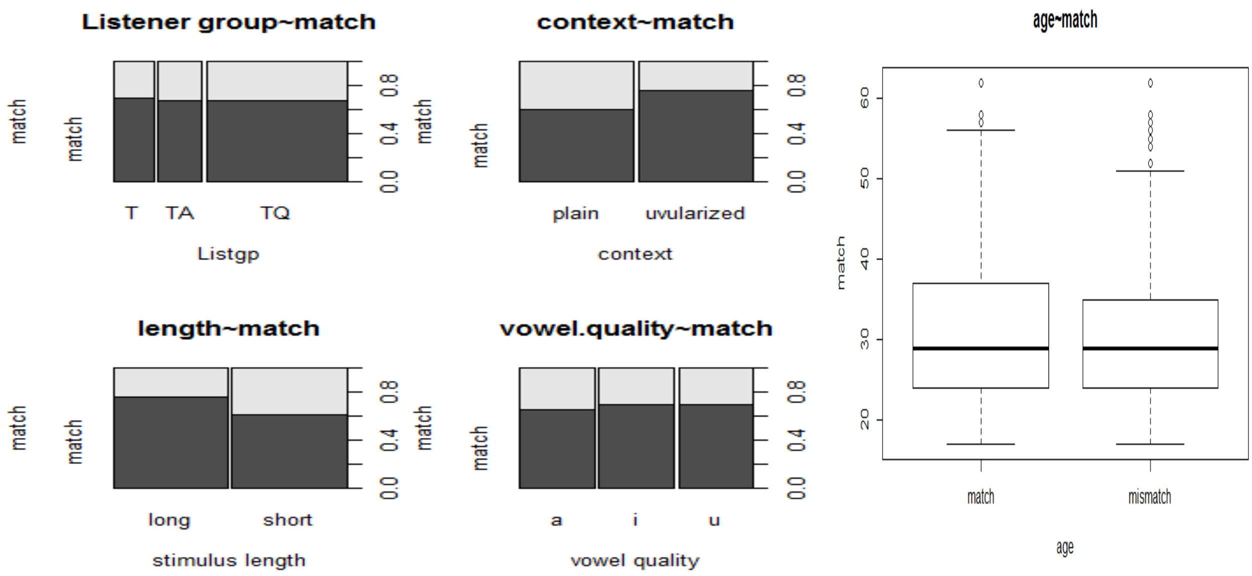


Figure 4-5: fixed effects interactions with match

As we have already seen from the raw data and now from figure 4-5, Listgp slightly shows some variability with the T group achieving higher degrees of match than the two other groups, an observation witnessed in the reduced model base.Listgp in table 4-10. Similarly, vowel quality reflects slight variation as well since vowel /a/ yields less match than vowels /i/ and /u/. The three remaining variables; age, length and context show more variability than Listgp and vowel quality do.

These variables are included in the maximal model along with six interactions which were either derived from the research hypotheses or because they reflect variability as shown in appendix 4-3. These include Listgp:length, Listgp:context, Listgp:vowel.quality, context:length, context:vowel.quality and age:vowel.quality where (:) signifies an interaction. The four interactions Listgp:age, context:age, length:age, length:vowel.quality were dropped from the maximal model as a simplification procedure since three interactions had the control variable.[20]

We follow the same graphical exploration method with the random effects to construct the random effects structure. Figures 4-6 and 4-7 below display the correlation of listener and stimulus with match.

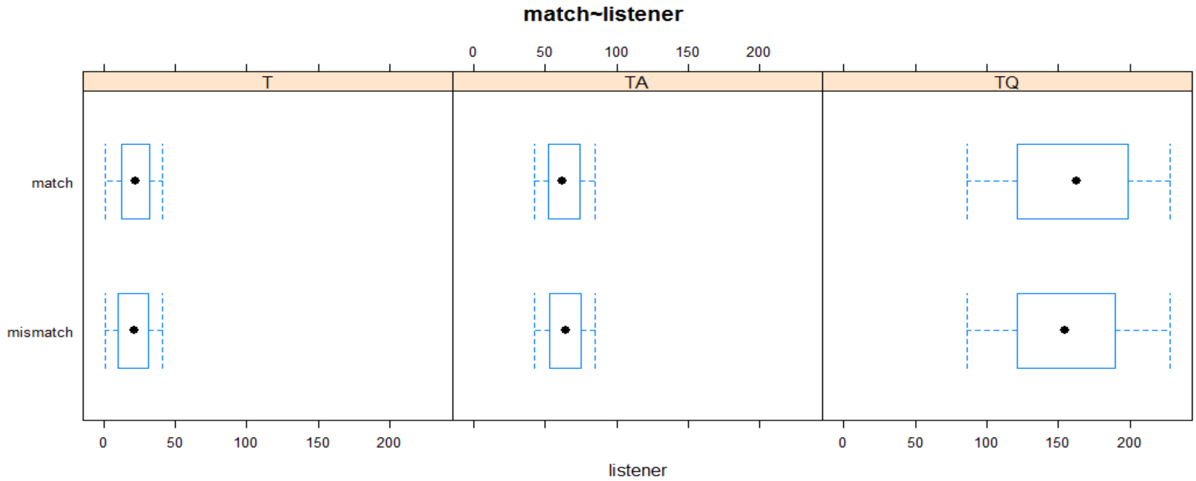


Figure 4-6: match correlation with listener

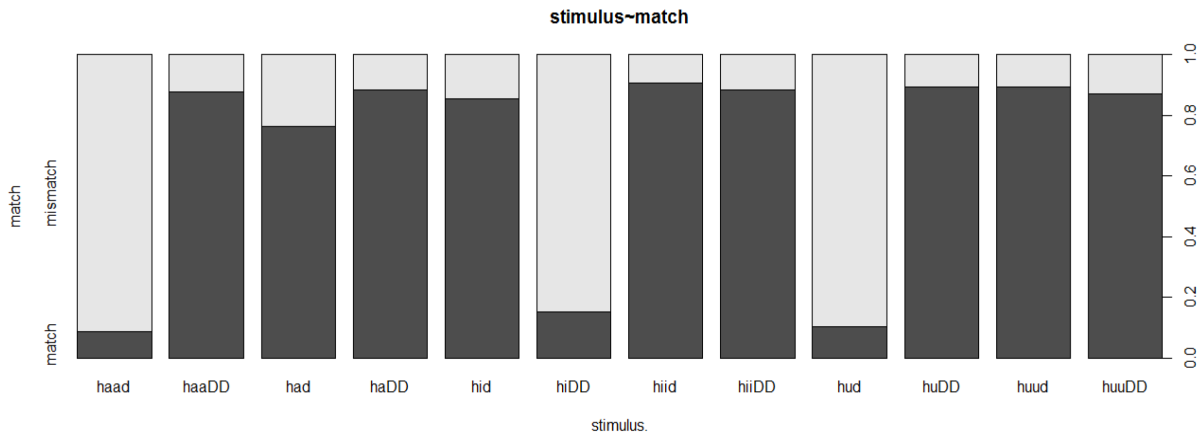


Figure 4-7: match correlation with stimulus

In addition, table 4-14 illustrates the random effects of listener and stimulus in the null model (i.e., the model with the intercept and without any variables) from which we extract the variance values of both effects.

[20] The variable length:vowel.quality was later added to the final model.

null.model<- glmer(match~ 1 + (1|listener)+ (1|stimulus), data = PAT, family = "binomial")

Groups	Name	Variance	Std.Dev.
Listener	(Intercept)	**1.680**	1.296
Stimulus	(Intercept)	**4.948**	2.224
Number of obs: 8208, groups: listener, 228; stimulus, 12			

Table 4-14: Random effects table of the null model

From table 4-14, we calculate the Variance partition coefficient VPC (Steele, 2008a) which is the total variance of both random effects as follows.

Total VCP=1.680+4.948= 6.628
∴ By-listener variance: 1.680/6.628= 0.2534, i.e.**, 25.3%**
∴ By-stimulus variance: 4.948/6.628= .7465, i.e., **74.65%**

Observations on random effects structure based on plots and null model:

1. Listener and stimulus variance values are both>0, which means that both have to be included in the final model's random effects' structure.
2. Stimulus variance is more than listener variance.
3. The random effects variance values indicate that almost 75% of the variance can be attributed to differences between stimulus and within listeners variables[21]. The remaining 25% of the variance can be explained by differences between listeners and within stimulus variables.

Step ii: Constructing the maximal model formula

The next step in the protocol of model selection is to build the maximal model but before doing so, we need to determine whether the variables of interest have 'within-unit random slopes' or 'between-unit random intercepts' where unit can be either a random subject or item (Barr et al, 2013). Baayen (2008, p. 290) states that "in general, predictors tied to subjects (age, sex, handedness, education level, etc) may require by-item random slopes, and predictors related to items (frequency, length, number of neighbors, etc) may require by-subject random slopes." Following Baayen's definition, we can categorize the variables involved in the PAT dataset as follows.

[21] The terms between-unit and within-unit in relation to the variables used are explained on the next page.

- ***Listgp*** is a ***between-listener*** since every listener belongs to a certain group and cannot be part of two or three groups at the same time and a within-stimulus variable as the stimulus presented to each of the three listener groups does not vary.

- ***Length*** is a ***within-listener*** variable since the stimulus length presented does not vary across listeners. It is also ***between-stimulus*** variable because a stimulus item can be either long or short but not both together.

- ***Context*** is a ***within-listener*** variable as stimulus context does not vary across listeners. However, it is ***between-stimulus*** because each context level is different than the other; either emphatic, plain, pharyngeal or q.

- ***Vowel.quality*** is a ***within-listener*** variable as the vowel quality of the stimulus tokens are the same for all the listeners. It is also ***between-stimulus*** since each vowel quality level is different than the other two vowels; either a, i or u.

- ***Age*** is a ***between-listener*** variable because every listener has a certain age and a ***within-stimulus*** variable since the stimulus presented does not vary across young and old listeners.

Following Barr et al's (2013) notion of maximal random effects structure, the current design assumes both random slopes and intercepts; thus, the maximal model formula is as follows.

```
modelPATset<-glmer(match~Listgp+context+length+age+ vowel.quality + Listgp:length + List
gp:context + Listgp:vowel.quality +context:length + context:vowel.quality+ age:vowel.quality
+ (Listgp|stimulus)+(vowel.quality+length+context|listener) , data = PATset, family = "binom
ial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)
```

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.284792	1.654989	0.172	0.86337	
ListgpTA	0.342786	0.605126	0.566	0.57107	
ListgpTQ	0.308373	0.435176	0.709	0.47856	
contextuvularized	-4.195321	1.721005	-2.438	0.01478	*
lengthshort	1.834031	1.511649	1.213	0.22503	
age	-0.013337	0.009758	-1.367	0.17173	
vowel.qualityi	-3.664941	1.856179	-1.974	0.04833	*
vowel.qualityu	-2.706976	1.889257	-1.433	0.15191	
ListgpTA:lengthshort	-0.600504	0.551291	-1.089	0.27604	
ListgpTQ:lengthshort	-0.650172	0.402534	-1.615	0.10627	
ListgpTA:contextuvularized	0.205941	0.516040	0.399	0.68983	
ListgpTQ:contextuvularized	0.139500	0.368180	0.379	0.70477	
ListgpTA:vowel.qualityi	0.265055	0.615715	0.430	0.66684	
ListgpTQ:vowel.qualityi	0.208071	0.433322	0.480	0.63110	
ListgpTA:vowel.qualityu	0.462933	0.640003	0.723	0.46948	
ListgpTQ:vowel.qualityu	0.630412	0.459616	1.372	0.17019	
contextuvularized:lengthshort	0.602552	1.245700	0.484	0.62859	
contextuvularized:vowel.qualityi	4.261465	1.476958	2.885	0.00391	**
contextuvularized:vowel.qualityu	2.597769	1.544575	1.682	0.09259	.
age:vowel.qualityi	0.023681	0.011000	2.153	0.03134	*
age:vowel.qualityu	0.007302	0.011736	0.622	0.53380	
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 4-15: summary table of the maximal model modelPATset

The maximal model in table 4-15 was simplified as follows.

1. Only the variables of theoretical interest, as derived from the hypotheses, were included. Moreover, a data-driven approach was used where the model's interactions were derived from the variables of interest and research hypotheses.
2. The continuous variable, age, was scaled and centered (normalized).
3. Number of optimizer's iteration was increased to 2e5; control=glmerControl=list(maxfun=2ef) and nAGQ=1), which is very high, so as to facilitate convergence of the regression models.
4. Collinearity of variables was adhered to since no two variables were found to be collinear with each other.

Next, seven regression models were fitted, six of which using the dropterm command in the MASS package (Venables & Ripley, 2003) in R whereas the final model was fitted by hand since the interaction length:vowel.quality was added as a last step. The derived models are provided in appendix 4-3. The model of the best fit is the final model in step_7 as shown below in table 4-16.

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	**1.492582**	**0.865270**	**1.725**	**0.084529**	**.**
contextuvularized	**-3.648663**	**0.896210**	**-4.071**	**4.68e-05**	*******
age	-0.013691	0.009707	-1.410	0.158399	
lengthshort	-0.776080	0.962151	-0.807	0.419892	
vowel.qualityi	**-5.136480**	**1.280216**	**-4.012**	**6.02e-05**	*******
vowel.qualityu	**-2.892269**	**1.232041**	**-2.348**	**0.018898**	*****
contextuvularized:vowel.qualityi	**4.731263**	**1.279049**	**3.699**	**0.000216**	*******
contextuvularized:vowel.qualityu	**2.269573**	**1.309578**	**1.733**	**0.083085**	**.**
lengthshort:vowel.qualityi	**3.041907**	**1.337849**	**2.274**	**0.022982**	*****
lengthshort:vowel.qualityu	**2.475961**	**1.396428**	**1.773**	**0.076217**	**.**
age:vowel.qualityi	**0.023302**	**0.010959**	**2.126**	**0.033477**	*****
age:vowel.qualityu	0.007495	0.011709	0.640	0.522098	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 4-16: summary table of step_7 (simple effects of variables and interactions)

Table 4-16 shows that two variables and three interactions were found either significant or near significant. These include ***context***, ***vowel quality***, ***context:vowel.quality***, ***length:vowel.quality*** and ***age:vowel.quality*** (bold-faced in the table). These variables and interactions explain the 75% of variance evident in the by stimulus differences since ***vowel quality i***, ***length*** and ***vowel.quality*** are all by-stimulus variables. On the other hand, the only by-listener variable having an effect on match is ***age***, albeit only when interacting with ***vowel quality***, thus, the 25% variance reported in table 4-14. One striking observation is that ***Listgp*** is not reported in the final model as having any statistical significance. Each of the significant or near significant variables and interactions is plotted using the effects package (Fox, 2003) in R and is interpreted below.

Figure 4-8 below displays a plot of all the significant effects in step_7 where the x-axis represents one of the two independent variables of an interaction whereas the levels of the other variable are shown in two panels. Moreover, the y-axis represents the probability of matched responses on a scale of 0 to 1 with 1 denoting a match (where match means that the observed responses match those observed in the corpus mappings) and 0 denoting a mismatch. Thus, degrees of match are higher if higher than 0.5 on the y-axis and lower if lower than 0.5 given the contrast coding scheme of match and mismatch.

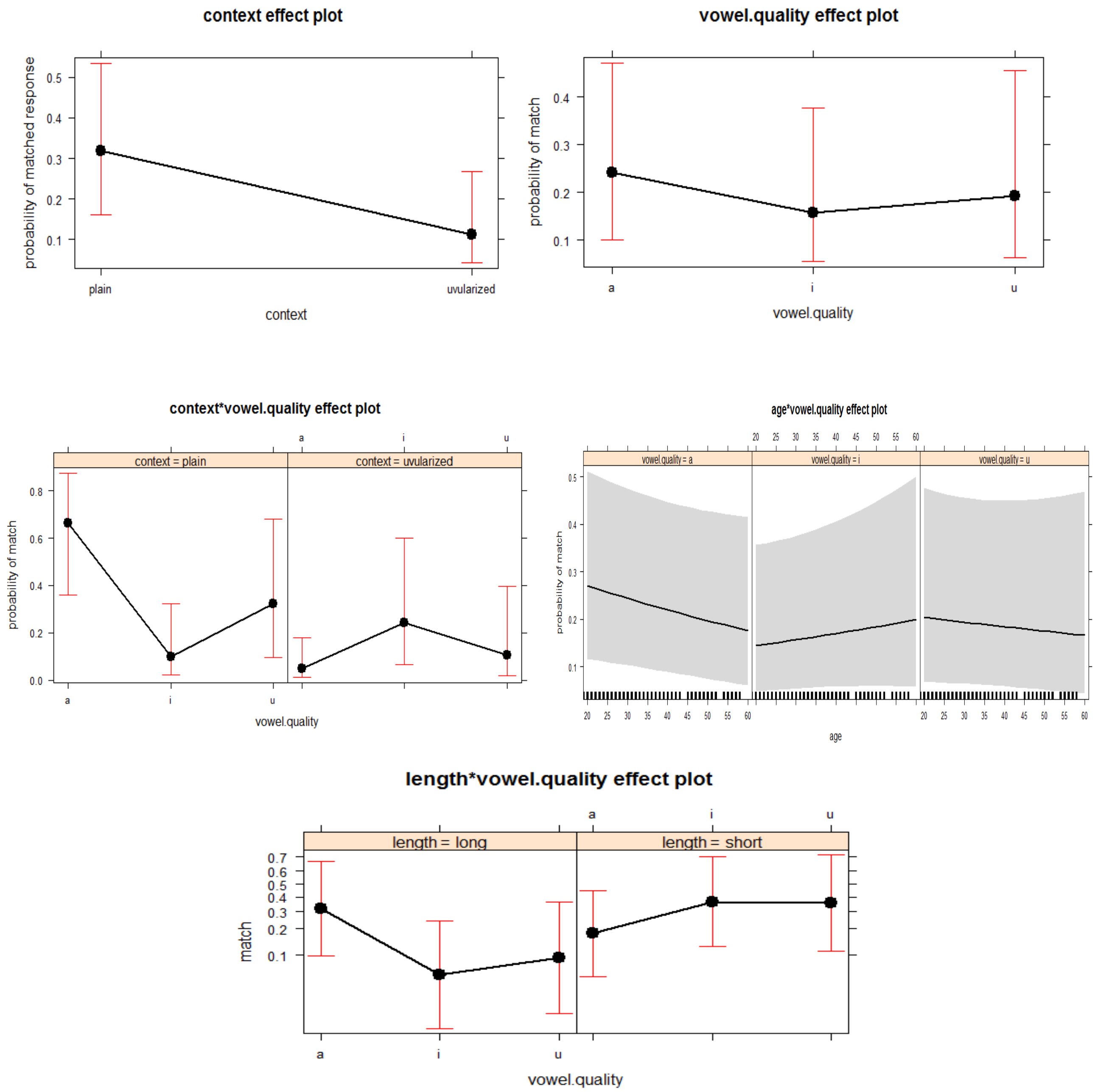


Figure 4-8: effect displays for all the interactions in model step_7, mismatch=0, match=1

The results of table 4-16 and the effect displays in figure 4-8 suggest that probably the only effect we might want to put a strong interpretation to in the model step_7 is the significant interaction between context (uvularized) and vowel quality (i) illustrated in figure 4-9 below.

This interaction is a significant simple effect at the uvularized i level where uvularized i (i.e. hiDD) is less likely to match at a p-value of 0. Regarding the main effects of the two factors in this significant interaction, we can set these aside since there does not seem to be any obvious overall main effect of context or vowel quality in figure 4-9.

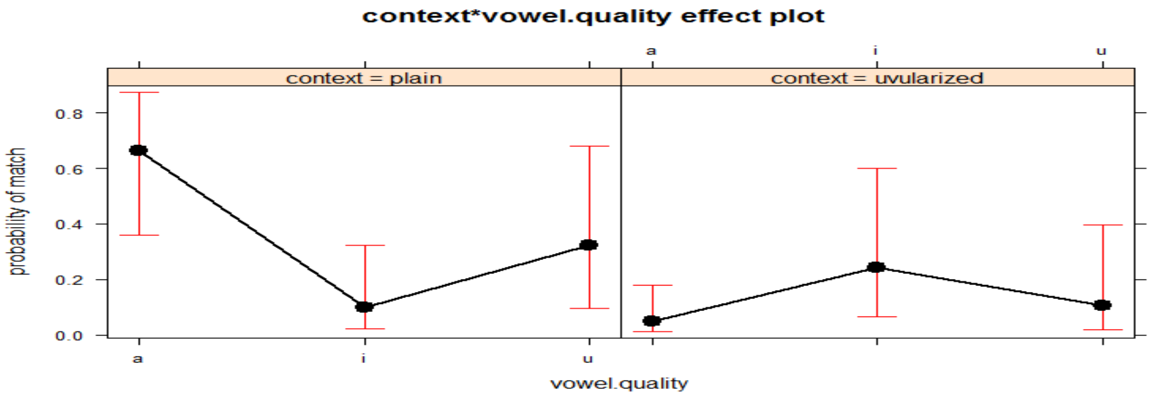


*Figure 4-9: context*vowel.quality effect plot*

Similarly, in figure 4-10 the significant interaction between length and vowel quality at the level of short i suggests that short i (i.e. hid/hiDD) is less likely to match than long i (i.e. hiid/hiiDD), this reflects that in the raw data (figure 4-7) hiid/hiiDD both have high match scores, but hid/hiDD are different from each other, and the low match for hiDD pulls down the average score for short i. Likewise, the marginal interaction for short u reflects the fact that in the raw data (figure 4-7) huud/huuDD both have high match scores, but hud/huDD are different from each other, and the low match for hud pulls down the average score for short u.

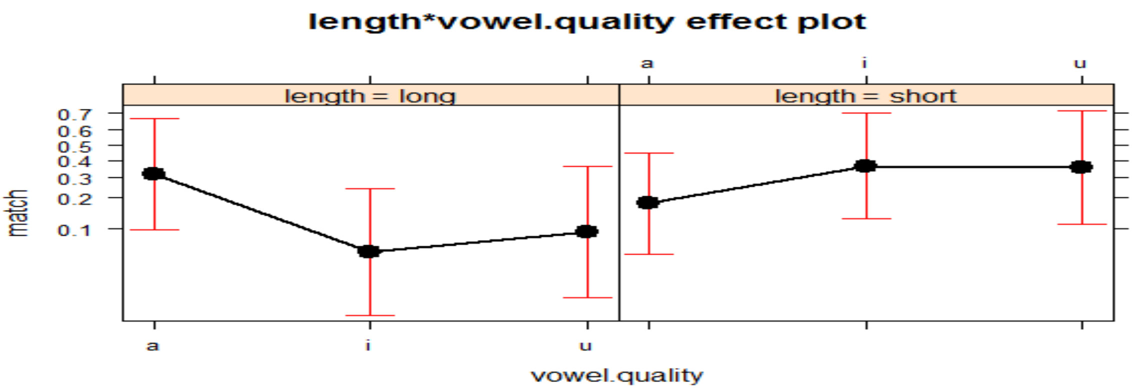


*Figure 4-10: length*vowel.quality*

Overall, these three effects (contextuvularized:vowel.qualityi, lengthshort:vowel.qualityi and lengthshort:vowel.qualityu) tell us what we can see in figure 4-7: that there are just three items that stand out as having much lower match (haad, hiDD and hud).

One interaction is also reported as being significant in table 4-16 which is the interaction between age and vowel quality i given in figure 4-14. This interaction can be interpreted such that there is an age difference between the people who matched/mismatched their responses to the /i/ words (i.e. hid, hiDD, hiid and hiiDD).

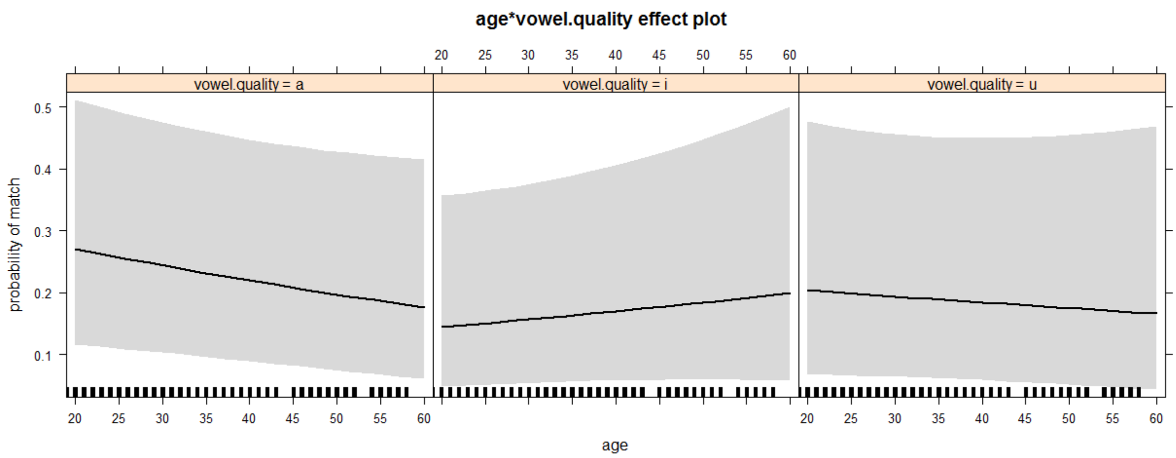


*Figure 4-11: age*vowel.quality effect plot*

Step iii: Model validation

The final step in the protocol is validating our selected model which is done here via inspecting plots for normality and homoscedasticity; tests for goodness of fit. Figure 4-12 shows that the residuals of model step_7 are homoscedastic. That is, the residuals points are centering around the 0 line without forming patterns on the positive or negative lines. Moreover, the solid line almost overlays the dashed line indicating a good fit. The second theoretical assumption of normality is also adhered to in figure 4-13 where the residuals histogram depicts a bell-curve shape.

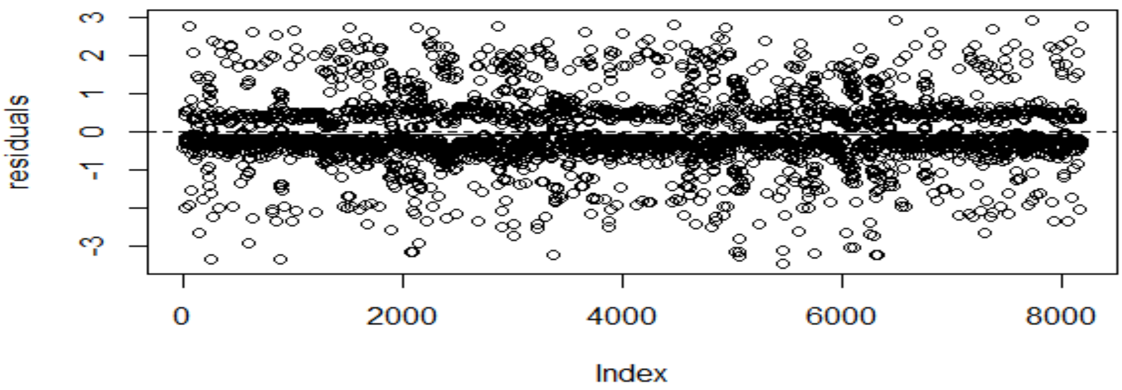


Figure 4-12: scatter plot of model step_7's residuals (testing for homoscedasticity)

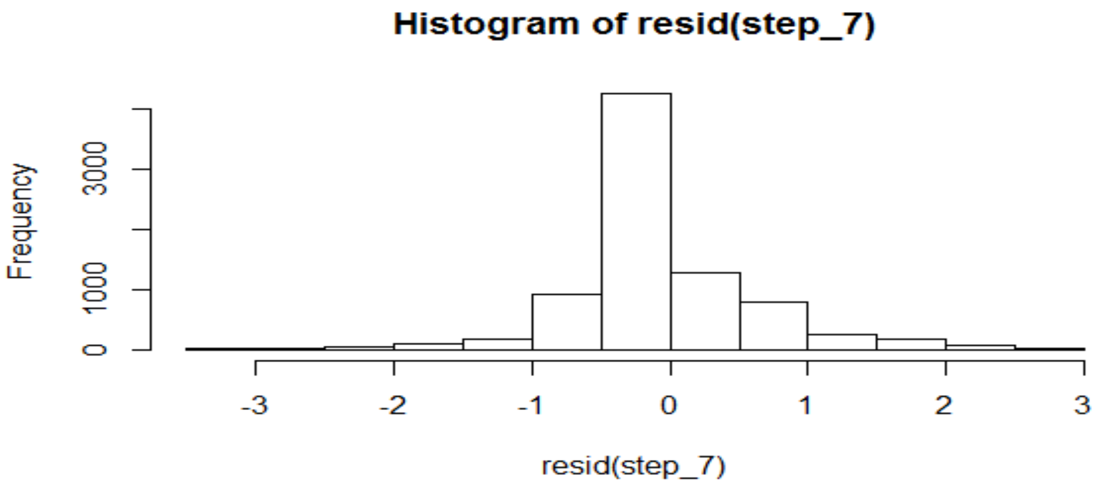


Figure 4-13: histogram of model step_7's residuals (testing for normality)

4.4 Discussion

To sum up, this chapter reports on the findings of a maximal generalized mixed effects modelling (GLMM) analysis performed in R (Team, 2015) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) to confirm the effects of predictors and/or interactions on the dependent variable ***match***, computed by comparing response vowels with corpus vowels. The fixed effects structure included the predictors ***Listgp***, ***context***, ***length***, ***vowel.quality*** and ***age*** as a control variable, which was scaled and centered, and the six interactions ***Listgp:context***, ***Listgp:length***, ***Listgp:vowel.quality***, ***context:length***, ***context:vowel.quality*** and **age:vowel.quality**. The interaction ***length:vowel.quality*** was added at a later stage to the final model. The maximal random effects structure assumed both intercepts and slopes, the slopes of the same fixed effects but not their interactions (Listgp|stimulus) and (vowel.quality+length+context|listener).

The three theoretical assumptions of collinearity, normality and homoscedasticity were adhered to when building the maximal model and when validating the final model. The significance level threshold of 5% (p>0.05) was used when selecting the final model and the results were plotted at a confidence level of 95% (Barr et al 2013).

The current study aimed to answer the following RQs reproduced from section 4.

RQ1: How close is the perception of the listeners to the observed mappings in the qualitative corpus?

- How do Turkish participants perceptually assimilate the Arabic long vs. short vowels to the Turkish short vowels?
- How do Turkish participants perceptually assimilate the Arabic plain vs. emphatic vowels?

RQ2: Does knowledge of Arabic (phonology) have any effect on perception?

In addressing RQ1, it was found that 68% of the categorizations in the PAT experiment matched those found in the corpus which means that there was almost 70% agreement between the listeners' perception of the PAT vowels and that of the Ottomans in the corpus data (chapter 3). This suggests that perception can account for most of the mappings, a result sustained by the tall bars in figure 4-7. The remaining 30% of mismatch may be attributed to different factors including the phonology and orthography.

In the logistic regression, we saw that length by itself did not have a significant effect on match but rather depended on the quality of the stimulus vowel in their effect on match. In the mapping of the short vowels, match occurred in the categorization of the vowels /a/$_A$>/e/$_T$ (76%), [aˤ]$_A$>/a/$_T$ (88%), /i/$_A$>/i/$_T$ (85%), [uˤ]$_A$>/u/$_T$ (89%) whereas mismatch occurred in the perception of [iˤ] $_A$>/e/ $_T$ (85%) and /u/$_A$>/u/$_T$ (90%) (predicted /y/$_T$). As for the long vowels, more match transpired, namely in [a:ˤ]$_A$ >/a/$_T$ (87%), /i:/$_A$>/i/$_T$ (90%), [i:ˤ]$_A$> /i/$_T$ (88%), /u:/$_A$>/u/$_T$ (89%) and [u:ˤ]$_A$ > /u/$_T$ (86%). However, only one long vowel mapping displayed mismatch in [a:]$_A$ >/e/$_T$ (91%).

The stimulus context (plain-emphatic) was found a very significant factor (in the R analysis), p<0. This is mirrored in the results, that match happened in the mappings of plain (/a/$_A$>/e/$_T$, /i/$_A$>/i/$_T$, /i:/$_A$>/i/$_T$, /u:/$_A$>/u/$_T$) and emphatic ([aˤ]$_A$>/a/$_T$, [uˤ]$_A$>/u/$_T$, [a:ˤ]$_A$ >/a/$_T$, [i:ˤ]$_A$> /i/$_T$, [u:ˤ]|$_A$ >/u/$_T$) vowels and similarly mismatch occurred in the categorization of plain (/a:/$_A$>/e/$_T$, /u/$_A$>/y/ $_T$) and emphatic vowels ([iˤ] $_A$>/e/$_T$).

The interaction of context with vowel.quality on match was also found very significant when the vowel in the stimulus was a uvularized /i/ such as in [i:ˤ]$_A$ > /i/$_T$, and significant when the vowel was uvularized /u/ such as in [uˤ]$_A$>/u/$_T$ and [u:ˤ]$_A$ > /u/$_T$. However, the categorization of the vowel /a/ was found to reflect lower degrees of match (mainly because of the mismatch of /a:/$_A$>/e/$_T$).

In order to explain these findings, we check the spectral elements (F1 and F2) of both Arabic and Turkish vowel space as participants heard them in the PAT experiment. In the PAT, the Turkish participants listened to different phonetic Arabic vowel categories including the three Arabic short vowel phonemes and their contrasting long ones in plain and emphatic settings. Then, they had to select the nearest phonetic vowel categories from the Turkish 8 vowel set. Figure 4-14 shows the acoustics of both Arabic and Turkish vowels on the vowel space whereas figure 4-15 depicts the vowel mappings in both the corpus and the PAT experiment.

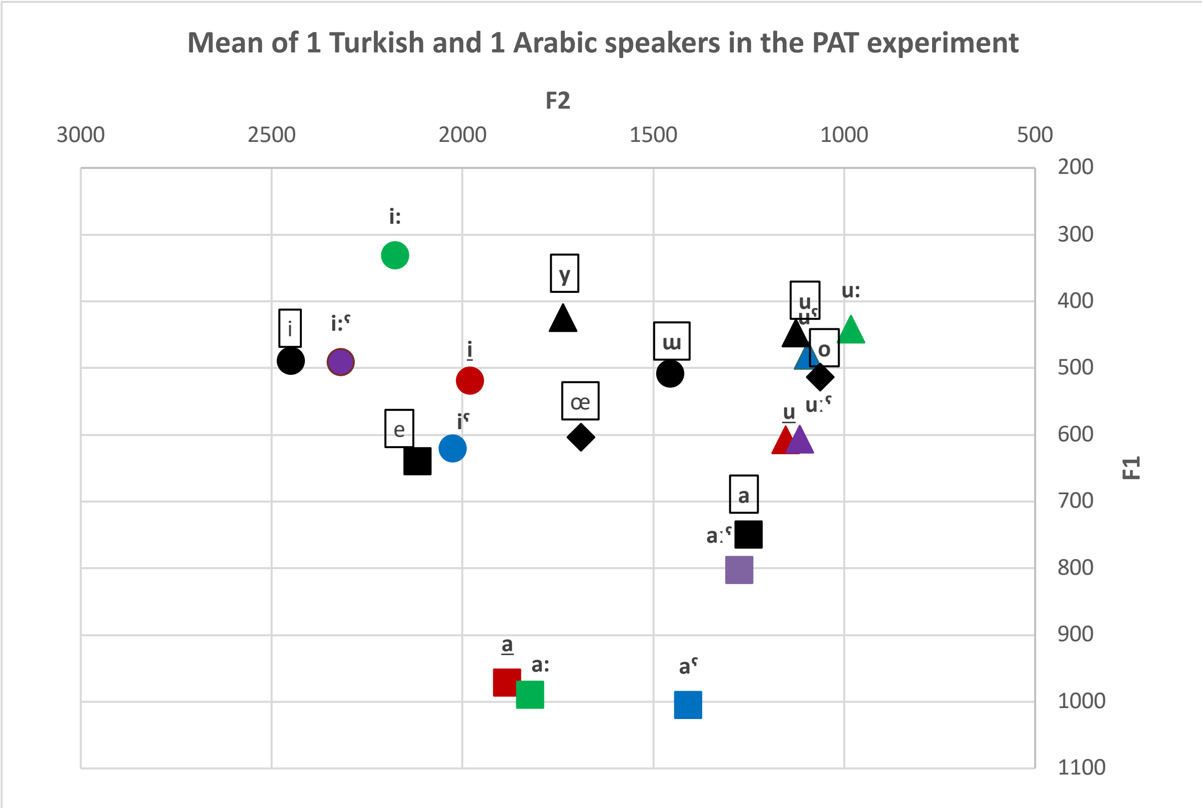


Figure 4-14: Mean frequency values of 1 Turkish speaker and 1 Arabic speaker's vowel formants plot (nonsense words)

Vowels in boxes are Turkish vowels and the rest are Arabic ones; underlined /a/, /i/and /u/ are Arabic vowels

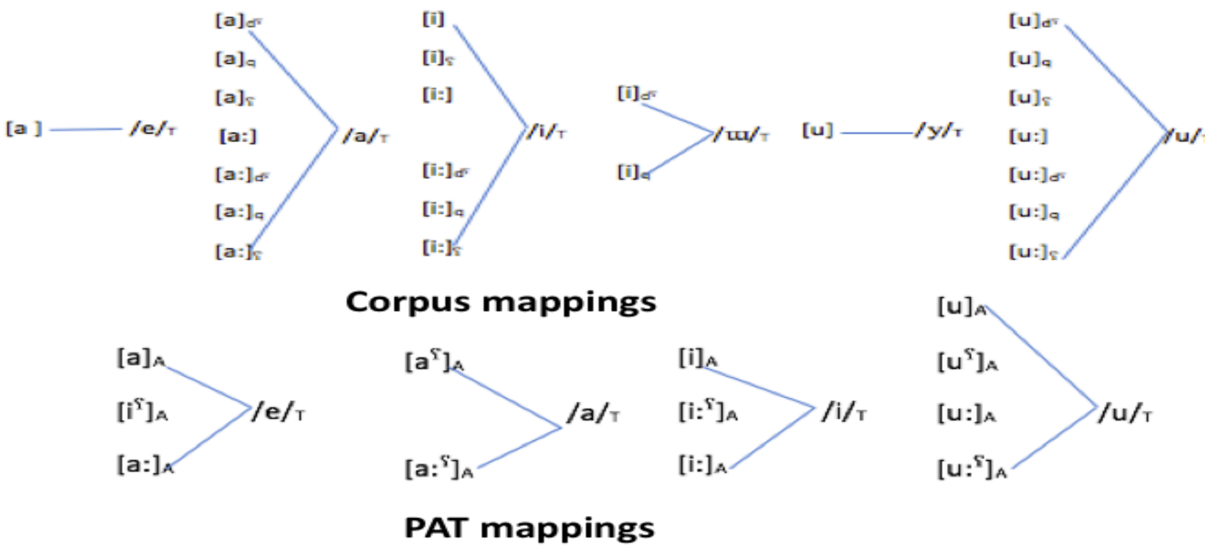


Figure 4-15: Corpus and PAT vowels

As can be seen from figures 4-14 and 4-15, the vowel distance i.e. backness between Arabic and Turkish vowels accounts for most of the mappings of the PAT experiment. Distance in terms of backness (F2) seems to be more important than height in the mapping probably because the listeners depend on their Turkish front-back distinction of vowel harmony (see chapter two). Arabic /a/ and /a:/, both phonetically front (as shown in figure 4-14 and in chapter 2), are closer in F2 [distance] to the Turkish $/e/_T$ than to Turkish $/a/_T$, meaning that $[a:]_A>/e/_T$ is phonetically grounded. In addition, Arabic $/a:/_T$ and $/e/_T$ agree in the two distinctive features of height and frontness (see 2.1.4 on VH). This means that the mismatched mapping $[a:]_A>/e/_T$ is both phonologically and phonetically grounded. On the other hand, both $[a^ʕ]_A$ and $[a:^ʕ]_A$ are both phonetically back in Arabic, hence closer to $/a/_T$. As for /i/, the two phonetic categories $/i:/_A$ and $[i:^ʕ]_A$ are near the phonetic space of $/i/_T$ and hence are perceived as such. The two remaining Arabic vowels $/i/_A$ and $[i^ʕ]_A$ are near two Turkish vowels, namely $/i/_T$ and $/e/_T$. The more common option for the listener would be to retain the vowel quality of the stimulus vowel and they map $[i]_A$ into $/i/_T$ which agree in height and frontness. On the other hand, in the case of $[i^ʕ]_A$, the listeners select $/e/_T$ which is lower (F1) than $/i/_T$ but closer in both F1 and F2 to $/e/_T$. This means that the mapping of $[i^ʕ]_A$ onto $/e/_T$ is phonetically motivated whereas $[i]_A$ onto $/i/_T$ is phonologically driven.

For the vowel /u/, most listeners in the three groups selected the vowel with the same vowel quality in Turkish, which is $/u/_T$ which is close to the four Arabic phonetic categories in F1 and F2 , i.e., $/u/_A$, $/u:/_A$, $[u^ʕ]_A$, $[u:^ʕ]_A$ on the vowel chart in figure 4-14. The vowel $/o/_T$ is also in close proximity to the four /u/ categories and some listeners selected it as a response vowel. Nevertheless, the numbers of these responses are minimal compared to $/u/_T$ (cf. table 4-5). Thus, we can say that the mapping the Arabic /u/ variants onto Turkish /u/ is phonetically supported. Moreover, most listeners are influenced by the vowel quality (distinctive features) of the two categories in both the source and native language. That is, they can hear that the Arabic vowel and Turkish vowel share the distinctive features of height and backness. This indicates that the mappings onto $/u/_T$ are also phonologically motivated.

4.5 Summary

To sum up, the mapping of most Arabic categories in Turkish is perceptually grounded as figure 4-14 shows. However, some mappings are only phonetically grounded (e.g., $[i^{ʕ}]_A>/e/_T$), others are only phonologically grounded (e.g., $[i]_A>/i/_T$) while still others are both phonetically and phonologically grounded (e.g., $[u]_A>/u/_T$).

Regarding RQ2, it was concluded that listener group does not have a main (significant) role in the mappings since the three groups manifested similar patterns with the T group reflecting more match than the two other groups; at 69.7% and p-value of 0.07, 67.8% and p-value of 0.27 and 67.6% and p-value of 0.17 for the T, TA and TQ groups respectively. The TA and TQ groups' similar mapping patterns indicate that knowledge of Arabic does not determine the quality of the vowels surrounding gutturals. Thus, perception did not vary by listener group, a result validated by figure 4-6. However, the role of the source language is seen in the fact that some mappings were phonologically only determined $[i]_A>/i/_T$.

One observation *vis-à-vis* listener group was that more variation was evident among the TQ group listeners only and detected in the mappings of each of the three Arabic vowels /a/, /i/ and /u/. This variation might be due to the large number of participants in the TQ group of 143 compared to the other two; T: 41 and TA: 44.

In the next chapter, a more realistic simulation of the borrowing process is presented and discussed in the form of a Simulated Borrowing experiment, which allows for the influence of orthography and segmental context. This is done to see whether the participants' mappings will closely match those seen in the corpus.

5 Simulated Borrowing- Audio only data

5.1 Introduction

The findings of the Perceptual Assimilation Task, in chapter four, where nonsense words were used showed that listener group did not a have significant effect on match. The researcher then conducted another experiment but with a mix of real and nonsense words to ascertain whether listener group has any effect on match when interacting with other variables such as stimulus length, context, vowel quality and frequency.

The Simulated Borrowing experiment was run online by Qualtrics portal just like the Perceptual Assimilation Task, PAT in two phases. As mentioned in chapter four, phase I was conducted by the researcher herself over a period of three weeks while phase II was run by the Qualtrics team over a period of seven weeks. During phase I, as in the PAT experiment, 54 participants took the online survey; however, only 26 met the survey's requirements and 28 responses were eliminated. When a participant skipped questions, did not answer all mandatory questions, provided nonsense responses (such as 'bbbb', etc) or did not consent to the terms of the survey, their data were not used. Then in phase II, different than the PAT whose entry was 520, 281 participants responded to the survey. Of these, only 51 responses were used whereas the remaining 230 were discarded based on the screening criteria described in phase I above.

The SB experiment was run in three conditions: audio only, A; written only, W; and audio+ written together, AW as elaborated in section 5.2.3. Similar to the PAT, the three listener groups in the A condition listened to recordings of all Arabic vowel categories; plain and emphatic, short and long, however, this time within randomized real and nonsense monosyllabic and polysyllabic words in addition to some distractors to test the engagement of the respondents. Some examples of the distractors are real words such as ***ibil***, camels; ***mihan***, professions; and ***niqam***, curses; and non-words like ***ti?im***, ***fi?ab*** and ***iTam*** among others.

The SB experiment has two main research questions as provided in **1** below.

(1) **RQ1:** Whether speakers of Turkish would generalize the residual effects (found in the ALT corpus) of emphatics[22]/gutturals on neighbouring vowels to actual non-borrowed words and to nonsense words.
RQ2: Whether knowledge of Arabic orthography and phonology plays a role in determining the quality of vowels neighbouring gutturals.

RQ1 in **1**. compares the results in the SB with the corpus mappings since the dependent variable is 'match', where match is defined in terms of match to the corpus mappings. This allows us to compare the degree of match between the SB and the corpus 'match', on the one hand, and the degree of match between the PAT and corpus 'match' on the other. RQ2 tests whether written stimuli would yield different responses from the TA and TQ participants in order to establish whether knowledge of Arabic writing and orthography have any effect on vowel mapping from Arabic to Turkish. For readability purposes, RQ1 is investigated in this chapter whereas RQs 2 is explored in chapter 6.

The subsequent sections are organized as follows. Section 5.1. sketches the research hypotheses. Section 5.2. presents the research methodology starting with the recording, stimuli used, review of the participants, rationale and procedure followed. Section 5.3. reports on the analysis of the experiment's results and discussion of RQ1. Section 5.4 concludes with a summary of the chapter.

5.1.1 Hypotheses

The main hypothesis of the SB experiment is the same as that of the PAT; that the patterns of assimilation in the experiment will match those observed in the research corpus. In other words, the ALT words can be mainly explained using the principles of the 'loanword adaptation as perception' argument (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992). However, the effect of listener group, if an effect is found, be them monolinguals, bilinguals or Quranic's Turkish speakers, would determine who carries out the borrowing similar to the Ottomans, thus by extension revealing which

[22] Here the term emphatics is used to represent gutturals in general unless otherwise stated.

loanword adaptation model, Perception only, Phonology only[23] or the hybrid model[24], could be used to explain the ALT corpus facts.

The patterns of categorizations observed in the study's corpus are reproduced from chapters 3 and 4 in table 5-1 and are reviewed below.

Context	*Vowel length*	a	i	u
Plain	*short*	i. a ---> e [dars~derʂ] 'lesson'	ii. i --->i [dʒin~dʒin] 'genie'	iii. u ---> y [ʃukr~ʃycyɾ] 'thankfulness'
	long	iv. a: ---> a [ʕan~an] 'instant'	v. i: --->i [fi:l~fil] 'elephant'	vi. u: ---> u [nuːr~nuɾ̥] 'light'
Emphatic	*short*	vii. ɑ/ aˤ ---> a [ðˤarf~zarf] 'envelope'	viii. iˤ ---> ɯ [tˤiˤp~tɯp] 'medicine'	ix. uˤ ---> u [tuħaf~tu̲haf] 'strange'
	long	x. a:ˤ ---> a [ʕaːr~aɾ̥] 'shame'	xi. i:ˤ ---> i [basˤi:ˤrah~basiret] 'foresight'	xii. u:ˤ ---> u [maðˤluːm~mazlum] 'wronged'

Table 5-1: Corpus patterns of assimilation in the ALT

According to table 5-1, $/a/_A$ occurring in the vicinity of a pharyngeal, uvular or emphatic consonant in an Arabic cognate appears in its Turkish counterpart as back $/a/_T$ (**iv, vii and x**), otherwise as front $/e/_T$ (**i**). Likewise, $/u/_A$ in the neighbourhood of a pharyngeal, uvular or emphatic in Arabic appears in Turkish as $/u/_T$ (**vi, ix and xii**), otherwise as $/y/_T$ (**iii**). The vowel $/i/_A$ occurring in the proximity of a uvular **q** or emphatic consonant (but not a pharyngeal) in Arabic appears in Turkish as $/\text{ɯ}/_T$ (**viii**), otherwise as $/i/_T$ (**ii, v, xi**).

In addition to the main hypothesis, one sub-hypothesis can be derived from the SB's RQ, which is given in **2.** below.

(2) Sub-hypothesis of the SB experiment.

H1: Speakers of Turkish generalize the residual effects of gutturals on neighbouring vowels to actual non-borrowed words and to nonsense words.

According to H1, the results of the SB experiment in the A condition are predicted to resemble those of the PAT. In both experiments, the percentage of 'match' is predicted to be more than that of 'mismatch' across the three listener groups. In the next sub-section, the

[23] (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001a, 2008)
[24] (Kenstowicz & Suchato, 2006; Smith, 2006; Change, 2008 and Dolus, 2013)

methodology is reviewed including the recording session, stimuli, participants, rationale and procedure.

5.2 Methodology

5.2.1 Recording and Stimuli

One recording session took place in the data lab in the Department of Language and Linguistic Science at the University of York. It was carried out using a Neumann U87i microphone, a TAC Scorpion mixing desk, M-Audio 24/96 Audio card and Adobe Audition CS5.5 on Windows 7 Pro x64 on PC software with 44.1 khz 16 bit sampling rate.

In this session, a set of real and nonsense educated Arabic, Classical/MSA, words totaling to 72 monosyllabic words were recorded by a native Arabic speaker of Syrian origin from Aleppo who was asked to read the words in MSA mode. The choice of MSA is based on the following argument. Firstly, most of the words in the study's corpus are similar to a higher variety than Arabic vernaculars. Secondly, these words are not the same as those collected by Tietze (1958 and 1999) of Arabic vernacular words which include words from Syrian, Egyptian, Lebanese, Iraqi and other Arabic varieties. Thirdly, the large number of words used in the corpus dates back to the second stage of loanword adaptation in which Classical Arabic was considered to correct the pronunciation of the words from the first stage which were borrowed via Persian (A Tietze, 1992). The choice of the speaker to be of Syrian origin stems from the following two observations. Firstly, the a>e pattern found in the corpus mappings and in the feminine construction in particular is similar to that found in Levantine Arabic. Second, the ratio of Arabic words of Syrian origin compared to those from other varieties in Tietze's (1958 and 1999) lists is higher being 72 words out of 216 compared to 2 Lebanese words, 1 Iraqi and 1 Egyptian. The remaining 140 words seem to have been borrowed from Classical Arabic/MSA or from other Arabic dialects which Tietze did not mention.

The classification of the stimuli words is illustrated in **3** and the actual stimulus material is given in table 5-2.

(3) Stimulus material consists of 2 blocks

1. ***Monosyllabic real and non-words:*** **72** words with plain and emphatic (guttural) variants (short and long vowels); 6 vowels (long+short) X 3 tokens (1 per condition) X 4 consonant types
2. ***Polysyllabic distractors (real& non-words):*** **12** words with plain and emphatic/guttural variants (short vowels); 1 vowel combination (short) X 3 tokens (of the shapes i-i, a-i and i-a, 1 per condition) X 4 consonant types

For the monosyllabic real and nonsense words given in table 5-2 below, 6 phonetic categories were tested which are illustrated in the first column of the table. These are the Arabic short /a/, /i/ and /u/ and long /a:/, /i:/ and /u:/ vowels. Then, four types of consonants are chosen, namely plain consonants in the 2nd column, emphatics and uvular q in the 5th column, and pharyngeals in the 8th column. Additionally, the Arabic word and English glossing are provided for each of the stimulus tokens. The stimulus words are transcribed in IPA font and next to each stimulus word, the predicted output in Turkish is given also in IPA transcription.

The polysyllabic words were dropped from the stimulus material and were not part of the ultimate analysis in the Simulated borrowing experiment. However, 12 polysyllabic words were only used as distractors as shown in (3) above.

In addition to table 5-2, the frequency of the stimulus material was checked and taken as a variable in the experiment. The Arabic words frequencies were obtained from arabiCorbus (Parkinson, 2009) , which is a database comprised of multiple corpora ranging from the Quran; Hadith, sayings of Prophet Mohammed; Medieval Science; some newspapers; modern literature; nonfiction novels; Islamic discourse; Egyptian colloquial; Penn Treebank and One Thousand and one night among others. The word frequencies are given in table 5-3.

Category	*Plain consonant* CV(V)C, CV(V)C_2C_2	*Arabic Word*	*English glossary*	*Emphatics & q* CV(V)C, CV(V)C_2C_2	*Arabic Word*	*English glossary*	*Pharyngeal & uvulars* CV(V)C, CV(V)C_2C_2	*Arabic word*	*English Glossary*
/a/	man>men barr>ber̥ faθθ>fez/fev	مَن بَر فَذ	Who Land Unrivaled	dˤabb>dap qad>kat sˤabb>sap qabb>kap tˤall>tal qatt>kat	ضَب قَد صَب قَب طَل قَت	giant lizard (Uromastyx) Indeed+ waist Passionate Chief Dew Dry grass; fodder+ lying	ʕan>an ʕajj>aj ʕall>al	عَن عَي عَل	About Inarticulate Maybe
/i/	min>min nidd>nit zirr>zir̥	مِن نِد زِر	From/of Rival Button	dˤid>dɯt qid>kɯt sˤill>sɯl qinn>kɯn sˤirr>sɯr̥ qibb>kɯp	ضِد قِد صِل قِن صِر قِب	Against Cummerbund/belt Egyptian copra Chicken coop Severe cold A bone protruding from the back	ʕijj>ij ʕihh>ih ʕizz>iz/is	عِي عِه عِز	Aphasia **Non-word** Glory
/u/	muð>myz dubb > dyp sull>syl	مُذ دُب سُل	Since Bear Tuberculosis	**sˤ**ujj>suj qud>kut sˤumm>sum qunn>kun dˤurr>dur̥ qull>kul	صُي قُد صُم قُن ضُر قُل	**Non-word** Codfish Deaf A small mountain Harm Small, short	ħubb>hup ʕuθθ>uv ʕubb> up	حُب عُث عُب	Love Mite/moth (pl.) Sleeves (of a shirt)
/aː/	ʃaːðð>ʃad/ʃav/ʃaz haːm>haːm naːb>nap	شَاذ هَام نَاب	Odd Important Canine	dˤaːrr>dar̥ qaːr>kar̥ sˤaːll>sal qaːb>kap sˤaːw>sav qaːt> kat	ضَار قَار صَال قَاب صَاو قَات	Harmful Tar Zebra with loud sound Gap Empty Qat/khat/kat (plant)	ħaːrr>har̥ ʕaːl>al ħaːff>haf	حَار عَال حَاف	Hot High dry

/iː/	riːf>rif diːk>dic tiːn>tin	رِيف دِيك تِين	Countryside Rooster Fig (pl.)	tˤiːb>tip qiːr>kiɾ̥ tˤiːn>tin qiːd>kit tˤiːl>til qiːb>kip	طِيب قِير طِين قِيد طِيل قِيب	Scent Tar Clay Little amount Rope for tying camels Amount	ʕiːr>iɾ̥ ʕiːs>is ʕiːh>it	عِير عِيس عِيه	Camels White camels (certain camel breed) Calling for camels to stop doing something
/uː/	θuːm>sum kuːz>kus suːd>sut	ثُوم كُوز سُود	Garlic Cone Black (adj., pl.)	**t**ˤuːn>tun **q**uːb>kup **t**ˤuːb>tup **q**uːr>kuɾ̥ **t**ˤuːm>tum **q**uːf>kuf	طُون قُوب طُوب قُور طُوم قُوف	Water abundance Impetigo (disease) Brick Small mountains; certain mounds Death; grave Top part of ear/neck	ħuːt>hut ħuːb>hup ħuːh>huh	حُوت حُوب حُوه	Whale Sin **Non-word**

Table 5-2: Monosyllabic real and non-words

S.no	Word	Word in Arabic	English glossary	Frequency	Word nature
1	man /man/	مَن	Who	7.8 instances of من per 100,000 words in All.	Real
2	barr /barr/	بَر	Land	20.2 instances of بر per 100,000 words in All.	Real
3	faThTH/faθθ/	فَذ	Unrivaled	14.93 instances of فذّ per 100,000 words in All.	Real
4	Dhabb /dˤabb/	ضَب	giant lizard (Mastigure)	50.02 instances of ضب per 100,000 words in All.	Real
5	qad /qad/	قَد	Indeed+ waist	75.67 instances of قد per 100,000 words in All.	Real
6	Sabb /sˤabb/	صَب	Passionate	28.46 instances of صب per 100,000 words in All.	Real
7	qabb /qabb/	قَب	Chief	72.25 instances of قبّ per 100,000 words in All.	Real
8	Tall /tˤa̱ll/	طَل	Dew	35.89 instances of طل per 100,000 words in All.	Real
9	qatt /qatt/	قَت	Dry grass; fodder+ lying	25.53 instances of قت per 100,000 words in All.	Real
10	‘an /ʕan/	عَن	About	15.78 instances of عن per 100,000 words in All.	Real
11	‘ayy /ʕajj/	عَي	Inarticulate	25.59 instances of عيّ per 100,000 words in All.	Real
12	‘al /ʕall/	عَل	Maybe	20.98 instances of علّ per 100,000 words in All	Real
13	min /min/	مِن	From/of	59.07 instances of من per 100,000 words in All.	Real
14	nidd /nidd/	نِد	Rival	3.93 instances of ند per 100,000 words in All.	Real
15	zirr /zirr/	زِر	Button	1.51 instances of زر per 100,000 words in All.	Real
16	Dhid /dˤid/	ضِد	Against	54.5 instances of ضد per 100,000 words in All.	Real
17	qid /qid/	قِد	Cummerbund/belt	0.44 instances of قِد per 100,000 words in All.	Real
18	Sill /sˤill/	صِل	Egyptian copra	23.67 instances of صلّ per 100,000 words in All.	Real
19	qinn /qinn/	قِن	Chicken coop	2.74 instances of قن per 100,000 words in All.	Real
20	Sirr /sˤi̱rr/	صِر	Severe cold	3.02 instances of صر per 100,000 words in All.	Real
21	qibb /qibb/	قِب	A bone protruding from the back	8.5 instances of قب per 100,000 words in All.	Real
22	‘iyy /ʕijj/	عِي	Aphasia	6.56 instances of عي per 100,000 words in All.	Real
23	‘ihh /ʕihh/	عِه	---------------------------	**Not applicable**	**Non-word**
24	‘izz /ʕizz/	عِز	Glory	17.04 instances of عز per 100,000 words in All.	Real
25	muTH /muð/	مُذ	Since	0.64 instances of مذ per 100,000 words in All.	Real
26	dubb /dubb/	دُب	Bear	8.52 instances of دب per 100,000 words in All.	Real

27	sull /sull/	سُل	Tuberculosis	2.8 instances of سل per 100,000 words in All.	Real
28	Suyy **/s**ˤujj/	صُي	--	**Not applicable**	**Non-word**
29	qudd /qud/	قُد	Codfish	0.97 instances of قُد per 100,000 words in All.	Real
30	Summ /sˤumm/	صُم	Deaf	1.13 instances of صمّ per 100,000 words in All.	Real
31	qunn /qunn/	قُن	A small mountain	1.37 instances of قن per 100,000 words in All.	Real
32	Dhurr /dˤurr/	ضُر	Harm	0.98 instances of ضر per 100,000 words in All.	Real
33	qull /qull/	قُل	Small, short	13.09 instances of قل per 100,000 words in All.	Real
34	Hubb /ħubb/	حُب	Love	29.15 instances of حب per 100,000 words in All.	Real
35	‘uthth /ʕuθθ/	عُث	Mite/moth (pl.)	5.22 instances of عث per 100,000 words in All.	Real
36	‘ubb /ʕubb/	عُب	Sleeves (of a shirt)	10.92 instances of عب per 100,000 words in All.	Real
37	shaaTHTH /ʃa:ðð/	شَاذ	Odd	0.83 instances of شاذ per 100,000 words in All.	Real
38	haam /ha:m/	هَام	Important	3.27 instances of هام per 100,000 words in All.	Real
39	naab /na:b/	نَاب	Canine	0.44 instances of ناب per 100,000 words in All.	Real
40	Daar /dˤa:rr/	ضَار	Harmful	0.74 instances of ضار per 100,000 words in All.	Real
41	qaar /qa:r/	قَار	Tar	0.69 instances of قار per 100,000 words in All.	Real
42	Saall /sˤa:ll/	صَال	Zebra with loud sound	0.29 instances of صالّ per 100,000 words in All.	Real
43	qaab /qa:b/	قَاب	Gap	0.6 instances of قاب per 100,000 words in All.	Real
44	Saaw /sˤa:w/	صَاو	Empty	0.01 instances of صاو per 100,000 words in All.	Real
45	qaat /qa:t/	قَات	Qat/khat/kat (plant)	0.16 instances of قات per 100,000 words in All.	Real
46	Haar /ħa:rr/	حَار	Hot	1.84 instances of حار per 100,000 words in All.	Real
47	‘aal /ʕa:l/	عَال	High	5.5 instances of عال per 100,000 words in All.	Real
48	Haaff /ħa:ff/	حَاف	Dry	0.22 instances of حافّ per 100,000 words in All.	Real
49	riif /ri:f/	رِيف	Countryside	5.38 instances of ريف per 100,000 words in All.	Real
50	diik /di:k/	دِيك	Rooster	2.85 instances of ديك per 100,000 words in All.	Real
51	tiin /ti:n/	تِين	Fig (pl.)	0.55 instances of تين per 100,000 words in All.	Real
52	Tiib /tˤi:b/	طِيب	Scent	8.87 instances of طيب per 100,000 words in All.	Real
53	qiir /qi:r/	قِير	Tar	1.49 instances of قير per 100,000 words in All.	Real

54	Tiin /tˤiːn/	طِين	Clay	1.8 instances of طين per 100,000 words in All.	Real
55	qiid /qiːd/	قِيد	Little amount	6.4 instances of قيد per 100,000 words in All.	Real
56	Tiil /tˤiːl/	طِيل	Rope for tying camels	0.02 instances of طيل per 100,000 words in All.	Real
57	qiib /qiːb/	قِيب	Amount	0 instances of قيب per 100,000 words in All.	Real
58	‘iir /ʕiːr/	عِير	Camels	1.1 instances of عير per 100,000 words in All.	Real
59	‘iis /ʕiːs/	عِيس	White camels (certain camel breed)	0.54 instances of عيس per 100,000 words in All.	Real
60	‘iih /ʕiːh/	عِيه	Calling for camels to stop doing something	0.71 instances of عيه per 100,000 words in All.	Real
61	thuum /θuːm/	ثُوم	Garlic	0.34 instances of ثوم per 100,000 words in All.	Real
62	kuuz /kuːz/	كُوز	Cone	0.25 instances of كوز per 100,000 words in All.	Real
63	suud /suːd/	سُود	Black (adj., pl.)	16.07 instances of سود per 100,000 words in All.	Real
64	Tuun /tˤuːn/	طُون	Water abundance	1.33 instances of طون per 100,000 words in All.	Real
65	quub /quːb/	قُوب	Impetigo (disease)	0 instances of قوب per 100,000 words in All.	Real
66	Tuub /tˤuːb/	طُوب	Brick	0.53 instances of طوب per 100,000 words in All.	Real
67	quur /quːr/	قُور	Small mountains; certain mounds	0.13 instances of قور per 100,000 words in All.	Real
68	Tuum /tˤuːm/	طُوم	Death; grave	0.12 instances of طوم per 100,000 words in All.	Real
69	quuf /quːf/	قُوف	Top part of ear/neck	1.71 instances of قوف per 100,000 words in All.	Real
70	Huut /ħuːt/	حُوت	Whale	0.73 instances of حوت per 100,000 words in All.	Real
71	Huub /ħuːb/	حُوب	Sin	0.32 instances of حوب per 100,000 words in All.	Real
72	Huuh /ħuːh/	حُوه	----------------------------	**Not applicable**	**Non-word**

Table 5-3: Monosyllabic words frequencies

In the experiment, three parallel sets of lexical items were formed corresponding to the three stimulus presentation conditions so that each of the listener groups would perform the same set of tasks. Furthermore, only lexical items conforming to the Arabic syllable structure were selected per consonantal type.

For instance, for the monosyllabic words, the shapes **C**VC or **C**VC2C2 were chosen where the coda position is either filled with one consonant or a geminate since CVCC underlyingly is the canonical syllable shape of Arabic (Ratcliffe, 2013). For the three long Arabic vowels /a:/, /i:/ and /u:/, the same syllable shapes of the short vowels were used. That is, these included **C**VVC and **C**VVC_2C_2 in the different consonantal environments; plain, emphatics, q and pharyngeals.

The selection of the experiment's words was done in accordance with Greenberg's (1950) asymmetry of the patterning of root morphemes in Semitic given in **4**. below.

(4) Greenberg's asymmetry of the root morpheme patterning in Semitic

1. In the 1st and 2nd consonantal positions, identical and homorganic consonants are excluded. (e.g. *mmd, *#bm, *#gk and *#rl): OCP
2. In the 2nd and 3rd consonantal positions, only homorganic consonants are excluded but identical consonants (geminates) are allowed. (e.g. ʃkk 'to split', but *ʃkg)
3. In 1st and 3rd consonantal positions, identical and homorganic consonants are marked compared to other positions. (e.g. qlq)

According to Greenberg's generalizations, the occurrence of emphatics or uvulars with pharyngeals in the same word was avoided altogether (generalization 1.). In addition, words were carefully chosen not to be dialectal but rather educated Arabic (MSA/Classical Arabic). This was done through consulting an online Arabic dictionary, namely *arabdict* (2008) which is based on famous Arabic dictionaries and thesauruses in Arabic grammar and eloquence such as *almu'jam alwaseeT*, *almuheeT fi allughah*, *taj al'aruus*, *lisaan al'arab*, *mukhtaar aSSaHaaH*, *mu'jam al'aSwaat*, *kalimaat alqur'aan tafseer wa bayaan* to mention but a few.

Four more criteria were used when constructing the stimulus material. One, the predicted loanword in Turkish should not be a Modern Standard Turkish word, similar to them or ethically inappropriate in Turkish. This was done by checking an online modern Turkish dictionary (Turkish, 2015) and asking a native Turkish speaker to verify that they abide by the set criteria (usage and being appropriate). Two, the Arabic source words should not have been borrowed by the Ottomans. This was achieved by consulting an online etymological dictionary, namely Sevan (2007). Three, the Arabic source words were carefully chosen to be either nouns or adjectives but not verbs since the Arabic borrowed words into Turkish were either nouns or adjectives but never verbs. Four, some of the Arabic source words were deliberately chosen on the basis of similarity in pronunciation to existing ALT words.

To recap, the criteria used when selecting the stimuli are summarized as follows.

1. Source words should be Classical/Modern Standard Arabic but not dialectal ones.
2. Predicted loanwords (real and nonsense) should not be Modern Standard Turkish words and should be ethically appropriate.
3. Real Arabic source words should not have been borrowed by the Ottomans.
4. Arabic source words should conform to Greenberg's root morpheme patterning in Semitic.
5. Arabic source words should be either nouns or adjectives but not verbs.
6. Arabic source words should bear some resemblance to already integrated words in the ALT e.g., zirr (button) being similar to sirr (Arabic)/sɯr 'secret' (ALT), tiin 'figs' (Arabic) similar in pronunciation to *din* (ALT) 'religion', haam 'important' (Arabic) similar in pronunciation to tam (ALT) 'complete'.

In what follows, the two remaining stimuli tables are provided for polysyllabic distractors, tables 5-4 and 5-5, in line with the criteria given above.

Plain consonants CV**C**VC i-i, i-a, a-i	***Arabic word***	***English Glossary***	***Emphatics and q*** CV**C**VC i-i, i-a, a-i	***Arabic word***	***English Glossary***	***Guttural*** CV**C**VC i-i, i-a, a-i	***Arabic word***	***English Glossary***
ʔibil>ibil namir>nemir̥ mihan>mihen	إِبِل نَمِر مِهَن	camels tiger occupations	ʔi**t**ˤil>ɯtɯl tiqir>tɯkɯr̥ fatˤin> fatɯn naqih>nakih ʔitˤam> ɯtam niqam>nɯkam	إِطِل تقر) تِقِر (تقر فَطِن نَقِه إِطَم نِقَم	waist laughing shrewd Intelligent **Non-word** Indignation (pl.)	tiʕim>tiim waʕil>vail fiʕab>fiap	تِعِم وَعِل فِعَب	**Non-word** Ibex (mountain goat) **Non-word**

Table 5-4: Polysyllabic distractor real and non-words

S.no	Polysyllabic Word	Word in Arabic	English glossary	Frequency	Word nature
1	‘ibil /ʔibil/	إِبِل	Camel (pl.)	4.56 instances of إبل per 100,000 words in All.	Real
2	‘iTil /ʔitˤil/	إِطِل	Waist	2.18 instances of إطل per 100,000 words in All.	Real
3	tiqir /tiqir/	تِقِر	Laughing	6.91 instances of تقر per 100,000 words in All.	Real
4	ti'im /tiʕim/	تِعِم	------------	**Not applicable**	**Non-word**
5	namir /namir/	نَمِر	Tiger	1.45 instances of نمر per 100,000 words in All.	Real
6	faTin /fatˤin/	فَطِن	Shrewd	14.72 instances of فطن per 100,000 words in All.	Real
7	naqih /naqih/	نَقِه	Intelligent	14.55 instances of نقه per 100,000 words in All.	Real
8	wa'il /waʕil/	وَعِل	Ibex	40.89 instances of وعل per 100,000 words in All.	Real
9	mihan /mihan/	مِهَن	Occupations	1.48 instances of نقم per 100,000 words in All.	Real
10	‘iTam /ʔitˤam/	إِطَم	--------------	**Not applicable**	**Non-word**
11	niqam /niqam/	نِقَم	Indignation (pl.)	1.68 instances of نقم per 100,000 words in All.	Real
12	fi'ab /fiʕab/	فِعَب	--------------	**Not applicable**	**Non-word**

Table 5-5: Word frequency of polysyllabic distractor real and non-real words

5.2.2 Participants

As mentioned in the introduction, 51 responses were recorded after eliminating all the data that did not adhere to the research criteria. These were further divided among the three listener groups as follows: 18 T (monolinguals), 11 TA (bilinguals) and 22 TQ (Quranic Turkish speakers).

The selection criteria were rigorous and followed throughout the process of data collection in order to ensure uniformity of listeners within each of the listener groups. Noteworthy to mention is that none of the participants had any physical or language problems or impairments preventing them from undertaking the survey. However, some of them reported either wearing glasses, having their vision corrected, having undergone Cataract radiation therapy (one participant) or Rhoticism (one listener).

As explained in the PAT chapter, the three participant groups were chosen to simulate the linguistic background of the Ottomans during whose time Arabic words infiltrated into Turkish. The Ottomans spoke Turkish, Arabic in addition to Persian. Thus, one of the goals of this study is to check which listener group would resemble the grammar of the Ottomans.

In the T group (18), only monolingual participants were selected, meaning that if a participant spoke a language or languages other than Turkish, their data were not used. Such languages include German, Hebrew, Spanish, Kurdish, Armenian, Tatar and English. In this respect, the monolingual group comprised 'naïve listeners' who recognized Arabic words only as borrowings. As in the PAT, participants were asked to fill in a language questionnaire (appendix A) in which they were asked to identify their first, second and third language, their parents', which language the participants spoke at home and whether they knew Arabic from reciting the Quran. The T group participants in the survey came from different parts of Turkey including Aksaray, Ankara, Bursa, Gaziantep, Istanbul, Kırklareli, Konya, Mersin, Muğla, Rize, Sivas and Kütahya.

As for the TA group (11), only participants raised as bilinguals or whose parents or one of them had Arabic as their first language were chosen. These are the ones who indicated in the survey that they learned Arabic at any early age and spoke an Arabic dialect at home in addition to Turkish. The bilingual group participants in the survey came from places in Turkey such as Antakya, Düzce, İskenderun, Mardin and şanlıurfa.

Moreover, for the TQ group (22), the data for those who indicated they knew Arabic through Quranic recitation were incorporated, leaving out those who learned Arabic as a 2nd or a foreign language. As in the PAT survey, the participants in this group either learned Arabic at the age of 7 or 8 in madrasas (Quran schools), learned it in high school or as an elective course for the purposes of reading the Quran. The participants came from areas in Turkey such as Afyon, Ağrı, Antalya, Artvin, Yozgat, Diyarbakır, İstanbul, İzmir, Kayseri, Malatya, Samsun, Sivas and Trabzon.

The participants' sociolinguistic data and their language proficiency ratings are given in table 5-6 below.

Participant	Age	Gender	Highest level of education	T. Average (1-10)	A. Read. (1-10)	A. Comp. (1-10)	A. Speak. (1-10)	A. Write. (1-10)	A.Average (1-10)
T	19-56	5m:13f	5H,12C,0M,1 Other	9.94	N/A	N/A	N/A	N/A	N/A
TA	22-57	7m:4f	3H,6C,1M,1D	9.47	6.27	5.81	5.36	4.09	6.2
TQ	17-62	16m:6f	6H,12C,4M	9.12	7.36	3.27	3.22	3.9	5.374
Average	19:58	9m:8f	4.6H,10C,2M, 0.3D, 0.3 Other	9.51	6.815	4.54	4.29	3.995	5.787

Table 5-6: Summary of the language questionnaire

H: high school, C: college, M: master's, D: doctorate, O: other
A.: Arabic, T.: Turkish, Read.: reading, Write.: writing, Comp.: comprehension, Speak.: speaking

5.2.3 Rationale and procedure

The rationale of the SB experiment is to measure the degree of match in the SB to the corpus, which is then compared to the PAT in order to make inferences about the three models of loanword adaptation; i.e., the Perception only model, the Phonology only or the hybrid model of both Perception and Phonology. This would ultimately help in establishing which model to utilize as will be shown in the discussion chapter.

As mentioned in the introduction, the SB survey was run online via Qualtrics website in two phases; one by the researcher herself and a later one by the Qualtrics team. In the SB study, different types of stimulus were used: audio only (A), written only (W), and audio and written (AW). Worth mentioning is that the stimuli were presented in this order: A, AW and then W since the T group participants had to be redirected to the end of the survey after the AW task.

In the audio only task (A), all three listener groups listened to recordings of Arabic monosyllabic words (real or nonsense) produced by a native speaker of Arabic, and were then asked to write them down in Turkish spelling. This is a contextualized version of the PAT (monosyllabic words being real words in addition to nonsense words) which would enable the researcher to test whether the participants would rely only on their perception of the vowels in the source words.

This task was a replication of the PAT but using a wider mix of real and nonsense monosyllabic words and with a different task (write it down). In Turkish, ‘write it down’ is the same as ‘choose which vowel’ since the spelling system of Turkish is phonemic for vowels.

In the written only stimuli (W), the TA and TQ participants only saw a list of monosyllabic real and nonsense words written in Arabic script and were then asked to write them down in Turkish spelling. The T group participants were not asked to perform this task since they had no access to Arabic orthography.

In the audio-written task, all three participants: T, TA and TQ listened to recordings of Arabic real or nonsense monosyllabic words produced by a native speaker of Arabic, accompanied by the word on screen in Arabic script. They were asked to write down in Turkish the word they heard and saw. This task tested whether spelling made a difference. The purpose of this task was to gauge what the listeners did when they had only one condition to rely on, and which they relied on more when they had both.

The hypothesis here was that spelling would have an effect of some kind as bilingual participants or second language learners (TQ) would use their knowledge of Arabic orthography and/ phonology in perceiving the new words they heard.

The methodology section concludes with a comparison between the PAT and the audio condition of the SB experiment before presenting the results of each in the data analysis section. Tables 5-7 and 5-8 summarize the methodology used in each experiment

Research methodology

task name	stimuli	who does it?	what do they do?	why?
PAT **Perceptual Assimilation Task**	➢ audio (only) = recordings of all Arabic vowels (including emphatic allophones), produced by a native speaker of Arabic, in hVd words. 12 Arabic VX 3 times+ 4 cardinal vowels (Turkish) (X 3 distractors)	**T:** Monolingual Turkish speakers **TA:** Turkish-Arabic bilingual speakers of Arabic dialect in Turkey) **TQ:** Turkish speakers with some knowledge of Arabic mainly from recitation of the Qur'an	listen to the vowel - what vowel did you hear? = choose from a set of 8 Turkish vowels	this is the core of the evidence to support the 'loanword adaptation as perception' argument ***Hypothesis:*** the patterns of assimilation will match the mappings among vowels observed qualitatively in the loanword corpus

Table 5-7: PAT methodology

task name	stimuli	who does it?	what do they do?	why?
SB **Simulated Borrowing**	➢ audio (only) = recordings of Arabic words (real and nonsense), produced by a native speaker of Syrian Arabic Real monosyllabic words: 23 Nonsense monosyllabic words: 1	**T:** Monolingual Turkish speakers **TA:** Turkish-Arabic bilingual speakers of Arabic dialect in Turkey) **TQ:** Turkish speakers with some knowledge of Arabic mainly from recitation of the Qur'an	listen to the word - write down in Turkish the word they	-This is a contextualized version of the PAT. -This is a replication of the PAT but using real monosyllabic words and with a different task (write it down). In Turkish, 'write it down' is the same as 'choose which vowel' since the spelling system of Turkish is phonetic, for vowels at least. -***Hypothesis:*** the patterns of assimilation will match the mappings among vowels observed qualitatively in the loanword corpus.

Table 5-8: SB-audio methodology

In section 5.3.1.2. a logistic regression is run in order to measure the effects of the different variables examined and their interactions on the DV match. As elaborated in chapter 4 and mentioned in chapter 6, this procedure follows two main stages. In the first stage, the objective of the logistic regression along with the hypotheses are outlined. This is followed with a protocol (Zuur et al ((2009), Winter (2013), Baayen (2008) and Barr et al (2013)) for selecting the final model to be used in the analysis. The protocol itself involves three main steps, namely i) constructing the fixed and random effects structures, ii) constructing the maximal model and running the logistic regression models and finally iii) validating the results and reporting them.

5.3 Data analysis

5.3.1 Results and discussion

When the results of the Simulated Borrowing experiment were collected for analysis, these were entered in an excel spread sheet and coded per the different variables of the study. Statistical analysis was run on the data, however after the raw data was explored to see different patterns, especially those related to the listener group variable.

5.3.1.1 Raw data of PAT and SB-audio experiments

This subsection compares the results of the raw data for both the PAT and SB-audio experiments. Table 5-9 depicts actual counts of PAT listeners' tokens. The green shade indicates a match; i.e., the response vowel is the same as the predicted vowel based on the research corpus. On the other hand, the red shade designates a mismatch.

As can be seen from table 5-9, the three listener groups demonstrate the same patterns of assimilation as can be interpreted from the position of the red and green shades. This suggests that listener group as a variable does not have a significant effect on match as was shown in chapter 4. As for the assimilation patterns, only three vowel categories were mismatched by the three groups; namely [a:]$_{T}$>/e/$_{T}$ (predicted /a/), [iˤ]>/e/$_{T}$ (predicted /ɯ/) and [u>]</u/$_{T}$ (predicted /y/). In chapter 4, it was concluded that the perception only model could account for most of the assimilation patterns (68% match).

Listgp	T								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	130	99	3		3		10	1	246
had	20	98	1		2		1	1	123
haDD	110	1	2		1		9		123
a:	120	106	6	3	2	1	8		246
haad	10	106	3	1	2	1			123
haaDD	110		3	2			8		123
i	4	81	124	15	7	6		9	246
hid	1	2	110	5	3	2			123
hiDD	3	79	14	10	4	4		9	123
i:	3	3	229	6	3		2		246
hiid	1	1	115	3	2		1		123
hiiDD	2	2	114	3	1		1		123
u	3	2	6	7	218	8	2		246
hud	2	1	5	4	102	8	1		123
huDD	1	1	1	3	116		1		123
u:	4		2	5	229	4	2		246
huud	1		2	2	115	2	1		123
huuDD	3			3	114	2	1		123
Total	264	291	370	36	462	19	24	10	1476

Listgp	TA								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	152	88	4	3	1	4	7	5	264
had	35	88	1	3		2		3	132
haDD	117		3		1	2	7	2	132
a:	132	112	7	2	3	1	7		264
haad	14	110	5	2		1			132
haaDD	118	2	2		3		7		132
i	6	72	129	30	7	1	2	17	264
hid	1	7	111	10	3				132
hiDD	5	65	18	20	4	1	2	17	132
i:	2	4	236	7	5	6	2	2	264
hiid	2	2	119	2	2	4		1	132
hiiDD		2	117	5	3	2	2	1	132
u	2	5	8	7	221	20	1		264
hud	2	1	5	5	101	18			132
huDD		4	3	2	120	2	1		132
u:	4	4	6	2	232	6	8	2	264
huud	3		2	1	119	4	1	2	132
huuDD	1	4	4	1	113	2	7		132
Total	298	285	390	51	469	38	27	26	1584

Listgp	TQ								
SV/RV	a	e	i	ı	u	ü	o	ö	Total
a	441	345	11	12	17	12	15	5	858
had	63	335	7	8	9	3	3	1	429
haDD	378	10	4	4	8	9	12	4	429
a:	408	372	11	16	9	9	19	14	858
haad	35	363	5	6	6	3	3	8	429
haaDD	373	9	6	10	3	6	16	6	429
i	21	252	394	99	16	23	13	40	858
hid	6	15	364	26	3	9	5	1	429
hiDD	15	237	30	73	13	14	8	39	429
i:	13	20	760	24	20	13	4	4	858
hiid	7	8	386	9	8	6	2	3	429
hiiDD	6	12	374	15	12	7	2	1	429
u	13	12	18	43	697	53	12	10	858
hud	9	5	11	29	322	45	3	5	429
huDD	4	7	7	14	375	8	9	5	429
u:	12	14	18	16	745	14	31	8	858
huud	3	6	13	9	377	12	5	4	429
huuDD	9	8	5	7	368	2	26	4	429
Total	908	1015	1212	210	1504	124	94	81	5148

Table 5-9: Summary confusion matrix of the PAT results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; RV= response vowel; Listgp= listener group

T= monolingual listeners; TA= bilingual listeners; TQ= Turkish Quranic listeners

As for the SB experiment, the raw results are given in tables 5-10 through 5-12 where again the green shade reflects a match and the red shade a mismatch between listeners' responses and predicted mappings based on the research corpus. One difference between table 5-9 and tables 5-10 to 5-12 is that the responses in the PAT experiment are restricted to the 8 Turkish vowels since it is a forced choice task. Conversely, the SB was an open choice task since the listeners were not limited to the 8 Turkish vowels but rather extended their responses to include in addition to the 8 Turkish vowels long vowels (which violates native Turkish grammar), syllabified vowels (2 or even 3 identical or different vowel categories) and diphthongs. This is manifested in the second row of the SB tables below.

This means that although the stimulus words were monosyllabic, the response words produced could be monosyllabic, disyllabic or even polysyllabic. However, this was mostly done when the stimulus vowel was long. For example, for the input *Tiib* [ii], the responses included monosyllabic words such as dıf, pliv, tib, tüp and tur; and disyllabic words (since vowel length is not used in Turkish) such as tııb, tiip, sıyyıb, tayip, tayyib, tayyip, teyip, tıiyb, Tıyb, tıyib, tıyip, tıyp, tıyyb, tıyyıb, toyyip, tuib, tuyib, Tuyip and tuyyip.

Noteworthy to mention here is that the subset of the data shown in tables 5-10 to 5-12 and analyzed in the next section pertains to monosyllabic audio words only.

Listgp		T																												
SV/RV		**a**	**aa**	**a-a**	**a-e**	**Ai**	**a-ɪ**	**E**	**ee**	**e-i**	**i**	**ɪ**	**ii**	**ɪ-ɪ**	**i-i**	**i-i-i**	**o**	**ö**	**oa**	**o-a**	**oo**	**U**	**Ü**	**u-a**	**ui**	**u-i**	**uu**	**u-u**	**ü-ü**	**Total**
A	man	2						15			1																			18
	Dhabb	13															5													18
	Qad	16															1	1												18
	?an	13		1	1			2	1																					18
a:	SHaathth	6					1	9	2																					18
	Dhaarr	8	3																2	4				1						18
	Qaarr	12	6																											18
	Haarr	15	2				1																							18
I	Min							3			15																			18
	Dhidd							2			13	2									1									18
	Qidd							3			14	0										1								18
	?iyy	1				3	1	13			0																			18
i:	Riif										8		7		2	1														18
	Tiibb						2			2	2			2	1							1	1		1	6				18
	Qiirr						3				0			11	3											1				18
	?iirr						16			1	1																			18
U	Muthth										1						5	1				11	0							18
	Suyy																3					15								18
	Qudd																2				1	15								18
	Hubb	1						1									11	3				2								18
u:	Thuumm																1					15					1	1		18
	Tuunn																					10		1			4	2	1	18
	Quubb																					15					3			18
	Huutt																				1	12			1		4			18
	Total	**87**	**11**	**1**	**1**	**3**	**24**	**48**	**3**	**3**	**55**	**2**	**7**	**13**	**6**	**1**	**28**	**5**	**2**	**4**	**3**	**97**	**1**	**2**	**2**	**7**	**12**	**3**	**1**	**432**

Table 5-10: Summary confusion matrix of T group mapping patterns in the SB-Audio experiment

Listgp		TA																						
SV/RV		**A**	**Aa**	**a-ɪ**	**a-u**	**e**	**ee**	**e-i**	**I**	**ɪ**	**ɪe**	**ii**	**ɪ-ɪ**	**i-i**	**ɪ-v**	**o**	**ö**	**oa**	**ou**	**u**	**ü**	**ui**	**uu**	**Total**
a	man	2				9																		11
	Dhabb	9				2																		11
	qad	9							1								1							11
	?an	6				4			1															11
a:	SHaaththh	3				5	3																	11
	dhaarr	7	2													1		1						11
	qaarr	7	4																					11
	Haarr	7	4																					11
i	min								10		1													11
	dhidd					2			3	6														11
	qidd	1				1			8	1														11
	?iyy	1				4			6															11
i:	riif								10				1											11
	Tiibb			1					5	3		1	1											11
	qiirr								4	2		1	2	2										11
	?iirr	1		4				2	1	2			1											11
u	muthth															2			2	7	0			11
	Suyy			1					2			1			1					6				11
	qudd								1											10				11
	Hubb								2							2				6		1		11
u:	thuumm																			7		1	3	11
	Tuunn																			7			4	11
	quubb																			10			1	11
	Huutt				1															8		1	1	11
	Total	**53**	**10**	**6**	**1**	**27**	**3**	**2**	**54**	**14**	**1**	**3**	**5**	**2**	**1**	**5**	**1**	**1**	**2**	**61**		**3**	**9**	**264**

Table 5-11: Summary confusion matrix of TA group mapping patterns in the SB-Audio experiment

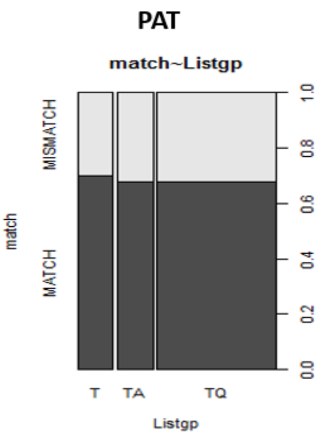
Listgp		TQ																																							
SV		a	aa	a-a	ai	a-ɪ	a-u	e	ea	ee	e-e	e-i	eɪ	e-ɪ	i	ɪ	ɪe	ii	ɪ-ɪ	i-i	ɪɪ	ɪ-ɪ	iu	i-ü	o	ö	oa	o-a	o-i	oö	u	ü	ua	ui	u-i	uɪ	uu	u-u	uü	w	Total
a	man	1					1	19							1																										22
	Dhabb	19																							3																22
	qad	21																													1										22
	?an	18	1	1				1								1																									22
a:	SHaathth	1	1					15		5																															22
	dhaarr	14	2																								2	3					1								22
	qaarr	14	7	1																																					22
	Haarr	13	8					1																																	22
i	min							3							18									1																	22
	dhidd							1							9	11															1										22
	qidd	1						2							17	2																									22
	?iyy	2			3	1		10	1				2		1	2																									22
i:	riif														14	1		7																							22
	Tiibb					1									7	7		1		1	1								1					2	1						22
	qiirr														6	6	1		5	2	1	1																			22
	?iirr	1				14		1			1	3		1	0		1																								22
u	muthth																								6						15	1									22
	Suyy														5	1									4						9				1	1				1	22
	qudd																														21						1				22
	Hubb							2	1						1										5	1				1	11										22
u:	thuumm																						1					1			15						4	1			22
	Tuunn														1																15						5		1		22
	quubb																														18						4				22
	Huutt																														17	1					4				22
	Total	105	19	2	3	16	1	55	2	5	1	3	2	1	80	31	2	8	5	3	2	1	1	1	18	1	2	4	1	1	123	2	1	2	2	1	18	1	1	1	528

Table 5-12: Summary confusion matrix TQ group mapping patterns in the SB-Audio experiment

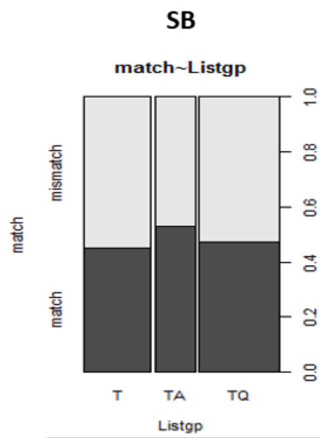
Studying the SB tables and figure 5-1, we can make the following observations.

1. Match across the three groups is high in both the PAT and the SB-audio task at 68% and 48% respectively. Figure 5-1 below compares the match results of both the PAT and SB tasks across and within the three listener groups.
2. In the PAT, the three listener groups showed the same mapping patterns with similar match scores where the T group slightly projected higher than the TA and TQ groups; T: 69.7%, TA: 67.8% and TQ:67.6%. However, in the SB-audio task, the TA group exhibited higher degrees of match at 53% followed by the TQ at 47% and T group at 45%.
3. For the /a:/ vowel in the plain environment, all three groups mismatched it to /e/ (predicted /a/) both in the PAT and SB-audio.
4. For the /u/ vowel in the plain environment, all three groups mismatch it to /u/ (predicted y <ü>) both in the PAT and SB-audio.
5. In the Sb-audio task, only the T group mismatched the vowel /u/ in the pharyngealized environment [uˤ]ʕ to /o/ (predicted /u/).

Percentage of match across the groups (PAT)	
match	mismatch
5587	2621
68.07%	31.93%



	match	mismatch
T	69.78%	30.22%
	1030	446
TA	67.80%	32.20
	1074	510
TQ	67.66%	32.34
	3483	1665



Percentage of match across the groups (SB)	
match	mismatch
584	640
47.71%	52.29%

	match	mismatch
T	45.14%	54.86%
	195	237
TA	53.03%	46.97%
	140	124
TQ	47.16%	52.84%
	249	279

Figure 5-1: correlation of Listgp with match in the PAT and SB-audio tasks

Moreover, the results were split by the vowel quality as shown in table 5-13 and figures 5-2 and 5-3. The following observations can be made.

1. In the PAT, vowels /i/ (70%,) and /u/ (69.9%) were matched to the corpus mappings more than /a/ (65%) whereas in the SB-audio task, /a/ (58%) and /u/ (60%) were matched more than was the /i/ (26%) to the corpus patterns. (table 5-13)
2. More variation is found in the TQ group for each of the three vowels (figure 5-3).

			match		mismatch					match		mismatch	
Task	**Vowel**	**Listgp**	**%**	**count**	**%**	**count**	**Task**	**Vowel**	**Listgp**	**%**	**count**	**%**	**count**
audio	a	T	59.722	86	40.277	58	**PAT**	a	T	66.666	328	33.333	164
		TA	56.818	50	43.181	38			TA	63.825	337	36.174	191
		TQ	57.954	102	42.045	74			TQ	65.326	1121	34.673	595
		average	58.165	79.33	41.83	56.66			**average**	65.272	595.333	34.727	316.666
	i	T	17.361	25	82.638	119		i	T	70.934	349	29.065	143
		TA	40.909	36	59.09	52			TA	69.507	367	30.492	161
		TQ	22.727	40	77.272	136			TQ	69.755	1197	30.244	519
		average	**26.999**	33.66	73	102.33			**average**	70.065	637.666	29.934	274.333
	u	T	58.333	25	41.666	119		u	T	71.747	353	28.252	139
		TA	61.363	36	38.636	52			TA	70.075	370	29.924	158
		TQ	60.795	40	39.204	136			TQ	67.89	1165	32.109	551
	average		60.16	33.66	39.83	102.33			**average**	69.904	629.333	30.095	282.666

Table 5-13: Percentage and count of match across Listgp per vowel in the audio and PAT tasks

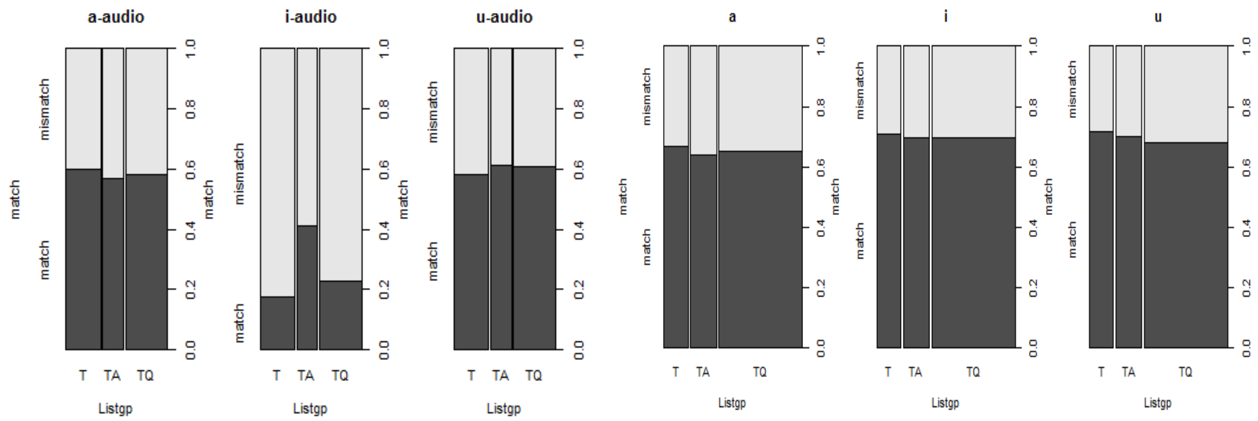


Figure 5-2: Barplots of the SB audio and PAT data split by data across the three listener groups

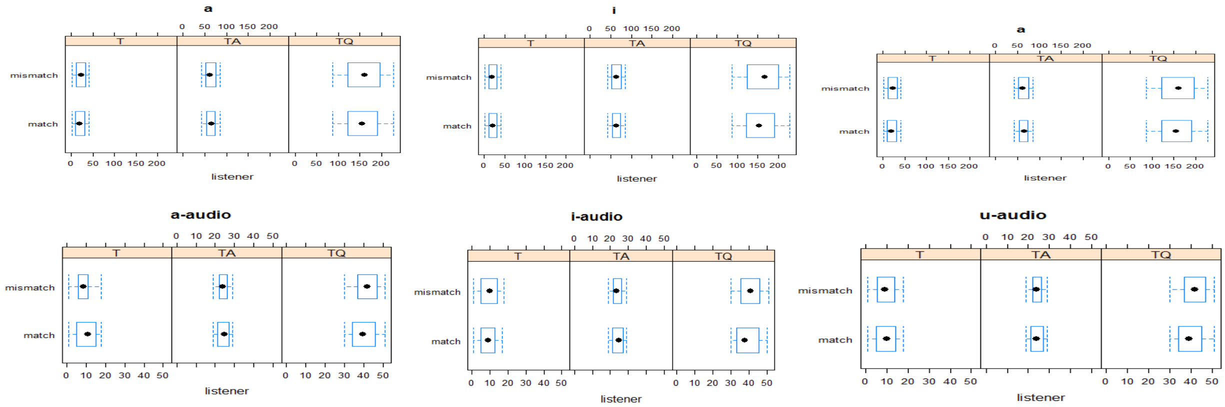


Figure 5-3: bwplots of match~Listgp per vowel in the audio and PAT tasks

In addition to the above, tables 5-14 to 5-16 illustrate the perceptual maps of the three listener groups in the SB-A experiment.

In the tables below, the counts after each arrow include the actual responses yielded by the listeners in addition to variations such as long vowels and syllabified vowels which indicate that the listeners can hear both the vowel duration and quality respectively. For example, in table 5-14, the listeners in the T group mapped $[a:]_{ʕ}$ onto $/a/_{T}$ in 15 cases, $/aa/_{T}$ (in two syllables) in 2 case and as $/a\text{-}ɪ/_{T}$ in 1 case. Moreover, solid arrows represent categorical (consistent) mapping which is indicated by a percentage between 80-100. Conversely, dashed arrows symbolize gradient (variable) mapping with a percentage below 80. One important point here is that, in contrast to the PAT experiment, the context of vowels in the SB-A experiment includes four categories, namely plain and gutturals; i.e., emphatics, q and pharyngeals, the gutturals being represented with subscripted $_{dˤ}$, $_{q}$ and $_{ʕ}$ respectively.

[a]→ (15/18) /e/$_T$ (83.33%) {e:15, a:2, i:1} [a:] --> (11/18) /e/$_T$ (61.11%) {e: 9, ee:2, a:6, a-ı:1}	[a]$_{d^{ʕ}}$ -->(13/18) /a/$_T$ (72.22%) {a:13, o:3} [a]$_q$ → (16/18) /a/$_T$ (88.88%) {a:13, a-a:1, a-e:1, e:2, ee:1} [a]$_ʕ$ → (15/18) /a/$_T$ (83.33%) {a: 13, a-a: 1, a-e: 1} [a:]$_{d^{ʕ}}$ → (18/18) /a/$_T$ (100%) {a:8, aa:3, oa:2, o-a:4, u-a:1} [a:]$_q$ → (18/18) /a/$_T$ (100%) {a:12, aa:6} [a:]$_ʕ$ → (18/18) /a/$_T$ (100%) {a:15, aa:2, a-ı: 1}	[u]--> (11/18) /u/$_T$ (61.11%) {u:11, o:5, ö:1, i:1} [u]$_{d^{ʕ}}$ →(15/18) /u/$_T$ (83.33%) {u:15, o:3} [u]$_q$ → (15/18) /u/$_T$ (83.33%) {u:15, o:2, oo:1} [u]$_ʕ$ --> (11/18) /o/$_T$ (61.11%) {o:11, ö:3, e:1, a:1, u:2 } [u:] → (17/18) /u/$_T$ (94.44%) {u:15, uu:1, u-u:1, o:1} [u:]$_{d^{ʕ}}$ → (17/18) /u/$_T$ (94.44%) {u:10, uu:4, u-u:2, u-a:1, ü-ü:1} [u:]$_q$ → (18/18) /u/$_T$ (100%) {u:15, uu:3} [u:]$_ʕ$ → (17/18) /u/$_T$ (94.44%) {u:12, ui:1, uu:4, oo:1}	[i] → (15/18) /i/$_T$ (83.33%) {i:15, e:3} [i]$_{d^{ʕ}}$-->(13/18) /i/$_T$ (72.22%) {i:13, e:2, ı:2, oo:1} [i]$_q$ --> (14/18) /i/$_T$ (77.77%) {i:14, e:3, u:1} [i]$_ʕ$ --> (13/18) /e/$_T$ (72.22%) {e:13, a:1, ai:3, a-ı:1} [i:] → (18/18) /i/$_T$ (100%) {i:8, ii:7, i-i:2, i-i-i:1} [i:]$_{d^{ʕ}}$ --> (12/18) /i/$_T$ (66.66%) {i:2, i-i:1, e-i:2, ui:1, u-i:6, a- ı:2, ı-ı:2, u:1,ü:1} [i:]$_q$ --> (14/18) /ɯ/$_T$ (77.77%) {ı-ı:11 a- ı: 3, i-i:3, u-i:1} [i:]$_ʕ$ → (16/18) /ɯ/$_T$ (88.88%) {a-ı: 16, i:1, e-i:1}

Table 5-14: perceptual maps of T group in the Audio condition in the Simulated Borrowing experiment

[a]→ (9/11) /e/$_{T}$ (81.81%) {a:2, e:9} [a:] --> (8/11) /e/$_{T}$ (72.72%) {e: 5, ee:3, a:3}	[a]$_{dˤ}$ → (9/11) /a/$_{T}$ (81.81%) {a:9, e:2} [a]$_{q}$ → (9/11) /a/$_{T}$ (81.81%) {a:9, i:1, ö:1} [a]$_{ʕ}$ --> (6/11) /a/$_{T}$ (54.54%) {a:6, e:4, i:1} [a:]$_{dˤ}$ → (10/11) /a/$_{T}$ (90.90%) {a:7, aa:2, oa:1, o:1} [a:]$_{q}$ → (11/11) /a/$_{T}$ (100%) {a:7, aa:4} [a:]$_{ʕ}$ → (11/11) /a/$_{T}$ (100%) {a:7, aa:4}	[u]→ (9/11) /u/$_{T}$ (81.81%) {u:7, ou: 2, o:2} [u]$_{dˤ}$ -->(6/11) /u/$_{T}$ (54.54%) {u:6, a-ɪ:1, i:2, ii:1, ɪ-v:1} [u]$_{q}$ → (10/11) /u/$_{T}$ (90.90%) {u:10, i:1} [u]$_{ʕ}$ --> (7/11) /u/$_{T}$ (63.63%) {u:6, ui:1, o:2, i:2} [u:] → (11/11) /u/$_{T}$ (100%) {u:7, uu:3, ui:1} [u:]$_{dˤ}$ → (11/11) /u/$_{T}$ (100%) {u:7, uu:4} [u:]$_{q}$ → (1/11) /u/$_{T}$ (100%) {u:10, uu:1} [u:]$_{ʕ}$ → (11/11) /u/$_{T}$ (100%) {u:8, ui:1, uu:1, a-u:1}	[i] → (10/11) /i/$_{T}$ (90.90%) {i:10, ɪ:1} [i]$_{dˤ}$-->(6/11) /ɯ/$_{T}$ (54.54%) {ɪ:6, i:3, e:2} [i]$_{q}$ --> (8/11) /i/$_{T}$ (72.72%) {i:8, ɪ:1, e:1 a:1 } [i]$_{ʕ}$ --> (6/11) /i/$_{T}$ (54.54%) {i:6, e:4, a:1} [i:] → (10/11) /i/$_{T}$ (90.90%) {i: 10, ɪ-ɪ:1} [i:]$_{dˤ}$ --> (6/11) /i/$_{T}$ (54.54%) {i:5, ii:1, ɪ:3, a-ɪ:1, ii:1} [i:]$_{q}$ --> (7/11) /i/$_{T}$ (63.63%) {i:4, ii:1, i-i:2, ɪ-ɪ:2, ɪ:2} [i:]$_{ʕ}$ --> (7/11) /ɯ/$_{T}$ (63.63%) {ɪ:2, a-ɪ:4, ɪ-ɪ: 1, i:1, e-i:2, a:1}

Table 5-15: perceptual maps of TA group in the Audio condition in the Simulated Borrowing experiment

[a]→ (19/22) /e/$_{T}$ (86.36%) {e:19, a:1, a-u:1, i:1} [a:] → (20/22) /e/$_{T}$ (90.90%) {e:15, ee:5, a:1, aa:1}	[a]$_{dˤ}$ → (19/22) /a/$_{T}$ (86.36%) {a:19, o:3} [a]$_{q}$ → (21/22) /a/$_{T}$ (95.45%) {a:21, u:1} [a]$_{ʕ}$ → (20/22) /a/$_{T}$ (90.90%) {a:18, aa:1, a-a:1, e:1, ɪ:1} [a:]$_{dˤ}$ → (22/22) /a/$_{T}$ (100%) {a: 14, aa:2, oa:2, o-a:3, ua:1} [a:]$_{q}$ → (22/22) /a/$_{T}$ (100%) {a:14, aa:7, a-a:1} [a:]$_{ʕ}$ → (21/22) /a/$_{T}$ (95.45%) {a:13, aa:8, e:1}	[u]--> (15/22) /u/$_{T}$ (53.57%) {u:15, o:6, ü:1} [u]$_{dˤ}$ -->(11/22) /u/$_{T}$ (50%) {u:9, u-i:1, uɪ:1, w:1, o:4, i:5, ɪ:1} [u]$_{q}$ → (22/22) /u/$_{T}$ (100%) {u:21, uu:1} [u]$_{ʕ}$ --> (11/22) /u/$_{T}$ (50%) {u:11, e:2, ea:1, i:1, o:5, ö:1, oö:1} [u:] → (21/22) /u/$_{T}$ (95.45%) {u:15, uu:4, u-u:1, o-1:1, iu:1} [u:]$_{dˤ}$ → (21/22) /u/$_{T}$ (95.45%) {u:15, uu:5, uü:1, i:1} [u:]$_{q}$ → (22/22) /u/$_{T}$ (100%) {u:18, uu:4} [u:]$_{ʕ}$ → (21/22) /u/$_{T}$ (95.45%) {u:17, uu:4, ü:1}	[i] → (19/22) /i/$_{T}$ (86.36%) {i:18, e:3, i-ü:1} [i]$_{dˤ}$-->(11/22) /ɯ/$_{T}$ (50%) {ɪ:11, i:9, e:1, u:1} [i]$_{q}$ --> (17/22) /i/$_{T}$ (77.27%) {i:17, ɪ:2, e:2, a:1} [i]$_{ʕ}$ --> (13/22) /e/$_{T}$ (59.09%) {e:10, ea:1, eɪ:2, i:1, ai:3, a:2, a-ɪ:1, ɪ:2} [i:] → (21/22) /i/$_{T}$ (95.45%) {i:14, ɪ:1, ii:7} [i:]$_{dˤ}$ --> (13/22) /i/$_{T}$ (59.09%) {i:7, ii:1, i-i:1, oi:1, a-ɪ:1, ɪ:7, ɪɪ:1, ui:2, u-i:1} [i:]$_{q}$ --> (13/22) /ɯ/$_{T}$ (59%) {ɪ:6, ɪe:1, ɪ-ɪ:5, ɪɪ:1, i:6, i-i:2, iu:1} [i:]$_{ʕ}$ --> (15/22) /ɯ/$_{T}$ (68.18%) {a-ɪ:14, ɪe:1, a:1, e:1, e-e:1, e-i:3, e-ɪ:1}

Table 5-16: perceptual maps of TQ group in the Audio condition in the Simulated Borrowing experiment

Next, a Generalized Linear Mixed Effects Modelling (GLMM) analysis is carried out in order to 1) confirm that Listener group has an effect and 2) check whether other factors (linguistic or sociolinguistic) determine the match.

5.3.1.2 *Logistic regression in R (GLMM)*

In the SB monosyllabic-audio experiment, data from 51 listeners were collected across the three listener groups yielding 1224 observations of 10 variables. The measures of association of these are given below in table 5-17. In the table below, the tokens for certain listeners and stimuli are given only as examples since it is difficult to fit the tokens of all the listeners and stimuli in the table.

Listener	Listgp	age	vowel.quality	st.vowel
T1 : 24	T :432	Min. :17	a:408	a :204
T10 : 24	TA:264	1st Qu.:24	i:408	aa:204
T11 : 24	TQ:528	Median :29	u:408	i :204
T12 : 24		Mean :32		ii:204
T13 : 24		3rd Qu.:37		u :204
T14 : 24		Max. :62		uu:204
(Other):1080				
Stimulus	**match**	**Length**	**context**	**freq.**
?an : 51	match :584	long :612	emphatic :306	Min. : 0.00
?iirr : 51	mismatch:640	short:612	pharyngeal:306	1st Qu.: 0.72
?iyy : 51			plain :306	Median : 1.41
dhaarr : 51			q :306	Mean :13.50
Dhabb : 51				3rd Qu.:10.60
dhidd : 51				Max. :75.67
(Other):918				

Table 5-17: descriptive statistics for the study's variables

The data were then analyzed in R (Team, 2015) via logistic regression, in particular generalized linear mixed effects modelling (GLMM) since the DV is binary (thus generalized) and the explanatory variables include a mix of fixed and random effects (hence mixed effects). In the GLMM, the function glmer from the package lme4 (Bates et al., 2015) in R was used in order to build the different models.

As a first step in the analysis, the RQ was redefined in operational terms. **RQ1** is reproduced in 1 below.

(1) Whether speakers of Turkish would generalize the residual effects of emphatics/gutturals on neighbouring vowels to actual non-borrowed words and non-words.

The RQ in 1 is in principle the same as the one investigated in the PAT which is rendered below in 2.

(2) How close is the perception of the listeners to the observed mappings in the qualitative corpus?

- How do Turkish participants perceptually assimilate the Arabic long vs. short vowels to the Turkish short vowels?
- How do Turkish participants perceptually assimilate the Arabic plain vs. emphatic vowels?
- Does knowledge of Arabic (phonology) have any effect on perception?

Hence, four sets of hypotheses are generated below.

- **H_1:** Vowel **length** has an effect on the mapping/perception of ALT words into Turkish.
- **H_2:** Stimulus **context** (-/+ emphatic) has an effect on the mapping of ALT words into Turkish by different Turkish **listener groups**.
- **H_3: Knowledge of Arabic** has an effect on the perception of ALT words into Turkish by different Turkish **listener groups**. It would be predicted that the TA and TQ groups would have a closer degree of match compared to the corpus mappings (i.e., Ottomans).
- **H_4:** Word **frequency**[25] has an effect on the perception of ALT words into Turkish.
- **H_5: Vowel quality** has an effect on the perception of ALT words into Turkish.

The five hypotheses above statistically suppose that **Listgp**, **length**, **context**, **freq.** and **vowel.quality** have effects on match (**H1, H2, H3, H4 and H5≠0**) whereas the null hypothesis would state that none of them has any effect on match (**H0=0**).

5.3.1.2.1 SB-audio data (monosyllabic words only)

Objective

Based on the previous section, the objective of the current logistic regression is to test whether Listgp, context, length, stimulus frequency and vowel quality contribute to matching assimilation patterns in the audio only monosyllabic data set to those predicted in the corpus across listener groups.

[25] Frequency is taken as a variable of interest instead of word nature both of which are collinear with each other. Preference is given to frequency since it is a continuous variable with actual frequencies of real and nonsense words. On the other hand, word nature is a dummy variable of the two levels real (=0) vs. non-words (=1).

Three Turkish listener groups' assimilation patterns of long, short, plain and emphatic Arabic vowel categories embedded in real monosyllabic words either match or mismatch corpus mappings depending on a number of explanatory qualitative and quantitative variables. These include the following variables where the response variable is match and the random effects are listener and stimulus.

- Listener group; T, TA and TQ: **Listgp**
- Stimulus length; long or short: **length**
- Stimulus context; emphatic, plain, pharyngeal, q: **context**
- Stimulus frequency (a continuous variable): **freq.**
- Age of the participant at the time of the experiment: **age**
- Stimulus **vowel quality: a, i** and **u**

In R modelling, the basic information pertaining to the construct of the fixed effects is given in the form of a formula. For example, baseListgp (3) starts with one variable of interest, here Listgp, along the random effects structure, followed by the data name (msba, denoting monosyllabic audio set) and the binomial family since the DV is binary.

(3) baseListgp<- glmer(match~ Listgp +(Listgp|stimulus) + (1|listener) , data = msba, family = "binomial)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.6678	0.5543	1.205	**0.2282**
ListgpTA	-0.7844	0.4522	-1.734	0.0828 .
ListgpTQ	-0.2392	0.3804	-0.629	0.5295
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table 5-18: basic model output summary

The output summary in table 5-18 shows that Listgp TA is near significant (p-value=0.08), a result we have already seen in the raw data in 5.3.1.1.

<u>**Graphical data exploration**</u>

This section examines plots visualizing the research's explanatory variables along with the other control variables of **age** in order to determine whether they have any relationship with the dependent variable match. This task should help in understanding the data better and is expected to facilitate in building and selecting the glmer models in R.

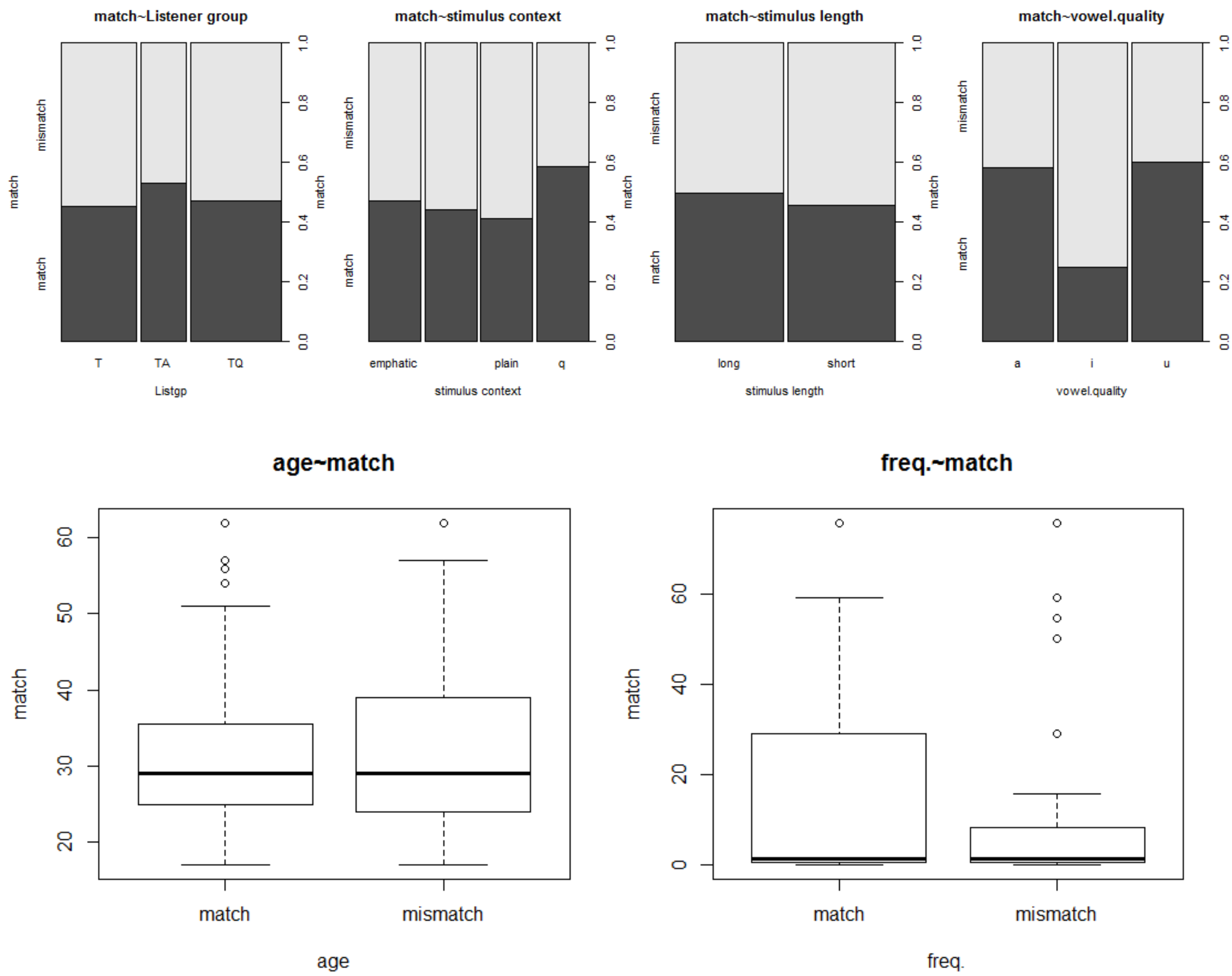


Figure 5-4: correlation between match and respectively Listgp, context, length & vowel quality

Figure 5-4 shows that the levels of each of the explanatory variables and the control variable reflect variation which, in turn, contributes to the probability of match. This sustains that these six variables are to be included in the optimal/maximal model when fitting the data.

The next step is to check if any interactions among these variables are of interest. Ten interactions were included in the maximal model including Listgp:context, Listgp:length, Listgp:freq., Listgp:vowel.quality, context:freq., context:vowel.quality, length:freq., length:vowel.quality, age:vowel.quality and freq.:vowel.quality, where (:) denotes an interaction. Some other interactions such as Listgp:age, length:age and age:freq. were deleted to simplify the maximal model especially that they involved a control variable and not a variable of interest.

In the last part of the analysis, the regression model that best fits/explains the SB audio data is given along with the model validation using the same methodology used in chapter 4 in section 4.2.3.

Step i: Determining the structures of the random and fixed effects

A. <u>Structure of the random effects</u>

The random structure was constructed first by determining whether the variables were within-unit or between-unit, where unit denoted either listener or stimulus and extrapolating which exploratory fixed effects to be included in the r models as per the research hypotheses.

1. Random effects variance

First, the null model (the model including only the intercept/constant and the random effect structure) was examined followed by plots of the two random effects. This was done so as to check the variance of both random effects and decide whether to include either or both effects.

(4) m0.null<- glmer(match~ 1 + (1|listener)+ (1|stimulus) , data = msba, family = "binomial", control = glmerControl(optimizer = "bobyqa"), nAGQ = 1)

Random effects:

Groups	Name	Variance	Std.Dev.
listener	(Intercept)	**0.5707**	0.7555
stimulus	(Intercept)	**3.7963**	1.9484
Number of obs: 1224, groups: listener, 51; stimulus, 24			

Table 5-19: Random effects variance summary

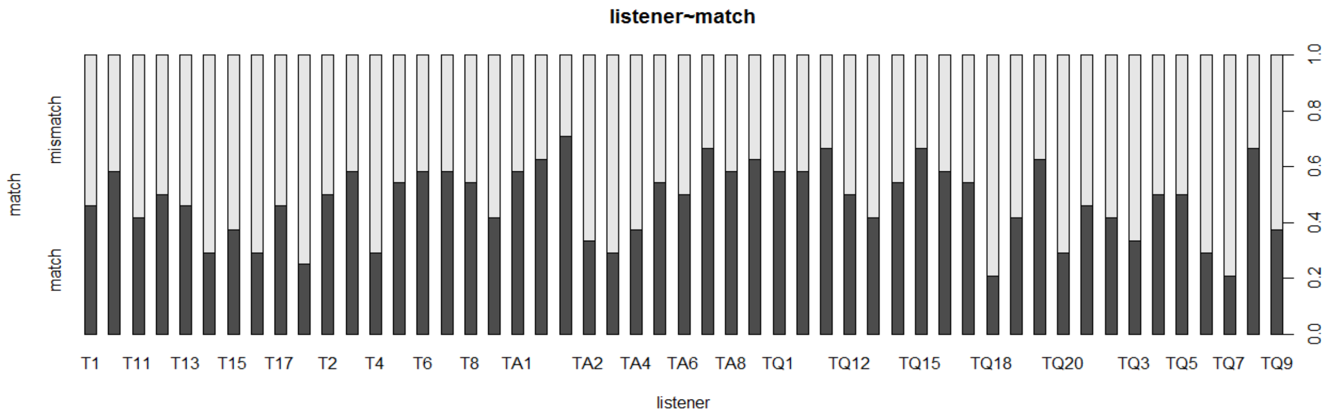


Figure 5-5: match~listener (null model)

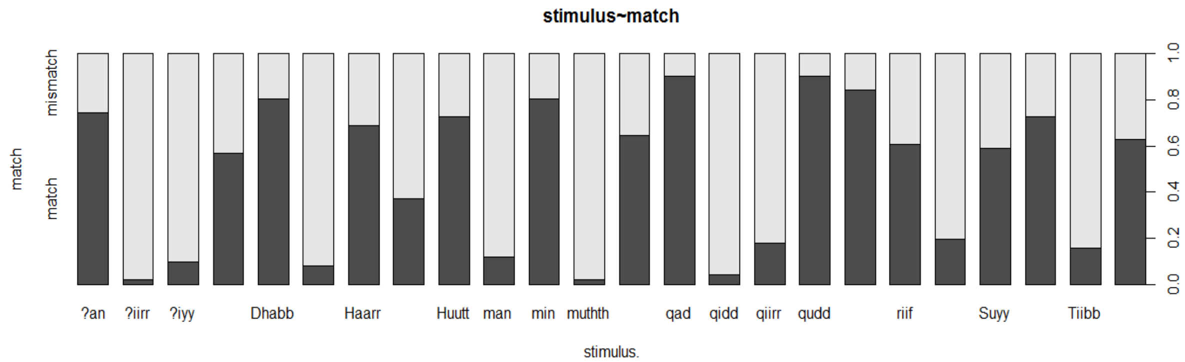


Figure 5-6: Random effects variance summary

The null model in 4 and table 5-19 along with figures 5-5 and 5-6[26] mirror the following.

1. Both variance values for listener and stimulus are >0 meaning they have to be part of the random effects structure of the final model. In addition, the variance in the model is attributed to both random effects.
2. The between-listener variance intercept in match is estimated as 0.5707, and the between-stimulus variance intercept is estimated as 3.7963. Thus, the total variance is 0.5707+ 3.7963 = 4.367. The variance partition coefficient, VPC (Steele, 2008b) for listener is 0.5707/4.367= 0.13, which indicates that 13% of the variance in match can be attributed to differences among listeners. On the other hand, the VPC for stimulus is 3.7963 /4.367= 0.869 indicating that almost 87% of the variance in match can be attributed to differences among stimulus tokens.

[26] In figure 5-5, only a representative sample of the listeners' population is given since the margins were too large to fit in the plot. The same scenario is found in the match~stimulus figure (5-6).

3. The variability in stimulus is larger than listener (0.5707 and 3.7963 respectively; also reflected in the bar plots), possibly because the number of stimulus units is less (24 items) compared to 51 listeners of 1224 observations.

B. Structure of the fixed effects. (maximal model: VOI +CVs)

Based on the hypotheses presented in 5.1.1., the current work's *variables of interest* ***VOI*** are Listgp, length, context, vowel quality and freq. Additionally, one control variable ***CV*** was included in the model, age. Furthermore, all interactions manifesting variation or related to the research hypotheses were included in the maximal model. Next, the random effect structure was revisited so as to determine which slopes were between-unit and which were within-unit.

Random slopes, random intercepts or both

According to Baayen (2008, p. 290), "in general, predictors tied to subjects (age, sex, handedness, education level, etc) may require by-item random slopes, and predictors related to items (frequency, length, number of neighbors, etc) may require by-subject random slopes." The classification of the present work's random effects is given below.

- ***Listgp*** is a ***between-listener*** (1|listener) variable since each listener belongs to a different listener group and a ***within-stimulus*** variable (1+Listgp|stimulus) because the same stimulus was presented to the three listener groups; stimulus does not vary across listener groups, at least in the audio task of the SB where all three groups receive the same stimulus.

- ***Length*** is a **within-listener** (1+length|listener) variable as each listener was presented with the same set of short and long vowels as stimulus. In other words, vowel length does not vary across the listeners. However, it is a **between-stimulus** (1|stimulus) variable since the stimulus can be either short or long but not both together; vowel length varies across stimulus.

- **Context** is also ***within-listener*** (1+context|listener) variable since all listeners get all the different contexts and the levels do not vary across the listeners. On the other hand, it is a ***between-stimulus*** (1|stimulus) variable as each context level has a different set of stimulus item; a context can be either emphatics, pharyngeals, plain or q but not two contexts or more at the same time.

Having determined the structures of both fixed and random effects, the maximal model presentation is in order as detailed in step ii.

#Step ii: Create the beyond optimal model (maximal model)

A. Include the VOIs+ CVs+ interactions of theoretical interest

In the maximal model (5) below, variables of interest, control variable and their interactions were incorporated as well as the random effects and the slopes in them. In addition, the intercept was included in the random effects structure.

(5) databasedmsba1<-glmer(match~Listgp + context+length+freq.+vowel.quality+ age+ Listgp:length + Listgp:freq. + Listgp:context + Listgp:vowel.quality + context:length+context:freq.+ context:vowel.quality+length:vowel.quality+ age:vowel.quality +freq.:vowel.quality + (Listgp|stimulus) + (context + length|listener) , data = msba1 , family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.5038843	0.8323136	-0.605	0.544912	
ListgpTA	0.1103299	0.9388776	0.118	0.906454	
ListgpTQ	0.2905367	0.8317390	0.349	0.726855	
contextpharyngeal	-2.1192874	0.8720994	-2.430	0.015095	*
contextplain	1.7333792	0.6972650	2.486	0.012920	*
contextq	-0.3689900	0.6947857	-0.531	0.595360	
lengthshort	-5.3857188	1.6539282	-3.256	0.001129	**
freq.	0.0911038	0.0346033	2.633	0.008468	**
vowel.qualityi	1.9252034	1.1781023	1.634	0.102226	
vowel.qualityu	-0.1753534	0.8712295	-0.201	0.840487	
age	-0.0004011	0.0149775	-0.027	0.978634	
ListgpTA:lengthshort	0.3020473	0.9820664	0.308	0.758415	
ListgpTQ:lengthshort	0.3789083	0.8867678	0.427	0.669167	
ListgpTA:freq.	-0.0040919	0.0164830	-0.248	0.803940	
ListgpTQ:freq.	-0.0212656	0.0163178	-1.303	0.192501	
ListgpTA:contextpharyngeal	-0.4673536	0.7538805	-0.620	0.535303	
ListgpTQ:contextpharyngeal	-0.4028556	0.7236783	-0.557	0.577748	
ListgpTA:contextplain	0.6074585	0.7913297	0.768	0.442699	
ListgpTQ:contextplain	1.3050237	0.7248495	1.800	0.071796	.
ListgpTA:contextq	0.1316402	0.7762492	0.170	0.865337	
ListgpTQ:contextq	-0.5072573	0.7592379	-0.668	0.504061	
ListgpTA:vowel.qualityi	-3.1809115	0.7612005	-4.179	2.93e-05	***
ListgpTQ:vowel.qualityi	-1.1275190	0.7314758	-1.541	0.123212	
ListgpTA:vowel.qualityu	-0.4860686	0.6492470	-0.749	0.454059	
ListgpTQ:vowel.qualityu	-0.6027517	0.6102960	-0.988	0.323330	
contextpharyngeal:lengthshort	-7.0386580	2.4992213	-2.816	0.004857	**
contextplain:lengthshort	5.9620366	1.3217566	4.511	6.46e-06	***
contextq:lengthshort	1.0655139	0.8912732	1.195	0.231893	
contextpharyngeal:freq.	0.7919074	0.2161283	3.664	0.000248	***

contextplain:freq.	-0.1422381	0.0343181	-4.145	3.40e-05	***
contextq:freq.	-0.0488685	0.0230343	-2.122	0.033875	*
contextpharyngeal:vowel.qualityi	4.6796594	1.5171519	3.085	0.002039	**
contextplain:vowel.qualityi	-5.1392891	1.0170320	-5.053	4.34e-07	***
contextq:vowel.qualityi	0.7816474	1.0452407	0.748	0.454571	
contextpharyngeal:vowel.qualityu	0.7541106	0.8065328	0.935	0.349787	
contextplain:vowel.qualityu	-3.4327476	0.8394297	-4.089	4.33e-05	***
contextq:vowel.qualityu	-2.0698468	0.8283821	-2.499	0.012466	*
lengthshort:vowel.qualityi	5.0648242	1.7496259	2.895	0.003794	**
lengthshort:vowel.qualityu	4.8843178	1.3682859	3.570	0.000357	***
vowel.qualityi:age	0.0733599	0.0238888	3.071	0.002134	**
vowel.qualityu:age	0.0244899	0.0174103	1.407	0.159536	
freq.:vowel.qualityi	-0.0701187	0.0431046	-1.627	0.103799	
freq.:vowel.qualityu	-0.5393535	0.1582398	-3.408	0.000653	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 5-20: maximal model output summary

Model interpretation is given in part C on fitting stepwise regression models below but first the model specification techniques that were adopted are in order.

B. Model simplification techniques

One note to mention before departing to the regression models is how the maximal model was simplified in order to deal with anticonservative and non-convergence issues (Barr et al., 2013).

1. Only variables of interest derived from the research hypotheses and their interactions were selected in addition to one control variable which reflected variation when interacting with the variables of interest.
2. Theoretical assumptions such as collinearity were taken in consideration. For instance, the variable ***st.vowel*** was removed from the maximal model since it is collinear with ***vowel.quality, length*** and ***context***.
3. The continuous variables ***age*** and ***freq.*** was scaled and centered[27].
4. The iteration number of the model was increased to 2e5, i.e., control=glmerControl(optCtrl=list(maxfun=2e5)) and nAGQ =1).

C. Fitting stepwise regression models

Two regression models were fitted using backward logistic regression. Specifically, one factor at a time was removed automatically using the dropterm and update commands in the MASS

[27] Centering a continuous variable entails selecting a number at which interpreting the intercept is meaningful. Often in R, users are prompted to scale and center continuous variables with warning messages.

package (Venables & Ripley, 2003) when it did not achieve 5% threshold of significance (p=<0.05). The repeated dropterm applications are given in **Appendix 5-2** and the final model is given in table 5-21.

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.4119516	0.7635410	-0.540	0.589523	
ListgpTA	0.2076415	0.6264908	0.331	0.740315	
ListgpTQ	-0.0374532	0.6251703	-0.060	0.952228	
contextpharyngeal	**-2.1892083**	**0.8867330**	**-2.469**	**0.013555**	*****
contextplain	**1.7069877**	**0.7094280**	**2.406**	**0.016122**	*****
contextq	-0.3974466	0.7111787	-0.559	0.576260	
lengthshort	**-5.0940069**	**1.5704405**	**-3.244**	**0.001180**	******
freq.	**0.0804284**	**0.0334369**	**2.405**	**0.016156**	*****
vowel.qualityi	**2.1100369**	**1.1950146**	**1.766**	**0.077446**	**.**
vowel.qualityu	-0.2792246	0.8721506	-0.320	0.748850	
age	-0.0002146	0.0149235	-0.014	0.988525	
ListgpTA:contextpharyngeal	-0.3810594	0.7561737	-0.504	0.614310	
ListgpTQ:contextpharyngeal	-0.2207218	0.7647029	-0.289	0.772859	
ListgpTA:contextplain	0.6280358	0.7855106	0.800	0.423986	
ListgpTQ:contextplain	**1.4300320**	**0.7489715**	**1.909**	**0.056220**	**.**
ListgpTA:contextq	0.1778608	0.7902864	0.225	0.821934	
ListgpTQ:contextq	-0.3956727	0.7931782	-0.499	0.617889	
ListgpTA:vowel.qualityi	**-3.3131208**	**0.7808000**	**-4.243**	**2.20e-05**	*******
ListgpTQ:vowel.qualityi	**-1.2639100**	**0.7592963**	**-1.665**	**0.095996**	**.**
ListgpTA:vowel.qualityu	-0.4334212	0.6104758	-0.710	0.477721	
ListgpTQ:vowel.qualityu	-0.3640794	0.6066042	-0.600	0.548378	
contextpharyngeal:lengthshort	**-7.2016469**	**2.4924847**	**-2.889**	**0.003860**	******
contextplain:lengthshort	**5.8285269**	**1.3116794**	**4.444**	**8.85e-06**	*******
contextq:lengthshort	0.9368842	0.9068071	1.033	0.301525	
contextpharyngeal:freq.	**0.7978377**	**0.2152008**	**3.707**	**0.000209**	*******
contextplain:freq.	**-0.1401666**	**0.0339762**	**-4.125**	**3.70e-05**	*******
contextq:freq.	**-0.0455771**	**0.0229981**	**-1.982**	**0.047504**	*****
contextpharyngeal:vowel.qualityi	**4.5839259**	**1.5110872**	**3.034**	**0.002417**	******
contextplain:vowel.qualityi	**-5.2603374**	**1.0244086**	**-5.135**	**2.82e-07**	*******
contextq:vowel.qualityi	0.7200049	1.0505651	0.685	0.493123	
contextpharyngeal:vowel.qualityu	0.7402418	0.8110165	0.913	0.361383	
contextplain:vowel.qualityu	**-3.4457779**	**0.8433414**	**-4.086**	**4.39e-05**	*******
contextq:vowel.qualityu	**-2.0579810**	**0.8279735**	**-2.486**	**0.012935**	*****
lengthshort:vowel.qualityi	**5.1509513**	**1.7328924**	**2.972**	**0.002954**	******
lengthshort:vowel.qualityu	**4.8654053**	**1.3551603**	**3.590**	**0.000330**	*******
vowel.qualityi:age	**0.0732599**	**0.0239829**	**3.055**	**0.002253**	******
vowel.qualityu:age	0.0243314	0.0174065	1.398	0.162164	
freq.:vowel.qualityi	**-0.0725039**	**0.0426335**	**-1.701**	**0.089012**	**.**
freq.:vowel.qualityu	**-0.5409178**	**0.1572894**	**-3.439**	**0.000584**	*******
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 5-21: step_2 model summary output

Table 5-18 shows that the four variables ***context***, ***length***, ***freq.*** and ***vowel quality*** have effects on match whereas ***Listgp*** and ***age*** do not. Furthermore, eight interactions were found significant or near significant, namely ***Listgp:context***, ***Listgp:vowel.quality***, ***context:length***, ***context:freq***., ***context:vowel.quality***, ***length:vowel.quality***, ***freq.:vowel.quality*** and ***age:vowel.quality***.

Figure 5-7 below represents a display of all the significant effects in step_2. As in chapter four, the dependent variable which is plotted on the y-axis is coded with the contrastive level 0 and 1 where 0=mismatch and 1=match. Match here is used when the response vowels yielded by the participants are the same as those observed in the corpus patterns. Moreover, the x-axis represents one of the two independent variables comprising an interaction whereas the other variable is represented with the lines in the middle of the two or three panes in each display.

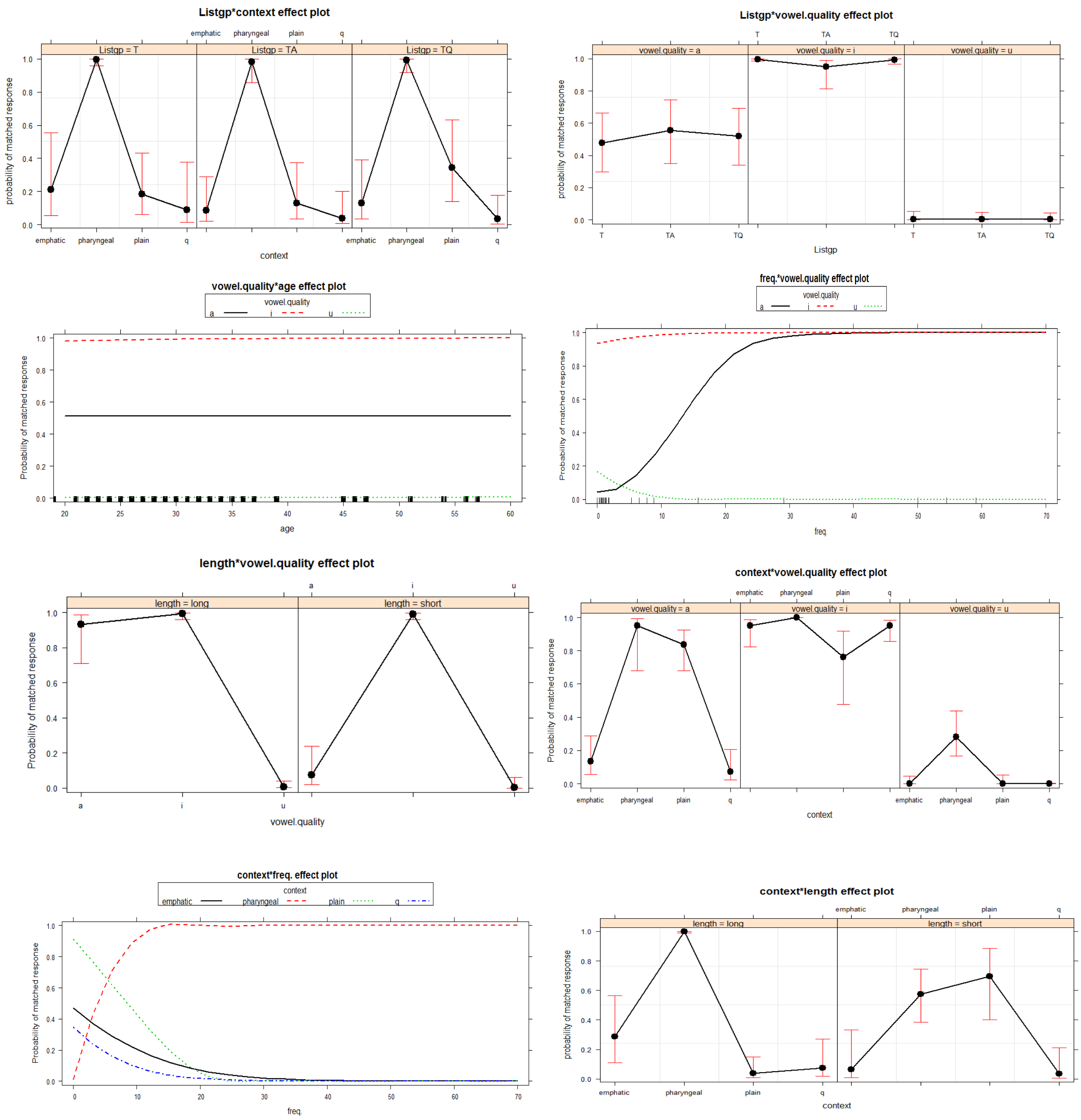


Figure 5-7:a display of all significant effects in step_2

In figure 5-8, we see that the three listener groups manifest similar patterns of match in the different contexts albeit with varying degrees. All three groups exhibited the highest degrees of match when the stimulus context had a pharyngeal consonant (?[28]an /ʕan/, Haarr and Huut) and lower degrees of match for emphatic (dhaarr and dhidd), plain (man, muthth and shaathth) and /q/ (qid and qiir), the last of which being matched the least. The listener groups, however, diverged in their assimilation of the emphatic and plain contexts. The T group displayed somewhat higher degrees of match for the emphatic context compared to the plain one in the words dhaar and Suyy compared to SHaathth and muthth in table 5-13, the TA group displayed somewhat higher degrees of match for the plain context in the words min and riif compared to dhidd and Tiib in table 5-14 and the TQ group manifested considerably higher degrees of match for the plain context in the words min and riif compared to the emphatic one in the words dhidd and Tiib (similar to the TA group but with varying degrees).

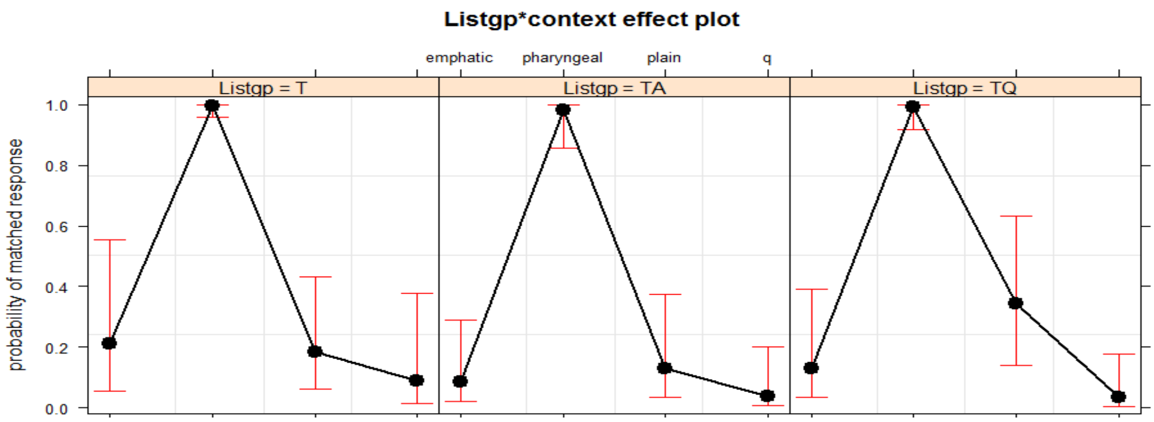


*Figure 5-8: Interaction between Listgp*context*

In figure 5-9, we can tease apart the effect of the interaction of Listgp and vowel quality in their effect on match. All groups reflected the highest degrees of match for the vowel /i/, the least for the vowel /u/ and 50-50 degrees of match for the vowel /a/. Moreover, the listener groups maintained the same degrees of match for the vowel/u/ reflected by the high matched tokens in green in tables 5-14, 5-15 and 5-16 in muthth, Suyy, qudd, Hubb, thuumm, Tuunn, quubb and Huutt; the TA group displayed the highest degrees of match followed by

[28] Transliteration was used here since the softward R does not allow use of IPA fonts.

the TQ and finally the T group for the vowel /a/ (in the words man, Dhabb, qad, ?an, dhaarr, qaarr and Harr in tables 5-14 to 5-16) and the T and TQ groups reflected higher degrees of match than the TA group for the vowel /i/ in the words min, dhidd, qidd, ?iyy, riif, Tiibb, qiir and ?iirr.

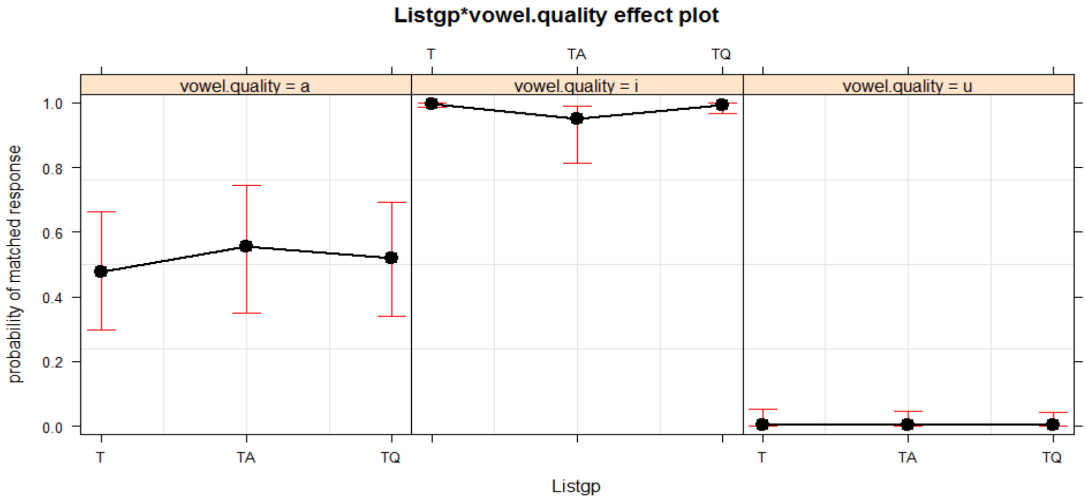


*Figure 5-9: Interaction between Listgp*vowel.quality (mismatch=0, match= 1)*

The same order of vowel quality regardless of Listgp is reflected in the interaction of vowel.quality with age in their effect on match in figure 5-10. We notice that regardless of age, listeners tended to display the highest degrees of match for the vowel /i/, the lowest for /u/ and 50-50 degrees of match (50%) for the vowel /a/. This can be interpreted such that all listeners, young and old, matched words with the vowel /i/ the highest which is shown as being significant in table 5-22.

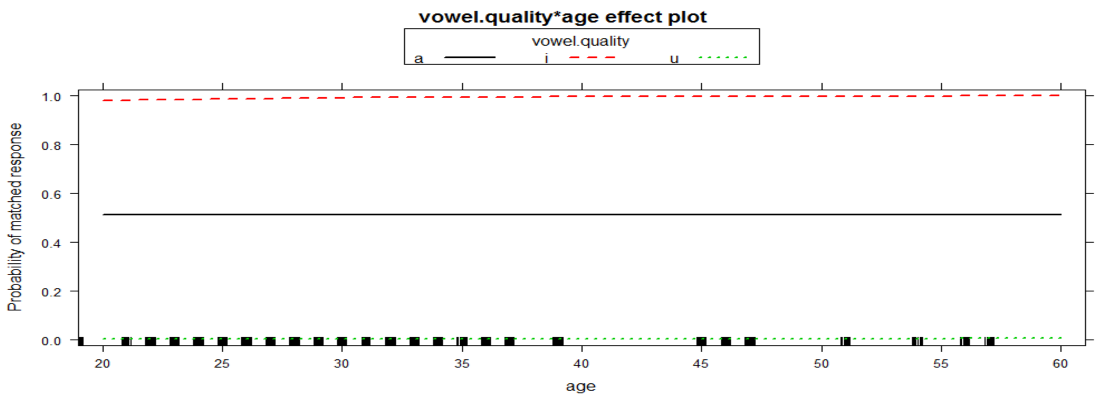


*Figure 5-10: Interaction of vowel.quality*age (mismatch=0, match= 1)*

Moreover, the same order of the three vowels (/i/, /a/ and /u/) is also reflected in the interaction of freq. and vowel quality when words are of high frequency (figure 5-11). That is, listeners tended to reflect higher degrees of match for the vowels /i/ (e.g. min with a frequency of 59,07/100,000 words and dhidd: 54,5, see table 5-7) and /a/ (man: 7,8 and Dhabb: 50,02) and lower degrees of match for the vowel /u/ in real words of high frequency (e.g. muthth: 0.64/100,000 words). However, with nonsense words or with words of less-frequency, listeners tended to manifest the highest degrees of match for the vowel /i/ (e.g. qiir: 1,49/100,000 words) still but the lowest for the vowel /a/ (e.g. Shaathth: 0.83) followed by /u/ (e.g. Suyy being a non-word). Thus, there seems to be a tendency for u words to be less frequent than i or a words.

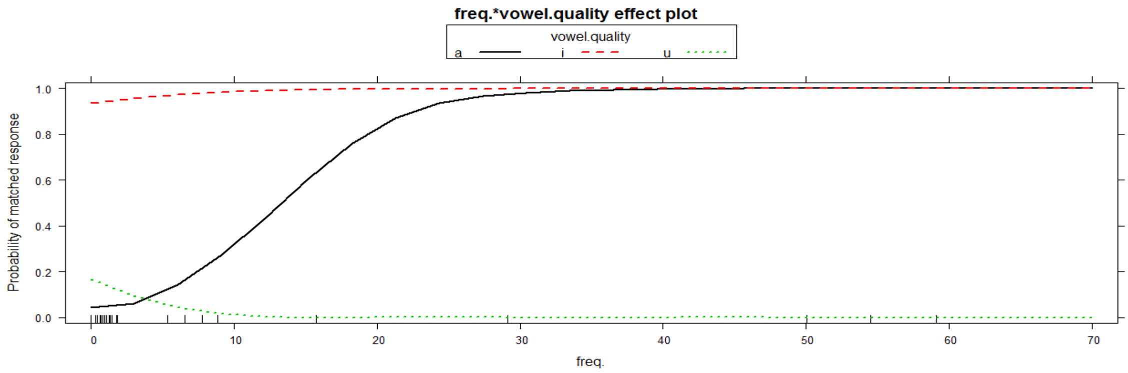


Figure 5-11: Interaction of freq. vowel.quality, (mismatch=0, match= 1)*

As for the interaction of length and vowel quality in figure 5-12, regardless of length, listeners' perception of the vowels /u/ and /i/ did not vary. That is, listeners reflected the highest degrees of match for the short and long /i/ alike and the least degrees of match for the short and long /u/ alike. Nevertheless, they displayed high degrees of match for the long vowel /a/ and low degrees of match for the short version.

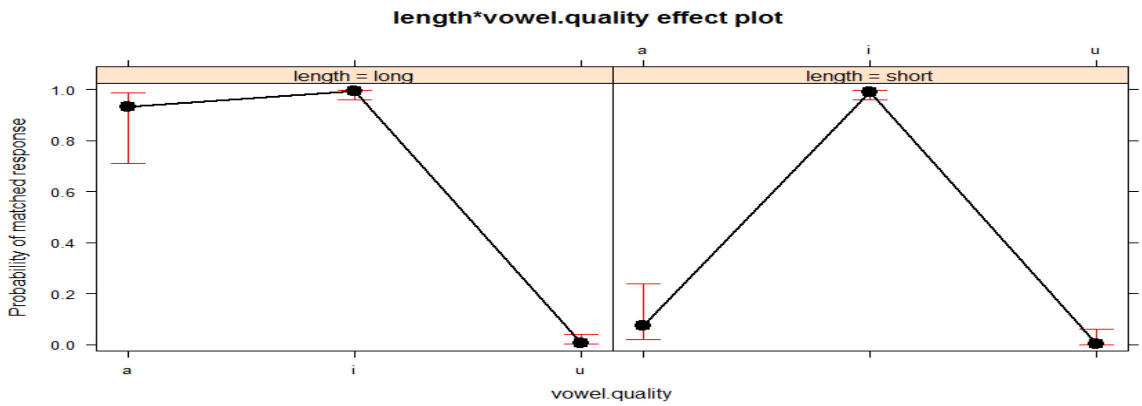


*Figure 5-12: Interaction of length*vowel.quality, (mismatch=0, match= 1)*

The last interaction involving vowel quality is that of context and vowel quality illustrated in figure 5-13 where listeners' perceptions of each context vary according to vowel quality. This interaction is related to the two interactions of Listgp*context in figure 5-8 and that of Listgp*vowel.quality in figure 8-9. That is, when the stimulus vowel was /a/, listeners perceived the pharyngeal and plain, in order, with higher degrees of match (higher than 0.5) and emphatic and q with lower degrees of match (less than 0.5). As for the vowel /i/, listeners assimilated the vowel with high degrees of match (higher than 0.5) in the order pharyngeal, emphatic and q (gutturals), and plain context. As for the vowel /u/, listeners reflected low degrees of match (lower than 0.5) for all four contexts, with pharyngeal being the highest and emphatic, plain and q all equally the lowest.

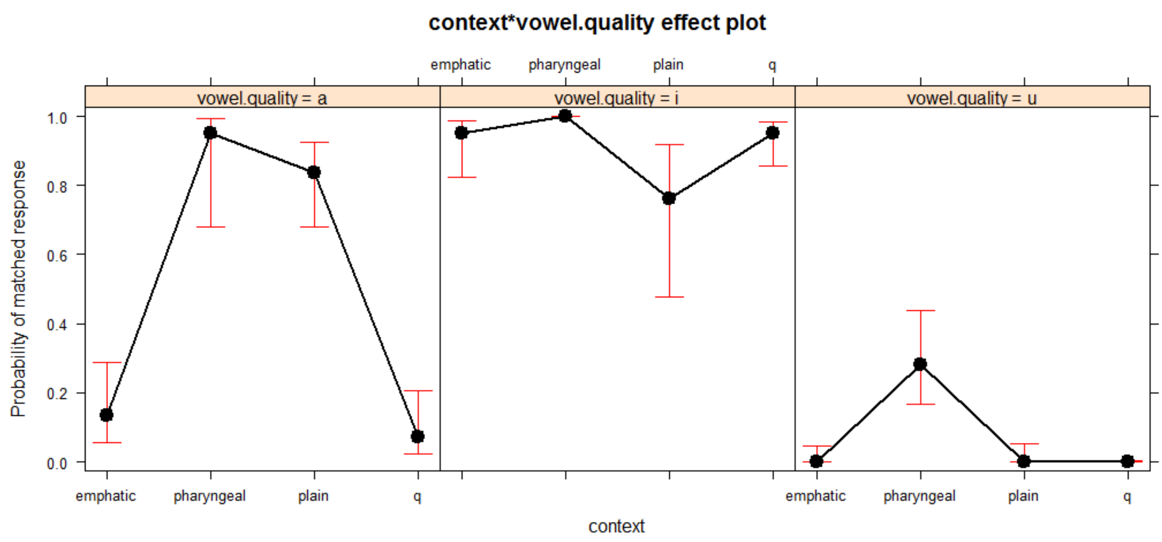


*Figure 5-13: Interaction of context*vowel.quality, (mismatch=0, match=1)*

As for the effect of the interaction of context and frequency, a clear pattern emerges as shown in figure 5-14. When the stimulus is a word of high frequency (high relative to other words exemplifying the same vowel phoneme), listeners tend to manifest the highest degrees of match (higher than 0.5) for the pharyngeal context (e.g. Hubb: 29,15 and ?an[29]: 15,78) but the lowest (lower than 0.5) for the q (e.g. qad: 75,67), emphatic (e.g. Tiib: 8,87 and dhabb: 50,02) and plain contexts (thuumm: 0.34 which is high relative to other u words). However, when the stimulus is a nonsense word or is of less frequency, listeners displayed the lowest degrees of match for pharyngeal (Huut: 0.), the highest for plain context (min: 59,07) and

[29] Here and throughout the thesis, the stimuli words were transliterated and non-IPA symbols were used because the software R does not read the IPA fonts.

displayed lower degrees of match (lower than 0.5) for the q (qaarr: 0.69) and emphatic contexts (Suyy: non-word) respectively.

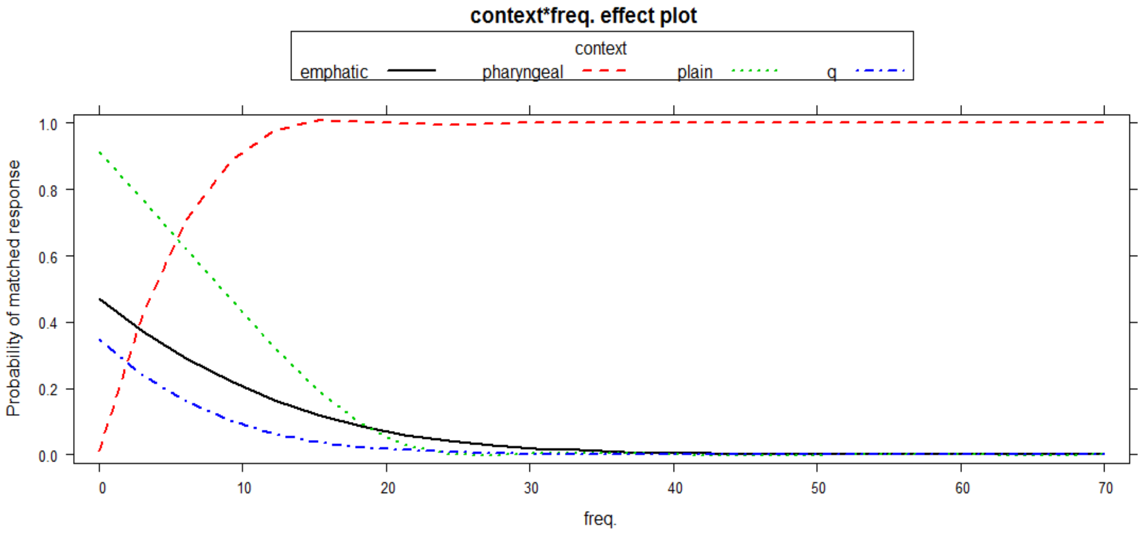


*Figure 5-14: Interaction of context*freq., (mismatch=0, match= 1)*

In the interaction of context with length in their effect on match in figure 5-15, when presented with long vowels, listeners exhibited the highest degrees of match in the pharyngeal context and the lowest in the plain one (e.g. Haar versus Shaathth). The match percentage the listeners yielded for the emphatic and q contexts (e.g., dhaarr, Tiibb, Tuunn versus qaarr, qiirr and quubb) when the vowels were long were below 50% indicating a low match rate, with q being matched lower than an emphatic. However, when the stimulus vowel was short, listeners tended to exhibit higher degrees of match for the plain context (e.g. man and min) followed by the pharyngeal one (higher than 0.5 on the y-axis) (e.g. ?an, ?iyy and Hubb) whereas they displayed lower degrees of match (below 0.5) for the emphatic (e.g. Dhabb, dhidd and Suyy) and q (e.g. qad, qidd and qudd), with q being the lowest.

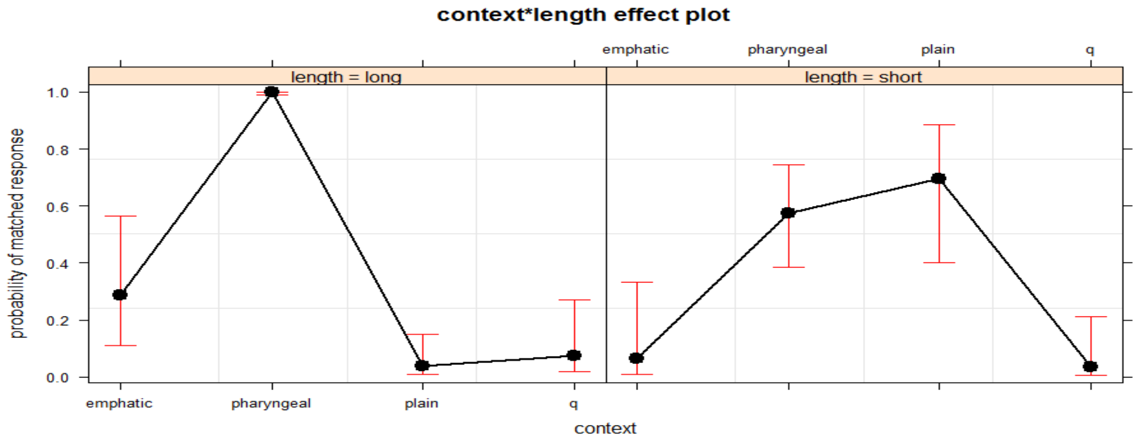


*Figure 5-15: Interaction of context*length*

Overall, these effects (length:vowel.quality, context:vowel.quality, context:length, Listgp:context, Listgp:vowel.quality, context:frequency and frequency: vowel. quality) reflect the mismatched mappings yielded differently by the three listener groups and reflected in figure 5-6. These mismatched mappings include ([a:]>/e/$_{\mathrm{T}}$, [u]>/u/$_{\mathrm{T}}$, [i]$_{\mathrm{d}ˤ}$>/i/$_{\mathrm{T}}$, [i]$_{\mathrm{q}}$>/i/$_{\mathrm{T}}$, [i]$_{ʕ}$>/e/$_{\mathrm{T}}$, [i:]$_{\mathrm{q}}$>/ɯ/$_{\mathrm{T}}$, [i:]$_{ʕ}$>/ɯ/$_{\mathrm{T}}$ and [u]$_{ʕ}$>/o/$_{\mathrm{T}}$).

The effect of age was found significant when interacting with vowel quality. The important take home message here is that regardless of age, all listeners young and old reflected the highest degrees of match for the vowel /i/, the lowest for the vowel /u/ and 50-50 for the vowel /a/.

<u>#Step iii: Model validation (goodness of fit)</u>

1. Plotting Residuals and interpreting them

After the best fitting model was selected, model validation was done. Since the response variable was categorical; residual plotting was chosen as the method of model validation.

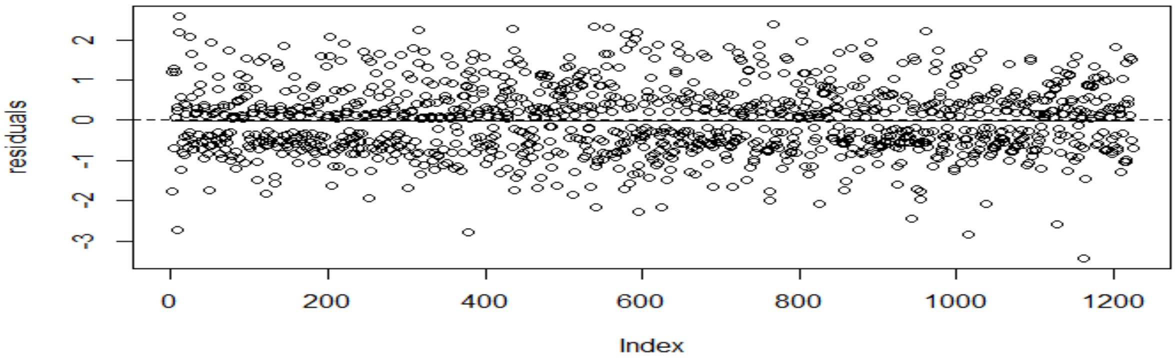


Figure 5-16: a scatter plot of residuals (testing for homoscedasticity)

In figure 5-16, residual points look homoscedastic since they are centering around the 0 line and the few non-centering points are not showing a pattern. In addition, the solid line overlays the dashed line, meaning that the model fit is good. Figure 5-17 below displays an almost normal distribution of maximalnoconstant model's residuals with few outliers at both ends.

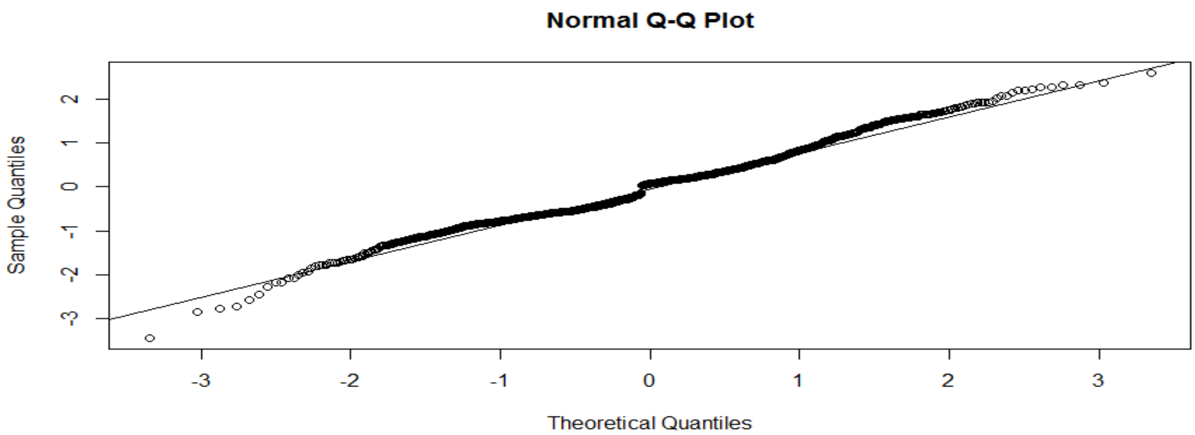


Figure 5-17: a Q-Q plot (testing for normality)

#Step iv: Reporting the results

A maximal generalized linear mixed effects modelling (GLMM) analysis was performed using R (Team, 2015) and *lme4* (Bates et al., 2015) to predict the relationship between **match** and the five independent variables of listener group ***Listgp***, consonants' ***context***, stimulus ***length***,

vowel quality, stimulus frequency ***freq.*** The fixed effects structure included all five exploratory variables, one control variable, namely ***age*** in addition to interactions reflecting variation. The random effects structure included both slopes and intercepts for listener and stimulus. The two theoretical assumptions of homoscedasticity and normality were adhered to and reflected in residuals plots and p-values with α-level of p>0.05 were used to evaluate variables

It was found that the probability of Turkish listeners matching assimilation patterns to ones predicted from the ALT corpus was dependent on the ***context*** and ***vowel quality***, the same result found in the PAT experiment, as well as ***frequency*** and ***length*** of the stimulus vowel. In addition, ***listener group*** was also found significant however when interacting with ***context*** and with ***vowel quality*** which was not witnessed in the PAT experiment. Moreover, the results also suggest that listeners' matched responses were also dependent on the interactions of ***context*** with the ***length*** of the stimulus vowel and its ***frequency,*** and of ***length*** and ***frequency*** with ***vowel quality***. The last common interaction between the PAT and SB-audio experiment is the interaction of ***age*** with ***vowel quality*** which was also found significant.

5.4 Discussion

This chapter examined the perception of three Turkish listener groups, T, TA and TQ, of real non-borrowed and nonsense Arabic monosyllabic words in order to see which group/s would have the closest perception of Arabic loanwords to that of Osmanlica speakers. This would, in turn, help in establishing which loanword adaptation model can be used to account for the corpus of Arabic loanwords in Turkish presented in chapter three.

The different assimilation patterns yielded by the three groups were provided in the form of tables in section 5.3.1. where it was found that listener group had an effect on the matched responses. In particular, the TA group displayed higher degrees of match by 53%, followed by the TQ group at 47% and the T group at 45%. This result shows that the TA group would be the closest to the Ottoman's perception of the ALT words followed by the TQ group, i.e., groups that know Arabic. The assimilation mappings of the three Turkish listener groups and those of the corpus (Ottomans) are given in figure 5-18 where the subscripts in the cognate vowels refer to the phonetic context and the subscripted “$_{T}$” in the mapped onto vowels refer to the recipient language, Turkish.

Figure 5-18 shows the mappings of each of the three listener groups. The TA mismatched only **five** vowel categories; /a:/>/e/$_{T}$ (predicted /a/$_{T}$), [i]$_{q}$>/i/$_{T}$ (predicted as /ɯ/$_{T}$), [i]$_{ʕ}$>/e/$_{T}$ (predicted as /i/$_{T}$), [i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted as /i/$_{T}$) and [u]>/u/$_{T}$ (predicted /y/$_{T}$). The TQ mismatched **six** vowel categories; /a:/>/e/$_{T}$ (predicted /a/$_{T}$), [i]$_{ʕ}$>/e/$_{T}$ (predicted /i/$_{T}$), [i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$), [i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted /i/$_{T}$), and [i:]$_{q}$>/ɯ/$_{T}$ (predicted /i/$_{T}$), and /u/>/u/$_{T}$ (predicted /y/$_{T}$). The T group mismatched the most (**eight** categories) /a:/>/e/$_{T}$ (predicted /a/$_{T}$), [i]$_{ʕ}$>/e/$_{T}$ (predicted /i/$_{T}$), [i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$), [i]$_{dˤ}$>/i/$_{T}$ (predicted /ɯ/$_{T}$), [i:]$_{q}$>/ɯ/$_{T}$ (predicted /i/$_{T}$), [i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted /i/$_{T}$), /u/>/u/$_{T}$ (predicted /y/$_{T}$) and [u]$_{ʕ}$>/o/ (predicted /u/$_{T}$).

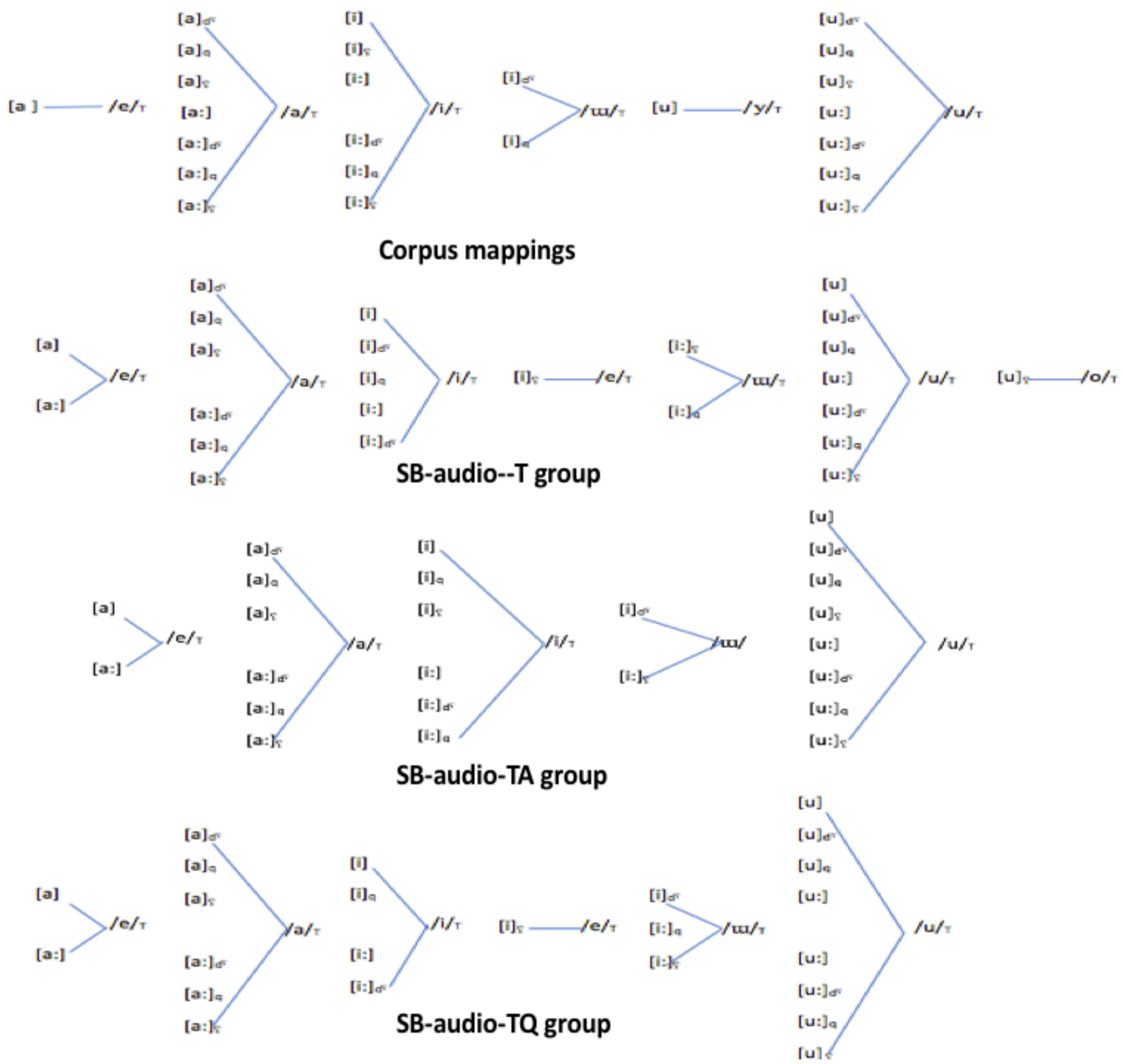


Figure 5-18: corpus and Turkish listener groups' mappings of Arabic loanwords into Turkish in the SB-audio task

The question that may be posed here is, what is the source of mismatch rendered by the listener groups? Is it only perception? We know from the raw data that the percentage of match in the audio task was 48% compared to 68% in the PAT, so other factors also play a role in the mappings. The listener groups were instructed to listen with their Turkish ears and borrow these words by writing them in Turkish spelling. However, they were not informed that the source of the stimulus words was Arabic; hence, an array of patterns emerged.

We have seen in chapter four in section 4.4., reproduced in figure 5-19, that perception accounted for almost 70% of the matched responses which was reflected in the distance between mapped categories in the vowel space of both Arabic and Turkish. The three categories mismtached in the PAT task were only $[a:]_{A}>/e/_{T}$, $[i^{\text{ʕ}}]_{A}>/e/_{T}$ and $/u/_{A}>/u/_{T}$.

PAT mappings

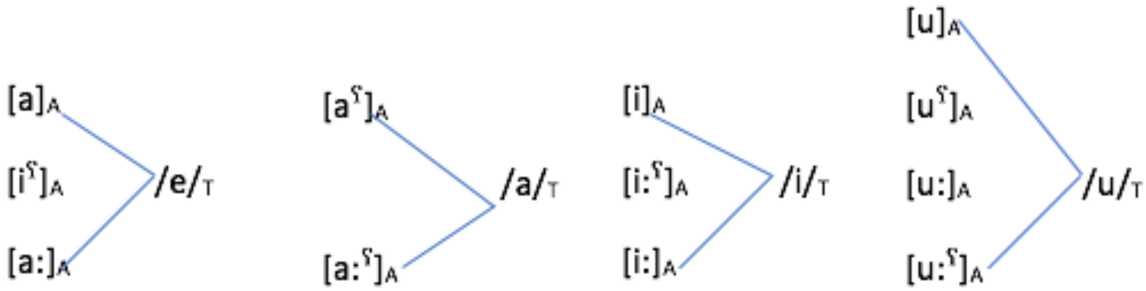


Figure 5-19: Turkish listener groups' perceptual mappings of Arabic loanwords into Turkish in the PAT experiment

In the same line, we revisit the vowel space of Arabic and Turkish but this time with real words in figure 5-20. In addition, the vowel categories mismatched by the three groups are presented in table 5-22. This information should help shed light on the source/s of mismatch in the perceived categorizations in the SB-audio only task.

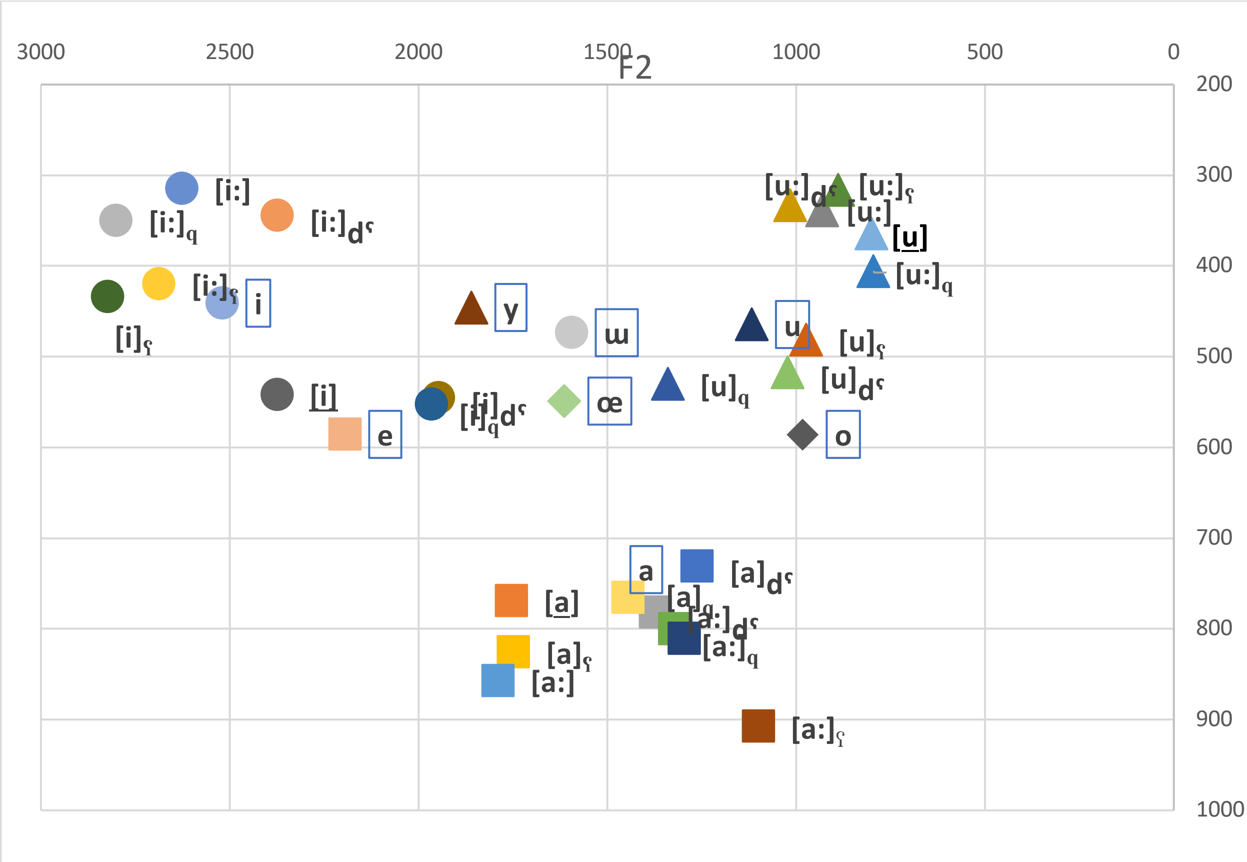


Figure 5-20: Mean frequency values of 2 Turkish speakers and 1 Arabic speaker's vowel formants plot (real words) in the SB-audio task. ***Vowels in boxes are Turkish vowels and the rest are Arabic ones; underlined /a/, /i/and /u/ are Arabic vowels*** *task.*

S.no	Mismatched category	Listgp
1.	[a:]>/e/$_{T}$ (predicted /a/$_{T}$)	T, TA, TQ
2.	[i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$)	T, TA, TQ
3.	[u]>/u/$_{T}$ (predicted /y/$_{T}$)	T, TA, TQ
4.	[i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted /i/$_{T}$)	T, TA, TQ
5.	[i]$_{ʕ}$>/e/$_{T}$ (predicted /i/$_{T}$)[30]	T and TQ
6.	[i:]$_{q}$>/ɯ/$_{T}$ (predicted /i/$_{T}$)	T and TQ
7.	[i]$_{dʕ}$>/i/$_{T}$ (predicted /ɯ/$_{T}$)	T
8.	[u]$_{ʕ}$>/o/$_{T}$ (predicted /u/$_{T}$)	T

Table 5-22: mismatched vowel categories in the SB-audio task

[30] In the mapping [i:]$_{ʕ}$>/a-ɯ/$_{T}$ (predicted /i/$_{T}$), the response vowels were considered as /ɯ/. Whenever the listeners responded with two vowels, the second vowel was only considered since long vowels are not allowed in Turkish.

In the first four patterns in table 5-22, all three groups converged in mismatching $[a:]_A$, $[i]_q$, [u] and $[i:]_ʕ$ respectively as $/e/_T$, $/i/_T$, $/u/_T$, $/ɯ/_T$ where $/a/_T$, $/ɯ/_T$, $/y/_T$ and $/i/_T$ were predicted. In figure 5-20, $[a:]_A$ is located somewhat centrally between $/a/_T$ and $/e/_T$; however, perceptually it is closer in F1 and F2 to $/a/_T$ than $/e/_T$. This means that the $[a:]_A>/e/_T$ is not phonetically supported. Nevertheless, the two vowels $[a:]_A$ and $/e/_T$ are both [-high] and [+front], meaning that this pattern is phonologically sustained. The $[i]_q$ is more front than $/ɯ/_T$ and closer to $/i/_T$; however, it is even closer to $/e/_T$ yet is mapped onto $/i/_T$ by the three groups. This indicates that the mapping in question ($[i]_q>/i/_T$) is not phonetically supported but rather phonologically determined. This is because $[i]_q$ and $/i/_T$ agree in the phonological features of height and frontness. In the third mapping in table 5-22, as shown in figure 5-20 $[u]_A$ is more back than $/y/_T$ and closer to $/u/_T$, meaning that this mapping is phonetically grounded and that the listeners depended on their Turkish ears (perception) here. In addition, the two vowels $[u]_A$ and $/u/_T$ are phonologically similar in the two features of height and backness whereas $[u]_T$ and $/y/_T$ are similar only in height. This means that the mapping $[u]_A>/u/_T$ is both phonetically and phonologically motivated. In the fourth pattern, listeners were confronted with the long vowel [i] neighbouring the voiced pharyngeal /ʕ/. All listener groups syllabified the monosyllabic word /ʕiːr/ into /a.ɯɾ̥/, shortened the long vowel and reflected the backing and lowering effect of the true guttural /ʕ/ by choosing the combination of the low $/a/_T$ in the first syllable and back $/ɯ/_T$ in the second syllable. This pattern reflects the effect of Turkish phonology as the listeners might have interpreted the backing effect of the guttural in a similar way to their Turkish vowel harmony of frontness-backness. As shown in figure 5-20, $/i/_T$ is phonetically closer in both F1 [height] and F2 [backness] to $[i:]_ʕ$ than $/ɯ/_T$ is. Moreover, $/i/_T$ and $[i:]_ʕ$ are phonologically similar in the two features of height and frontness whereas $/ɯ/_T$ and $[i:]_ʕ$ are similar only in height. Thus, we can conclude that the mapping $[i:]_ʕ>/ɯ/_T$ neither is phonetically nor phonologically grounded.

The T and TQ groups mismatched two patterns which the TA matched to the corpus mappings. These are $[i]_ʕ>/e/_T$ (predicted $/i/_T$) and $[i:]_q>/ɯ\text{-}ɯ/_T$ (predicted $/i/_T$). Regarding the mapping $[i]_ʕ>/e/_T$, $/i/_T$ is closer to $[i]_ʕ$ than $/e/_T$ is in terms of both F1 [height] and F2 [backness]. Furthermore, [i]$_{ʕ}$ and /i/$_T$ are similar in the two features of height and frontness whereas [i]$_{ʕ}$ and /e/$_T$ agree only in the phonological feature of frontness but not height. Hence, the mapping [i]$_{ʕ}$>/e/$_T$ neither is phonetically nor phonologically (phonology of Arabic) grounded. Similar to the treatment of [iː]$_{ʕ}$, we can assume that the source of this perceptual map is Turkish phonology.

In the categorization of the vowel [iː]$_q$, the two groups T and TQ resorted to both their native language perception and phonology as they syllabified the word *qiid* /qiːˤd/ into /qɯ.ɯd/ which implies shortening the vowel, and chose /ɯ/$_T$ (backer than /i/$_T$) in response to the emphasis spread from the /q/. Moreover, they rendered an identical vowel /ɯ/$_T$ in both syllables in the word /qɯ.ɯd/. As for the TA group, they matched the vowel [iː]$_q$ to the predicted category /i/$_T$. In figure 5-20, we can see that [iː]$_q$ is phonetically closer to /i/$_T$ than /ɯ/$_T$ is both in F1 and F2. Moreover, [iː]$_q$ and /i/$_T$ are phonologically similar in both features of height and frontness. Therefore, this suggests that the T and TQ groups might have translated the lowering and backing effect of /q/ as a backing effect similar to their Turkish vowel harmony's front-back distinction. On the other hand, the TA group utilized their access to both Arabic and Turkish and rendered the predicted category /i/$_T$.

The mapping of [i]$_{d^ʕ}$ onto /i/$_T$ was only mismatched by the T group where /ɯ/$_T$ was predicted. The two other groups with knowledge of Arabic matched it to its predicted category. In figure 5-20, we can see that [i]$_{d^ʕ}$ is in closer proximity in F1 and F2 to /e/$_T$ and /ɯ/$_T$ than /i/$_T$. This means that the perceptual map [i]$_{d^ʕ}$ >/i/$_T$ by the T group is not phonetically grounded. However, it is phonologically supported since [i]$_{d^ʕ}$ and /i/$_T$ share the same phonological features of height and frontness.

In the mapping of the vowel [u]$_{ʕ}$, the T group categorized it as /o/$_T$, a vowel not found in MSA/Classical Arabic whereas the TA and TQ responded with the predicted /u/$_T$. This shows that the TA and TQ used their knowledge of Arabic here whereas the T group depended on their native perception since the /o/$_T$ is lower in the vowel space reflecting the lowering effect of the /ʕ/. In figure 5-13, this was reflected in the interaction of context and vowel quality where listeners displayed higher degrees of match for the vowel /u/ in the pharyngeal context. We presume from the mismatched mapping of $[u]_{ʕ}>$ $/o/_{T}$ that this was the case because the TA and TQ matched the vowel $[u]_{ʕ}$ to the predicted category $/u/_{T}$. This pattern confirms that the TA and TQ groups are closer to the Ottomans' perception than the T group of mapping the $[u]_{ʕ}$ category.

5.5 Summary

To sum up, it was demonstrated through the analysis of real and nonsense monosyllabic words in the audio-only condition of the simulated borrowing experiment that the perceptual maps rendered in the SB-audio only condition can be explained by a mix of phonetics, phonology of Turkish and phonology of Arabic sustaining a hybrid model of both phonetics and phonology. In addition, it was also found that listener group, with the T, TA and TQ groups, had an effect on the matching of vowels neighbouring gutturals. Nevertheless, this effect included both the TA and TQ groups, the two groups with Arabic knowledge, and was linked with the interaction of ***listener group*** to the two variables of ***context*** and ***vowel.quality***. In addition, the four variables ***length***, ***context, frequency*** *and* ***vowel.quality*** were also found to play a role in the matched responses along with interactions ***context:length***, ***context:freq.***, ***context:vowel.quality***, ***length:vowel.quality***, ***freq.:vowel.quality*** and ***age:vowel.quality***. Table 5-23 presents a summary of the significant variables and interactions in the PAT experiment and SB-audio-only task.

S.no	*Variable/interaction*	PAT	SB-audio
1.	**Context**	✓	✓
2.	**Freq.**	NA	✓
3.	**Length**		✓
4.	**Vowel.quality**	✓	✓
5.	**Listgp:context**		✓
6.	**Listgp:vowel.quality**		✓
7.	**Context:length**		✓
8.	**Context:freq.**		✓
9.	**Context:vowel.quality**	✓	✓
10.	**Length:freq.**	NA	✓
11.	**Length:vowel.quality**	✓	✓
12.	**Vowel.quality:age**	✓	✓
13.	**Vowel.quality:freq.**	NA	✓

Table 5-23: A summary table of the significant variables and interactions in the PAT and SB-audio tasks, NA= not applicable

The effect of listener group in both the PAT and SB experiments could be interpreted as follows. In the PAT experiment with only nonsense words, listener group did not play any role in the mapping of ALT words to their predicted categories. Conversely, when real and nonsense words were introduced to the three listener groups in the SB experiment, the listener group role emerged since the TA in addition to the TQ exhibited closer degrees of match to the corpus facts. This also gives weight to the nature (real vs. nonsense) and ***frequency*** of stimulus words.

Similar to the PAT results, ***context*** and the ***vowel quality*** prominently contributed to the matched responses along with the two interactions ***length:vowel quality*** and ***age:vowel.quality***. However, some other interactions were found significant in the SB-audio condition but not the PAT such as ***Listgp:context***, ***Listgp:vowel.quality***, ***context:length***, ***context:freq.***, ***context:vowel.quality*** and ***length:vowel.quality***.

Thus far, the TA and TQ groups could be said to be closer in the SB-audio task in their perception of the ALT predicted facts to the Ottomans grammar than the T group listeners. Having said this, the perception of all three groups was high across the three groups implying that some pieces of the puzzle are still not in place.

Hence, in the next chapter, the analysis of the simulated borrowing experiment is resumed with the stimulus presentation conditions audio-only, audio-written and written-only, and with monosyllabic only with the aim of checking whether knowledge of Arabic writing system and phonology have any effect on the matching of vowels neighbouring gutturals.

6 Simulated Borrowing: audio, audio-written and written data

6.1 Introduction

In chapter 5, only audio stimuli were analyzed and the findings showed that variables related to the TA and TQ groups, i.e., groups with knowledge of Arabic, had effects on the response variable match. In this chapter, audio-written and written data are also analyzed in addition to the audio stimuli with a mix of real and nonsense monosyllabic words to measure whether knowledge of the spelling system of Arabic (orthography) would affect the degree of match of Arabic source words to those predicted based on the qualitative corpus as given in RQ2 in 1 below.

(1) RQ2: Whether knowledge of Arabic orthography and/or phonology play a role in determining the quality of vowels neighbouring gutturals.

In RQ2, two alternative outcomes are possible. The first is that knowledge of Arabic grammar and spelling does not play any role in determining neighbouring vowels' quality as represented by the null hypothesis. If this were true, either the perception only or phonology only models could be resorted to to explain the residual effects of gutturals under study. Alternatively, the second prediction states that either the TA and/or TQ or both groups would yield similar patterns to the Ottomans' as represented by the dependent variable 'match'. If it is the TA versus TQ, then this should reveal further information on the knowledge of Arabic that the Ottomans had; i.e., whether it was written, for religious purposes and/or spoken. Furthermore, if the second prediction proves correct, then the hybrid model would be the one to account for the corpus patterns. It is expected that the second prediction is most likely to prove correct since from the results of the SB audio experiment, it was found that the groups with knowledge of Arabic had effects on the variable match compared to the monolingual Turkish group. This implies that there must be another factor in addition to the perceptual one that the TA and TQ groups relied on, which is knowledge of Arabic phonology or grammar in general of which knowledge of Arabic spelling is a fine-grained component.

This chapter is organized as follows. Section 6.1 opens the chapter with the main research question investigated. Section 6.2 highlights the research methodology including the stimuli, participants and procedures adopted. Section 6.3 presents preliminary raw data analysis based on observations on audio, audio-written and written stimuli. Section 6.4 reports on the analysis of audio and audio-written data using logistic regression in R while section 6.5 offers a logistic regression analysis of the written data. Section 6.7 closes with a summary of the chapter.

6.2 Research methodology

The same methodology presented in chapter 5, section 5.3, was adopted in chapter six. However, it is important to note that in this chapter all three stimuli types (audio, audio-written and written) are examined where first the raw data of the audio and audio-written conditions are compared to each other. Then the raw data of the written condition are explored in the data analysis section in 6.3.

The recording was done by the same native Syrian Arabic speaker in chapter 5 (5.3.1) and the participants were the same three groups mentioned in chapters 4 and 5 (5.3.2). These included Turkish monolingual speakers (T), bilingual speakers (TA) and Turkish speakers with Arabic knowledge through Quranic recitation (TQ).

The stimuli involved recording 72 words with plain and emphatic (guttural) vocalic variants (short and long vowels); 6 vowels (long+short) X 3 tokens (1 per condition) X 4 consonant types (emphatic, plain, pharyngeal and q). Stimuli presentation was of three types: audio only (A) to the T, TA and TQ groups; audio-written to all three groups, and written presented to the TA and TQ groups only since the T group did not have access to Arabic.

Of note is that participants' engagement in the Simulated Borrowing tasks was tested through posing some mandatory arbitrary questions such as '*give a number less than 5*' and '*give a number bigger than 6*', inter alia. Hence, no distractors, no repetition or randomization of questions were used.

A summary table of the research methodology is provided in table 6-1 which consists of the stimuli type, participants, procedure and rationale including hypotheses and predictions.

Stimuli	***Who does it?***	***What do they do?***	***Why?***
Audio (only) = recordings of Arabic words (real and nonsense), produced by a native speaker of Arabic =Real monosyllabic words= 23 Nonsense monosyllabic words=1	T: Monolingual Turkish speakers TA: Turkish-Arabic bilingual speakers of Arabic dialect in Turkey) TQ: Turkish speakers with knowledge of Arabic from recitation of Qur'an	write down in Turkish the word you hear	-This is a contextualized version of the PAT. This is a replication of the PAT but using real and nonsense monosyllabic words and with a different task (write it down). In Turkish, 'write it down' is the same as 'choose which vowel' since the spelling system of Turkish is phonetic, for vowels at least. -**Hypothesis:** patterns of assimilation will match the mappings among vowels observed in the corpus.
Audio + written = recordings of Arabic words (real and nonsense), produced by a native speaker of Arabic, accompanied by the word on screen in Arabic script =Real monosyllabic words= 23 Nonsense monosyllabic words=1	T: Monolingual Turkish speakers TA: Turkish-Arabic bilingual speakers of Arabic dialect in Turkey) TQ: Turkish speakers with knowledge of Arabic from recitation of Qur'an	write down in Turkish the word you hear and see	-**Hypothesis:** knowledge of Arabic spelling will have an effect on the degree of match. -TA and TQ groups are expected to exhibit closer matching patterns to the predicted patterns based on their knowledge of both Arabic phonology and orthography than the T group. This task also tests what people rely on more when they have two conditions (audio & audio-written stimuli), i.e., perception or phonology. - **Prediction:** If the hybrid model is correct, then knowledge of Arabic orthography will result in a closer match to the corpus mappings; whereas, if the perception only model is correct, knowledge of Arabic will make no difference or result in a lower degree of match.
Written (only) = Arabic words (real and nonsense), produced on screen in Arabic script Real monosyllabic words= 23 Nonsense monosyllabic words=1	TA: Turkish-Arabic bilingual speakers of Arabic dialect in Turkey) TQ: Turkish speakers with knowledge of Arabic from recitation of Qur'an	write down in Turkish the word you see	-This is to tease apart the effects of orthography vs. perception as it tests what participants do when they have only one condition to rely on (written stimuli). - It also measures whether TA and TQ would have similar patterns to the Ottomans based on their knowledge of Arabic spelling.

Table 6-1: Summary of the Simulated Borrowing methodology

The methodology used in the logistic regression parts of this chapter (6.4 and 6.5) follows the same procedure used in chapters 4 and 5. That is, first the sample size, objectives and hypotheses are stated and summarized. Then a protocol of model selection is followed (Zuur et al (2009), Winter (2013), Baayen (2008) and Barr et al (2013)). The protocol involves the three steps of i) building the fixed and random effects structure, ii) constructing the maximal model and running logistic regression models until stopping at the final model where no further variables or interactions are significant any more, and iii) validating the results and reporting them.

In the next section, the raw data of the experiment's three conditions of audio, audio-written and written stimuli are inspected and generalizations are drawn on the effect of listener group on match.

6.3 Raw data analysis

In this part, the results of the raw data for the three SB conditions of audio, audio-written and written stimuli are reviewed and compared. Pivot tables were created in excel 365 and match percentage plots were made in R software (Team, 2015). Confusion matrix tables of groups' mappings are given as tables 6-2 through 6-7 and 6-19 and 6-20. These portray actual counts of participants' tokens in the three conditions of the SB experiment. The green shade indicates a match where the response vowel is the same as the predicted vowel based on the research corpus. On the other hand, the red shade designates a mismatch where the response vowel is not the same as the predicted vowel.

On each of the eight tables, the first column displays the stimulus vowels, long and short, in their plain and guttural environments next to the stimulus words exemplifying these vowels. The second row to the right of the first two columns features the response vowels as mapped by the participants. These vowels include the eight Turkish vowels in addition to long vowels, syllabified vowels (2 or even 3 identical or different vocalic categories) and diphthongs. The reason for the variety of the responses is due to the fact that the SB experiment was an open choice task. Thus, despite the fact that all the stimulus words were all monosyllabic, the participants sometimes yielded either monosyllabic, disyllabic or polysyllabic responses. For instance, for the input word *SHaathth [aa]*, 6 listeners produced words with the vowel [a] such as şarz, şar, şaf and şaz; 1 with a disyllabic word, namely *şahıs* [a-*ı*]; 9 with the vowel [e] such as şez, şed, şerr, şef, şel and şew and 2 with the vowel [ee] such as şeer, şeev, şees and şeef.

Tables 6-2 to 6-4 represent the mappings of the T, TA and TQ groups in the SB-audio only (A) condition reproduced from chapter 5. Tables 6-5 to 6-7 depict the mappings of the three groups in the SB-audio-written (AW) condition followed by generalizations on the vowel categorizations and percentage of match across and within listener groups in both the audio and audio-written conditions. These are followed by observations on the raw data and generalizations on the vowel mappings and percentage of match across and within the two listener groups.

Listgp		T																												
SV/RV		a	aa	a-a	a-e	ai	a-ı	e	ee	e-i	i	ı	ii	ı-ı	i-i	i-i-i	o	ö	oa	o-a	oo	u	ü	u-a	ui	u-i	uu	u-u	ü-ü	Total
a	man	2						15			1																			18
	Dhabb	13															5													18
	qad	16															1	1												18
	?an	13		1	1			2	1																					18
a:	SHaathth	6					1	9	2																					18
	dhaarr	8	3																2	4				1						18
	qaarr	12	6																											18
	Haarr	15	2				1																							18
i	min							3			15																			18
	dhidd							2			13	2									1									18
	qidd							3			14	0										1								18
	?iyy	1				3	1	13			0																			18
i:	riif										8		7		2	1														18
	Tiibb						2			2	2			2	1							1	1		1	6				18
	qiirr						3				0			11	3											1				18
	?iirr						16			1	1																			18
u	muthth										1						5	1				11	0							18
	Suyy																3					15								18
	qudd																2				1	15								18
	Hubb	1						1									11	3				2								18
u:	thuumm																1					15					1	1		18
	Tuunn																					10		1			4	2	1	18
	quubb																					15					3			18
	Huutt																				1	12			1		4			18
	Total	87	11	1	1	3	24	48	3	3	55	2	7	13	6	1	28	5	2	4	3	97	1	2	2	7	12	3	1	432

Table 6-2: Table 6-2: Summary confusion matrix of T group mappings in the SB_Audio experiment (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; RV= response vowel; Listgp= listener group

T= monolingual listeners

Listgp		TA																					
SV/RV		a	aa	a-ɪ	a-u	e	ee	e-i	i	ɪ	ɪe	ii	ɪ-ɪ	i-i	ɪ-v	o	ö	oa	ou	u	ü	ui	uu	Total
a	man	2				9																		11
	Dhabb	9				2																		11
	qad	9							1								1							11
	?an	6				4			1															11
a:	SHaathth	3				5	3																	11
	dhaarr	7	2													1		1						11
	qaarr	7	4																					11
	Haarr	7	4																					11
i	min								10		1													11
	dhidd					2			3	6														11
	qidd	1				1			8	1														11
	?iyy	1				4			6															11
i:	riif								10				1											11
	Tiibb			1					5	3		1	1											11
	qiirr								4	2		1	2	2										11
	?iirr	1		4				2	1	2			1											11
u	muthth															2			2	7	0			11
	Suyy			1					2			1			1					6				11
	qudd								1											10				11
	Hubb								2							2				6		1		11
u:	thuumm																			7		1	3	11
	Tuunn																			7			4	11
	quubb																			10			1	11
	Huutt				1															8		1	1	11
	Total	**53**	**10**	**6**	**1**	**27**	**3**	**2**	**54**	**14**	**1**	**3**	**5**	**2**	**1**	**5**	**1**	**1**	**2**	**61**		**3**	**9**	**264**

Table 6-3: Summary confusion matrix of TA group mappings in the SB-Audio experiment (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; RV= response vowel; Listgp= listener group

TA= bilingual listeners

Listgp		TQ																																							
SV		a	aa	a-a	ai	a-ɪ	a-u	e	ea	ee	e-e	e-i	eɪ	e-ɪ	i	ɪ	ɪe	ii	ɪ-ɪ	i-i	ɪɪ	ɪ-ɪ	iu	i-ü	o	ö	oa	o-a	o-i	oö	u	ü	ua	ui	u-i	uɪ	uu	u-u	uü	w	Total
a	man	1					1	19							1																										22
	Dhabb	19																							3																22
	qad	21																													1										22
	?an	18	1	1				1								1																									22
a:	SHaathth	1	1					15		5																															22
	dhaarr	14	2																								2	3					1								22
	qaarr	14	7	1																																					22
	Haarr	13	8					1																																	22
i	min							3							18									1																	22
	dhidd							1							9	11															1										22
	qidd	1						2							17	2																									22
	?iyy	2			3	1		10	1				2		1	2																									22
i:	riif														14	1		7																							22
	Tiibb					1									7	7		1		1	1								1					2	1						22
	qiirr														6	6	1		5	2	1	1																			22
	?iirr	1				14		1			1	3		1	0		1																								22
u	muthth																								6						15	1									22
	Suyy														5	1									4						9				1	1				1	22
	qudd																														21						1				22
	Hubb							2	1						1										5	1				1	11										22
u:	thuumm																						1					1			15						4	1			22
	Tuunn														1																15						5		1		22
	quubb																														18						4				22
	Huutt																														17	1					4				22
	Total	105	19	2	3	16	1	55	2	5	1	3	2	1	80	31	2	8	5	3	2	1	1	1	18	1	2	4	1	1	123	2	1	2	2	1	18	1	1	1	528

Table 6-4: Summary confusion matrix of TQ group mappings in the SB-Audio experiment (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; Listgp= listener group*, *TQ= Turkish listeners with Quranic recitation knowledge

Listgp		T																																						
SV		a	aa	a-a	ae	a-e	a-i	ao	a-o	e	ea	ee	ı	i	ii	o	oo	u	uu	zr	i-i	ai	ı-ı	aı	ö	ui	ü	üü	uv	au	ü-ü	ue	a-ı	ı-ı	aı	ıı	u-ı	ı -i	e-ı	Total
a	faththa	2								14			2																											18
	Tall	16														2																								18
	qatt	18																																						18
	?al	13	1	1	1	1				1																														18
a:	naabb	7	5							2	3	1																												18
	Saaww	11	2					3	2																															18
	qaatt	10	8																																					18
	Haaff	10	1		1					5		1																												18
i	zirr												13	4						1																				18
	Sirr									1			15												1											1				18
	qibb	4								2			12																											18
	?izz	1			1					8			2	2	1							1		2																18
i:	tiinn													7	8			1			1													1						18
	Tiill													4	2						6		1			1								1		1	1		1	18
	qiibb												4	2	1						5		4													1			1	18
	?iih	1					9						1	2	1						1												1		1			1		18
u	sull																	13									5													18
	dhurr																	18																						18
	qull															4		14																						18
	?ubb	1	1													9		6												1										18
u:	suud																	8	7								2	1												18
	Tuumm																	12	5												1									18
	quuff																1	12	5																					18
	Huuh																1	13	1										2			1								18
Total		94	18	1	3	1	9	3	2	33	3	2	49	21	13	15	2	97	18	1	13	1	5	2	1	1	7	1	2	1	1	1	1	2	1	3	1	1	2	432

Table 6-5: Summary confusion matrix of the SB-audio-written results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; Listgp= listener group**, **T= monolingual listeners

Listgp		TA																			
SV		a	aa	ae	e	ee	ı	i	ii	o	u	uu	i-i	ö	ü	ei	üu	uü	a-ı	ıı	Total
a	faththa				10		1														11
	Tall	10			1																11
	qatt	7			3		1														11
	?al	5		1	5																11
a:	naabb	1	4		4	2															11
	Saaww	7	4																		11
	qaatt	7	4																		11
	Haaff	5	2	1	1	2															11
i	zirr						5	4							1					1	11
	Sirr						11														11
	qibb						8	2	1												11
	?izz				1		2	7								1					11
i:	tiinn							10					1								11
	Tiill						3	6	1				1								11
	qiibb						4	5					2								11
	?iih						1	6	1				2						1		11
u	sull										8			1	2						11
	dhurr										10	1									11
	qull						1			1	9										11
	?ubb									1	10										11
u:	suud										7	3					1				11
	Tuumm									1	9	1									11
	quuff										10	1									11
	Huuh										9	1						1			11
Total		42	14	2	25	4	37	40	3	3	72	7	6	1	3	1	1	1	1	1	264

Table 6-6: Summary confusion matrix of the SB-audio-written results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)
Red shade indicates response vowel≠predicted vowel (mismatch)
SV= stimulus vowel; Listgp= listener group*, *TA= bilingual listeners

Listgp		TQ																												
SV		a	aa	a-a	ae	a-e	a-i	e	ee	ı	i	ii	o	u	uu	i-i	ai	e-i	ü	i-e	NA	üü	uv	uü	u-e	ou	u-u	uw	ıı	i-e-e	Total
a	faththa							22																							22
	Tall	22																													22
	qatt	20											1	1																	22
	?al	15	2			1	1	3																							22
a:	naabb	11	2					5	4																						22
	Saaww	15	5	1	1																										22
	qaatt	14	8																												22
	Haaff	13	7					1		1																					22
i	zirr							1		7	14																				22
	Sirr							1		21																					22
	qibb									18	2	1		1																	22
	?izz							3		9	8	1					1														22
i:	tiinn									1	14	5						1		1											22
	Tiill									9	9	3				1															22
	qiibb									12	5	2				1													1	1	22
	?iih	1								10	4	1				1	2	1		1	1										22
u	sull													15					6					1							22
	dhurr										1			20	1																22
	qull													20					2												22
	?ubb	1											2	16												3					22
u:	suud													12	5				1			1	1	2							22
	Tuumm													14	5	1									1		1				22
	quuff													16	6																22
	Huuh										1			13	4				1				1	1				1			22
Total		112	24	1	1	1	1	36	4	88	58	13	3	128	21	4	3	2	10	2	1	1	2	4	1	3	1	1	1	1	528

Table 6-7: Summary confusion matrix of the SB-audio-written results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; Listgp= listener group, TQ= Turkish listeners with Quranic recitation knowledge

In addition to the previous tables, I include the perceptual maps yielded by the three groups in the Audio-written condition of the Simulated Borrowing experiment (henceforth SB-AW). These are given in tables 6-8 through 6-10.

[a]--> (14/18) /e/$_T$ (77.77%) {a:2, e:14}	[a]$_{d^ʕ}$ → (16/18) /a/$_T$ (88.88%) {a:16, o:2} [a]$_q$ → (18/18) /a/$_T$ (100%) {a:18} [a]$_ʕ$ → (17/18) /a/$_T$ (94.44%) {a:13, aa:1, a-a:1, ae:1, a-e:1, e:1} [a:] → (15/18) /a/$_T$ (83.33%) {a:7, aa:5, e:2, ea:3, ee:1} [a:]$_{d^ʕ}$ → (18/18) /a/$_T$ (100%) {a: 11, aa:2, ao:3, a-o:2} [a:]$_q$ → (18/18) /a/$_T$ (100%) {a:10, aa:8} [a:]$_ʕ$ --> (12/18) /a/$_T$ (66.66%) {a:10, aa:1, ae:1, e:5, ee:1}	[u]--> (13/18) /u/$_T$ (72.22%) {u:13, ü:5} [u]$_{d^ʕ}$ →(18/18) /u/$_T$ (100%) {u:18} [u]$_q$ --> (14/18) /u/$_T$ (77.77%) {u:14, o:4} [u]$_ʕ$ --> (9/18) /o/$_T$ (50%) {o:9, u:6, a:1, aa:1, au:1} [u:] → (15/18) /u/$_T$ (83.33%) {u:8, uu:7, ü:2, üü:1} [u:]$_{d^ʕ}$ → (18/18) /u/$_T$ (100%) {u:12, uu:5,uü:1} [u:]$_q$ → (17/18) /u/$_T$ (94.44%) {u:12, uu:5, oo:1} [u:]$_ʕ$ → (17/18) /u/$_T$ (94.44%) {u:13, uu:1, ue:1, uv:2, oo:1}	[i] --> (5/18) /ɯ/$_T$ (72.22%) {i:4, zr:1, ı:13} [i]$_{d^ʕ}$→(16/18) /ɯ/$_T$ (88.88%) {ı:15, e:1, ıı:1, ö:1} [i]$_q$ --> (12/18) /ɯ/$_T$ (66.66%) {ı:12, e: 2, a:4} [i]$_ʕ$ --> (9/18) /e/$_T$ (50%) {ae:1, e:8, i:2, a:1, ii:1, ai:1, ı:2, aı:2} [i:] → (16/18) /i/$_T$ (88.88%) {i:7, ii:8, i-i:1, u:1, ı-ı:1} [i:]$_{d^ʕ}$ --> (13/18) /i/$_T$ (72.22%) {i:4, ii:2, i-i:6, ui:1, ı-ı:2, u-ı:1, ıı: 1, e-ı:1} [i:]$_q$ --> (8/18) /i/$_T$ (44.44%) {i:2, ii:1, i-i:5, ı:4, ı-ı:4, ıı: 1, e-ı: 1} [i:]$_ʕ$ --> (14/18) /i/$_T$ (77.77%) { i:2, ii:1, i-i:1, ı-i:1, a-i:9, a:1, ı:1, a-ı:1, aı:1 ı -i:1}

Table 6-8: Perceptual maps of the T group in the Audio-written condition in the Simulated Borrowing experiment

[a]→ (10/11) /e/$_T$ (90.90%) {e:10, ı:1} [a:] --> (6/11) /e/$_T$ (54.54%) {a:1, aa:4, e:4, ee:2}	[a]$_{dˤ}$ → (10/11) /a/$_T$ (90.90%) {a:10, e:1} [a]$_q$ --> (7/11) /a/$_T$ (63.63%) {a:7, e:3} [a]$_ʕ$ --> (6/11) /a/$_T$ or /e/$_T$ (50%) {a:5, ae:1, e:5} [a:]$_{dˤ}$ → (11/11) /a/$_T$ (100%) {a: 7, aa:4} [a:]$_q$ → (11/11) /a/$_T$ (100%) {a:7, aa:4} [a:]$_ʕ$ --> (8/11) /a/$_T$ (72.72%) {a:5, aa:2, ae:1, e:1, ee:2}	[u]--> (8/11) /u/$_T$ (72.72%) {u:8, ö:1, ü:2} [u]$_{dˤ}$ →(11/11) /u/$_T$ (100%) {u:10, uu:1} [u]$_q$ → (9/11) /u/$_T$ (81.81%) {u:9, o:1, ı:1} [u]$_ʕ$ → (10/11) /u/$_T$ (90.90%) {u:10, uu:1} [u:] → (11/11) /u/$_T$ (100%) {u:7, uu:3, üu:1} [u:]$_{dˤ}$ → (10/11) /u/$_T$ (90.90%) {u:9, uu:1, o:1} [u:]$_q$ → (11/11) /u/$_T$ (100%) {u:10, uu:1} [u]$_ʕ$ → (11/11) /u/$_T$ (100%) {u:9, uu:1, uü:1}	[i] --> (6/11) /ɯ/$_T$ (54.54%) (i:4, ı:5, ıı:1, ü:1) [i]$_{dˤ}$→ (11/11) /i/$_T$ (100%) {ı:11} [i]$_q$ --> (8/11) /ɯ/$_T$ (72.72%) {ı:8, i:2, ii:1} [i]$_ʕ$ --> (8/11) /i/$_T$ (72.72%) {i:7, ei:1, ı:2, e:1} [i:] → (11/11) /i/$_T$ (100%) {i:10, i-i:1} [i:]$_{dˤ}$ --> (8/11) /i/$_T$ (72.72%) {i:6, ii:1, i-i:1, ı:3} [i:]$_q$ --> (7/11) /i/$_T$ (72.72%) {i:5, i-i:2, ı:4} [i:]$_ʕ$ → (9/11) /i/$_T$ (81.81%) { i:6, ii:1, i-i:2, ı:1}

Table 6-9: Perceptual maps of the TA group in the Audio-written condition in the Simulated Borrowing experiment

[a]→ (22/22) /e/$_{T}$ (100%) {e:22}	[a]$_{d^{ʕ}}$ → (22/22) /a/$_{T}$ (100%) {a:22} [a]$_{q}$ → (20/22) /a/$_{T}$ (90.90%) {a:20, o:1, u:1} [a]$_{ʕ}$ → (19/22) /a/$_{T}$ (86.36%) {a:15, aa:2, a-e:1, a-i:1, e:3} [a:] --> (13/22) /a/$_{T}$ (59.09%) {a:11, aa:2, e:5, ee:4} [a:]$_{d^{ʕ}}$ → (22/22) /a/$_{T}$ (100%) {a: 15, aa:5, a-a:1, ae:1} [a:]$_{q}$ → (22/22) /a/$_{T}$ (100%) {a:14, aa:8} [a:]$_{ʕ}$ → (20/22) /a/$_{T}$ (90.90%) {a:13, aa:7, e:1, ı:1}	[u]--> (16/22) /u/$_{T}$ (72.72%) {u:15, uü:1, ü:6} [u]$_{d^{ʕ}}$ →(21/22) /u/$_{T}$ (95.45%) {u:20, i:1, uu:1} [u]$_{q}$ → (20/22) /u/$_{T}$ (90.90%) {u:20, ü:2} [u]$_{ʕ}$ → (19/22) /u/$_{T}$ (86.36%) {u:16, o:2, a:1, ou:3} [u:] → (20/22) /u/$_{T}$ (90.90%) {u:12, uu:5, uv:1, uü:2, ü:1, üü:1} [u:]$_{d^{ʕ}}$ → (21/22) /u/$_{T}$ (95.45%) {u:14, uu:5,i-i:1, u-e:1,u-u:1} [u:]$_{q}$ → (22/22) /u/$_{T}$ (100%) {u:16, uu:6} [u]$_{ʕ}$ → (20/22) /u/$_{T}$ (90.90%) {u:13, uu:4, uv:1, uü:1, uw:1, i:1, ü:1}	[i] --> (14/22) /i/$_{T}$ (63.63%) {e:1, ı:7, i:14} [i]$_{d^{ʕ}}$→(21/22) /ɯ/$_{T}$ (95.45%) {e:1, ı:21} [i]$_{q}$ → (18/22) /ɯ/$_{T}$ (81.81%) {ı:18, i:2, ii:1, u:1} [i]$_{ʕ}$ --> (10/22) /i/$_{T}$ (45.45%) {i:8, ii:1, ai:1, e:3, ı:9} [i:] → (21/22) /i/$_{T}$ (95.45%) {i:14, ii:5, e-i:1, i-e:1 ı:1} [i:]$_{d^{ʕ}}$ --> (13/22) /i/$_{T}$ (59.09%) {i:9, ii:3, i-i:1, ı:9} [i:]$_{q}$ --> (13/22) /ɯ/$_{T}$ (59.09%) {ı:12, ıı:1, i:5, ii:2, i-i:1,i-e-e:1} [i:]$_{ʕ}$ --> (10/22) /i/$_{T}$ or /ɯ/$_{T}$ (45.45%) { i:4, ii:1, i-i:1, ai:2, e-i:1, i-e:1, ı:10, a:1, nill:1 }

Table 6-10: Perceptual maps of the TQ group in the Audio-written condition in the Simulated Borrowing experiment

6.3.1 Generalizations on the vowel categorizations of listener groups in the SB-audio and audio-written conditions

1. **Mapping of vowel /a/:** Listener groups tend to match the vowel /a/ categories mostly as predicted, followed by the vowel /u/ variants in both SB-audio and SB-audio-written conditions.

1.1 In the audio condition, the three listener groups categorized the variants of the vowel /a/ as predicted except for the plain long variant /a:/ which they mapped to the Turkish vowel /e/ as in table 6-11. This is the same categorization the three groups displayed in the PAT experiment and is reflected with a grey shade in table 6-8. As for the audio-written condition, only the TA group mismatched the the vowel /a:/ since the TA participants were evenly divided between /a:/ and /e/ as shown in table 6-9 (grey-shaded cell). They also categorized the pharyngealized vowel /a^{ʕ}/ as either /a/ or /e/ (yello-shaded cell). Interestingly, the monolingual group and the TQ group straighforwardly mapped the /a:/ this time as predicted.

Vowel length	Consonant context	T group mapping of /a/	TA group mapping of /a/	TQ group mapping of /a/
Short	Plain	/a/>/e/✓	/a/>/e/✓	/a/>/e/✓
	Emphatic	/aˤ/>/a/✓	/aˤ/>/a/✓	/aˤ/>/a/✓
	q	/aˤ/>/a/ ✓	/aˤ/>/a/ ✓	/aˤ/>/a/ ✓
	ʕ	/aˤ/>/a/ ✓	/aˤ/>/a/ ✓	/aˤ/>/a/ ✓
Long	Plain	/a:/>/e/ ×	/a:/>/e/ ×	/a:/>/e/ ×
	Emphatic	/a:/> /a/✓	/a:ˤ/> /a/✓	/a:/> /a/✓
	q	/a: ˤ/>/a/ ✓	/a:ˤ/>/a/ ✓	/a: ˤ/>/a/ ✓
	ʕ	/a: ˤ/>/a/ ✓	/a:ˤ/>/a/ ✓	/a: ˤ/>/a/ ✓

Table 6-11: vowel /a/ categorization in the SB-audio-written condition, a (✓) indicates a matched response and a (×) a mismatched one.

Vowel length	Consonant context	T group mapping of /a/	TA group mapping of /a/	TQ group mapping of /a/
Short	Plain	/a/>/e/✓	/a/>/e/✓	/a/>/e/✓
	Emphatic	/aˤ/>/a/✓	/aˤ/>/a/✓	/aˤ/>/a/✓
	q	/aˤ/>/a/✓	/aˤ/>/a/✓	/aˤ/>/a/✓
	ʕ	/aˤ/>/a/✓	/aˤ/>/a/✓, /e/×	/aˤ/>/a/✓
Long	Plain	/a:/>/a/✓	/a:/> /a:/×, /e/×	/a:/>/a/✓
	Emphatic	/a:/>/a/✓	/a:/> /a/✓	/a:/>/a/✓
	q	/a:ˤ/>/a/✓	/a:ˤ/>/a/✓	/a:ˤ/>/a/✓
	ʕ	/a:ˤ/>/a/✓	/a:ˤ/>/a/✓	/a:ˤ/>/a/✓

Table 6-12: vowel /a/ categorization in the SB-audio condition, a (✓) indicates a matched response and a (×) a mismatched one.

2. **Mapping of vowel /i/:** The categorization of the vowel /i/ was rather *messy* and *noisy* across the three listener groups compared to the mappings of /a/ and /u/. The TA group was the one that showed the least mismatch especially in the audio-written condition which was almost as predicted except for the plain short vowel /i/ in the Arabic word /zirr/ (button) which they mismatched to /ɯ/ probably because the stimulus word sounded like the Turkish word /sɯr̥/ (secret), a categorization which they shared with the T group but not the TQ group (table 6-14).

2.1. In the audio condition, the three groups mismatched the three pharyngealized vowels [iˤ]$_q$, [iˤ]$_ʕ$ and [i:ˤ]$_ʕ$ (where the subscript indicates the environment) which they mapped to /i/, /e/ and /a-ɯ/ respectively, albeit not to the same degree, as indicated with the grey shade in table (6-13).

Vowel length	Consonant context	T group mapping of /i/	TA group mapping of /i/	TQ group mapping of /i/
Short	Plain	/i/>/i/✓	/i/>/i/✓	/i/>/i/✓
	Emphatic	/iˤ/>/i/× predicted /ɯ/	/iˤ/>/ɯ/✓	/iˤ/>/ɯ/✓

	q	/iˤ/>/i/× predicted /ɯ/	/iˤ/>/i/× predicted /ɯ/	/iˤ/>/i/× predicted /ɯ/
	ʕ	/iˤ/>/e/×	/iˤ/>/e/×	/iˤ/>/e/×
Long	Plain	/i:/>/i/ ✓	/i: />/i/ ✓	/i: />/i/ ✓
	Emphatic	/i:ˤ/>/u-i/×	/i:ˤ/>/i/ ✓	/i:ˤ/>/i/ ✓; /i:ˤ/>/ ɯ/ ×
	q	/i:ˤ/>/ɯ-ɯ/×	/i:ˤ/>/i/ ✓	/i:ˤ/>/i/ ✓; /i:ˤ/>/ ɯ/ ×
	ʕ	/i:ˤ/>/a-ɯ/× predicted /i/	/i:ˤ/>/a-ɯ/ × predicted /i/	/i:ˤ/>/a-ɯ/ × predicted /i/

Table 6-13: vowel /i/ categorization in the SB-audio condition, a (✓) indicates a matched response and a (×) a mismatched one.

Vowel length	Consonant context	T group mapping of /i/	TA group mapping of /i/	TQ group mapping of /i/
Short	Plain	/i/>/ɯ/×	/i/>/ɯ/×	/i/>/i/✓
	Emphatic	/iˤ/>/ɯ/✓	/iˤ/>/ɯ/✓	/iˤ/>/ɯ/✓
	Q	/iˤ/>/ɯ/✓	/iˤ/>/ɯ/✓	/iˤ/>/ɯ/✓
	ʕ	/iˤ/>/e/×	/iˤ/>/i/✓	/iˤ/>/ɯ/×
Long	Plain	/i:/>/i:/×	/i:/>/i/✓	/i:/>/i/✓
	Emphatic	/i:ˤ/>/i-i/×	/i:ˤ/>/i/✓	/i:ˤ/>/i/✓, /ɯ/×
	Q	/i:ˤ/>/ɯ/×	/i:ˤ/>/i/✓	/i:ˤ/>/ɯ/×
	ʕ	/i:ˤ/>/a-i/×	/i:ˤ/>/i/✓	/i:ˤ/>/ɯ/×

Table 6-14: vowel /i/ categorization in the SB-audio-written condition, a (✓) indicates a matched response and a (×) a mismatched one.

2.2. In the audio-written condtion, the matching patterns of the listener groups varied. The TA group reflected more matching pattens as in table 6-14 compared to the audio condition. The same applies to the TQ group which yielded more matching patterns compared to the audio condition, however, with a lesser degree of matching compared to the TA group. As for the T group, their patterns did not seem to be affected by the stimulus presentation condition since they had more mismatch anyway.

3. <u>Mapping of the vowel /u/:</u>

All three groups mismatched the /u/ vowel in the plain environment to /u/ (predicted y <ü>) both in the audio and audio-written conditions, similar to the PAT results. This is reflected in tables 6-15 and 6-16.

Only the T group mismatched the vowel /u/ in a pharyngealized environment [uˤ] to /o/ (predicted /u/) both in the audio and audio-written conditions as depicted in tables 6-15 and 6-16 (yellow shaded cell).

Vowel length	Consonant context	T group mapping of /u/	TA group mapping of /u/	TQ group mapping of /u/
Short	Plain	/u/>/u/× predicted /y/	/u/>/u/× predicted /y/	/u/>/u/× predicted /y/
	Emphatic	/u/>/u/✓	/u/>/u/✓	/u/>/u/✓

	Q	/u/>/u/✓	/u/>/u/✓	/u/>/u/✓
	ʕ	/u/>/o/×	/u/>/u/✓	/u/>/u/✓
Long	Plain	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	Emphatic	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	Q	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	ʕ	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓

Table 6-15: vowel /u/ categorization in the SB-audio condition, a (✓) indicates a matched response and a (×) a mismatched one.

Vowel length	**Consonant context**	**T group mapping of /u/**	**TA group mapping of /u/**	**TQ group mapping of /u/**
Short	Plain	/u/>/u/×	/u/>/u/×	/u/>/u/×
	Emphatic	/u/>/u/✓	/u/>/u/✓	/u/>/u/✓
	Q	/u/>/u/✓	/u/>/u/✓	/u/>/u/✓
	ʕ	/u/>/o/×	/u/>/u/✓	/u/>/u/✓
Long	Plain	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	Emphatic	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	Q	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓
	ʕ	/u:/>/u/✓	/u:/>/u/✓	/u:/>/u/✓

Table 6-16: vowel /u/ categorization in the SB-audio-written condition, a (✓) indicates a matched response and a (×) a mismatched one.

Match across and within listener groups:

Match and listener group per stimulus presentation condition: Figure 6-1 and table 6-17 represent the match results across and within the three listener groups in the SB-audio and SB-audio-written conditions as explained below.

The average percentage of match in the audio-written condition is higher than in the audio condition at 61.9% to 48.44%.
The TA group incurred the highest degree of match in both audio-written and audio condition at 66.287% and 53.03% respectively, followed by the TQ group at 64.89% and 47.159% while the T group achieved the least degrees of match at 54.39% and 45.13% respectively.

		match		**mismatch**				**match**		**mismatch**	
Condition	**Listgp**	**%**	**count**	**%**	**count**	**Condition**	**Listgp**	**%**	**count**	**%**	**count**
Audio	T	45.138	195	54.861	237	**AW**	T	54.398	235	45.601	197
	TA	53.03	140	46.969	124		TA	66.287	175	33.712	89
	TQ	47.159	249	52.84	279		TQ	64.895	342	35.104	185
Total/average		48.44	222	51.56	258	**Total/average**		61.9	250.7	38.1	157

Table 6-17: Cross tabulation of match and Listgp in the audio and audio-written conditions

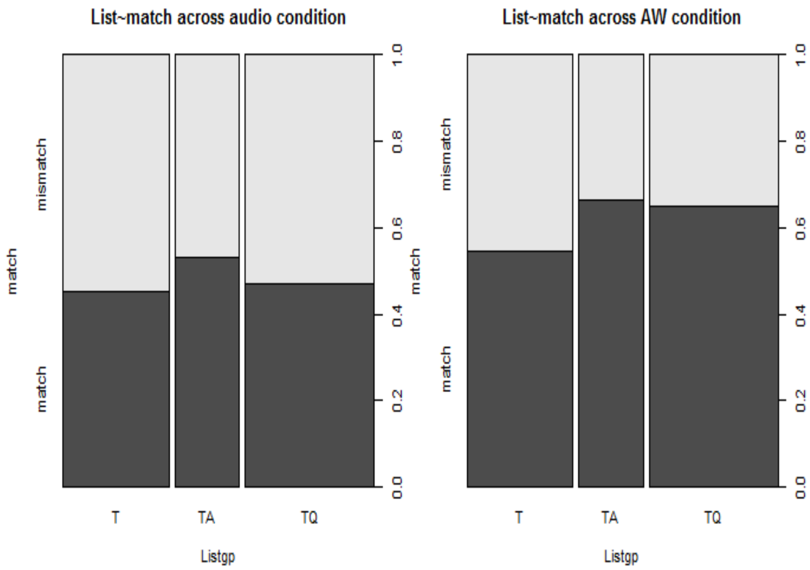


Figure 6-1: barplot of match ~Listgp percentage in audio and audio-written conditions

Match and listener group per vowel: Figure 6-2 and 6-3, and table 6-18 illustrate the match results across and within the three listener groups in the SB-audio and SB-audio-written conditions split by vowel quality. The following observations can be made.

Audio condition:

i. The percentage of match for the vowel /a/ in the audio condition is higher than mismatch at 58.16% (≅60%) across the three groups. The results of match in descending order are T: 59.72%, TQ: 57.95% and TA: 56.82%.
ii. The percentage of match for the vowel /i/ in the audio condition is less than mismatch at 26.99% match across the three groups; TA: 40.90%, TQ: 22.73% and T: 17.36%.
iii. The percentage of match for the vowel /u/ in audio condition is higher than mismatch at 60.16% match across the three groups. The groups' results in descending order are TA: 61.36%, TQ: 60.79% and T: 58.33%.

Thus, the percentage of match for the vowels /a/ and /u/ across the three groups in the audio condition is higher than mismatch (almost 60% match). However, the percentage of match for the vowel /i/ is lower than mismatch across the three groups (almost 30%) in the audio condition.

Audio-written condition:

i. The percentage of match for the vowel /a/ in the audio-written condition is higher than mismatch at 67.61% across the three groups. The groups' results in descending order is TQ: 75%, T: 68.75% and TA: 59.09%,
ii. The percentage of match for the vowel /i/ in the audio-written condition is a bit higher than mismatch at 50.41%. The groups' results in descending order of match are as follows: TA: 64.77%, TQ: 53.14% and T: 33.33%.
iii. The percentage of match for the vowel /u/ in the audio-written condition is higher than mismatch at 67.52%. The groups' results in descending order is TA: 75%, TQ: 66.47 and T: 61.11%.

Hence, the percentage of match for the vowels /a/ and /u/ across the three groups is high, almost 70% and is a bit higher than mismatch for the vowel i, almost 50%, across the three groups in the audio-written condition. The order of the listener groups of match for the vowels /i/and /u/ in both audio and audio-written condition is TA, followed by TQ and then T group. However, it is the T group followed by the TQ and then TA group for the vowel /a/ in the audio condition, and TQ, T and TA group for the vowel /a/ in the audio-written condition. The bwplot in figure 6-3 shows that the TQ group displayed more variation for the three vowels both in the audio condition and audio-written, followed by the TA group and finally the T group.

			match		mismatch					match		mismatch	
Condition	**Vowel**	**Listgp**	**%**	**count**	**%**	**count**	**Condition**	**Vowel**	**Listgp**	**%**	**count**	**%**	**count**
audio	a	T	59.722	86	40.277	58	AW	a	T	68.75	99	31.25	45
		TA	56.818	50	43.181	38			TA	59.09	52	40.909	36
		TQ	57.954	102	42.045	74			TQ	75	132	25	44
		Avg.	58.165	79.33	41.83	56.66			**Avg.**	67.61	94.33	32.38	41.66
	i	T	17.361	25	82.638	119		i	T	33.333	48	66.666	96
		TA	40.909	36	59.09	52			TA	64.772	57	35.227	31
		TQ	22.727	40	77.272	136			TQ	53.142	93	46.857	82
		Avg.	26.999	33.66	73	102.33			**Avg.**	50.41	66	49.58	69.66
	u	T	58.333	25	41.666	119		u	T	61.111	88	38.888	56
		TA	61.363	36	38.636	52			TA	75	66	25	22
		TQ	60.795	40	39.204	136			TQ	66.477	117	33.522	59
		Avg.	60.16	33.66	39.83	102.33			**Avg.**	67.52	90.33	32.47	45.66

Table 6-18: Percentage and count of match across Listgp per vowel in audio and audio-written conditions

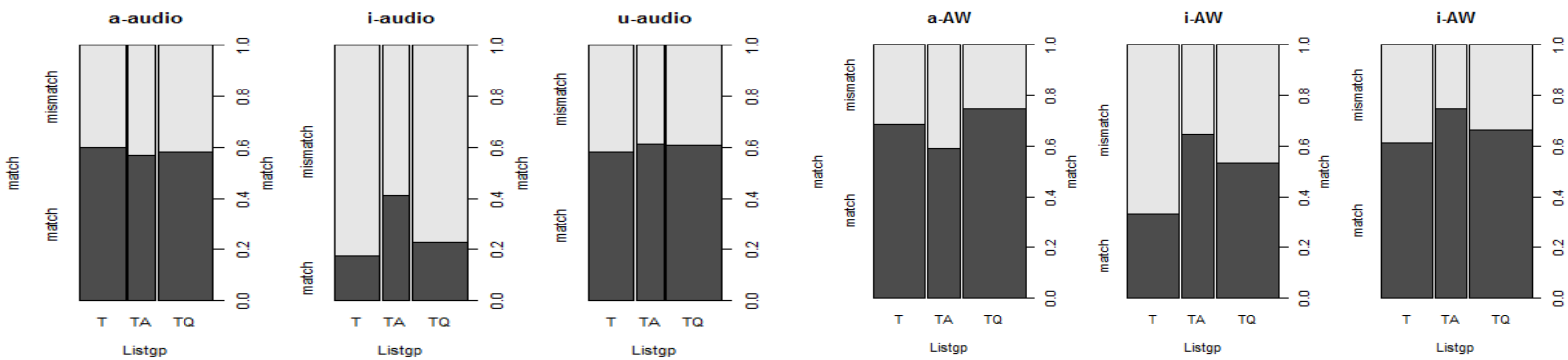


Figure 6-2:Barplots of the audio and audio-written data split by data across the three listener groups

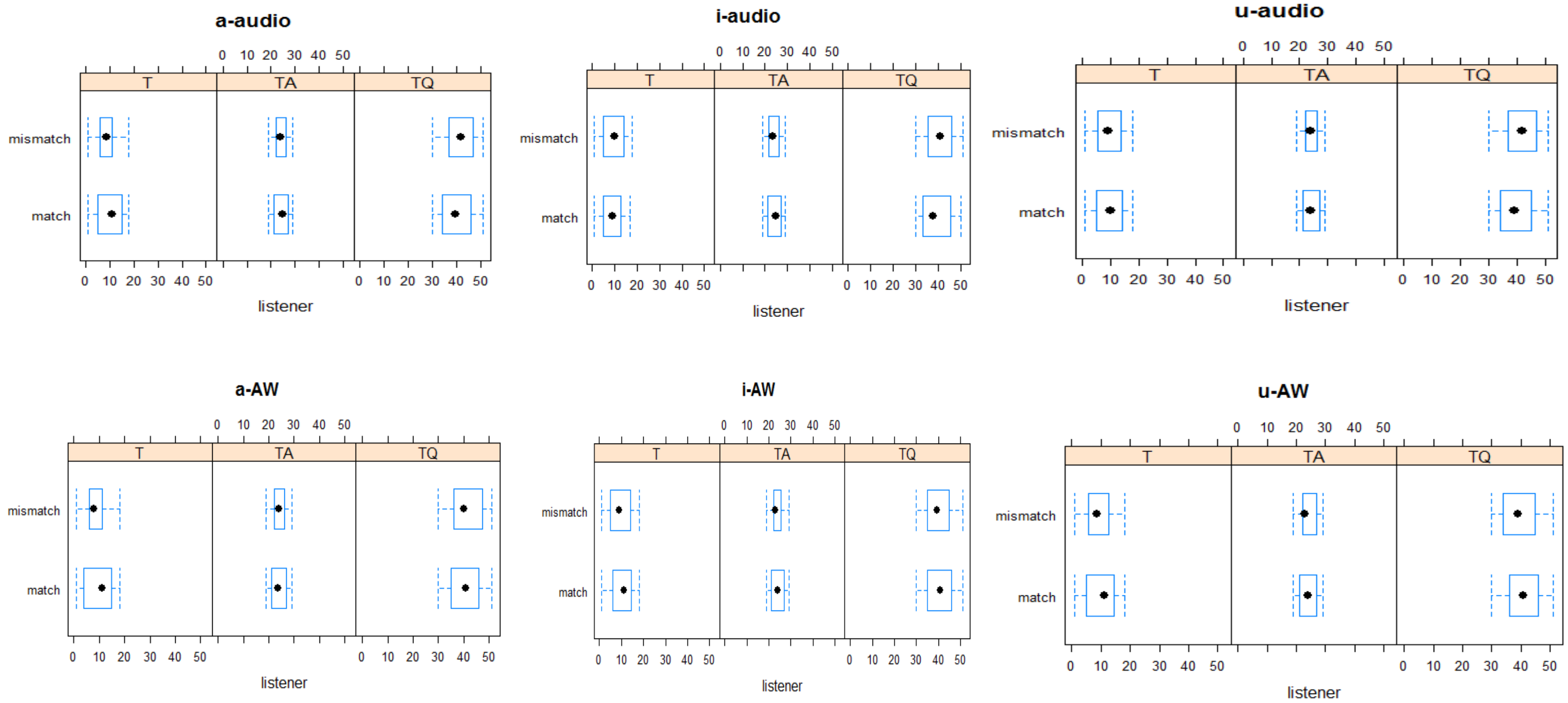


Figure 6-3: bwplots of match~Listgp per vowel in the audio and audio-written conditions

Listgp		TA																						
stimulus		a	aa	a-a	a-a-ı	a-e	a-ı	e	ea	ee	e-e	i	NA	ü	u-e	ö	ı	u	u-a	ıı	uu	i-i	a-u	Total
a	barr	3						7				1												11
	Sabb	5						4				1	1											11
	qabb	7				1		2				1												11
	?ayy	4			1		1	3	1			1												11
a:	haamm	4						6		1														11
	Saall	4	1	2				3			1													11
	qaabb	8						2		1														11
	?aall	5	1	2				2		1														11
i	nidd											11												11
	Sill	2										9												11
	qinn	1										10												11
	?ihh	1										8					1			1				11
i:	diik											11												11
	Tiin											7					2			1		1		11
	qiidd	1										9										1		11
	?iis	1										8					2							11
u	dubb													1		1		9						11
	Summ													2		1		7					1	11
	qunn															1		10						11
	?uth															1		10						11
u:	kuuzz													2		2		6	1					11
	Tuubbb															1		7	2		1			11
	quurr															1		8	1		1			11
	Huubb							1							1	1		7			1			11
Total		46	2	4	1	1	1	30	1	3	1	77	1	5	1	9	5	64	4	2	3	2	1	264

Table 6-19: Summary confusion matrix of the SB-written results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)

Red shade indicates response vowel≠predicted vowel (mismatch)

SV= stimulus vowel; Listgp= listener group*, *TA= bilingual listeners

Listgp		TQ																																														
stimulus		a	aa	a-a	aa-e	a-e	a-e-a	ae-e	ai	a-i	a-ı	b	e	ee	e-e	ee-e	eu	i	i-a	i-e	NA	ü	u-e	w	ı	u	ıi	uu	i-i	i-a-e	üü-e	ü-e	a-i-a	o	o-a-a	ı-ı	ıi-e	üü	i-e-a	ı-e-e	ü-a-e	u-a-e	ia	ü-ü	e-a	ü-a-a	ıı	ı-e	Total
a	barr											1	16		2			1	1		1																												22
	Sabb	13		1		2							2		2					1		1																											22
	qabb	16		1		3						1	1																																				22
	?ayy	12	1	1		1				2			1					1		2			1																										22
a:	haamm	2	1			1	1			1			10	2	1	1					1			1																									22
	Saall	10	5	1		1						1	1	1			1			1																													22
	qaabb	10	4		1	4		1	1						1																																		22
	?aall	11	3	4		1				1								1		1																													22
i	nidd			1									1					17	1	2																													22
	Sill	1								1		1						15		3							1																						22
	qinn																	18		3												1																	22
	?ihh	2																10		3					4		1		1														1						22
i:	diik											1						10		4							3		1	1						1			1										22
	Tiin									2								13		3					1		1																				1	1	22
	qiidd	1									1				1			8	1	1					5		2		1				1																22
	?iis														1			10	1	3					3		1									1	1			1									22
u	dubb	1		1																		6	4			9						1																	22
	Summ	1		1													1	1				4	1			10		2				1																	22
	qunn												1					1				3	2			15																							22
	?uth										1									1		3	1			16																							22
u:	kuuzz					1																5	3			6		1			1			1				2			1			1					22
	Tuubbb														1							3	1			10		4				1			1							1							22
	quurr	1		1																		3	1			10		4																	1	1			22
	Huubb	2				1																2				11		5														1							22
Total		83	14	12	1	15	1	1	1	7	2	5	33	3	9	1	2	106	4	28	2	30	14	1	13	87	9	16	3	1	1	4	1	1	1	2	1	2	1	1	1	2	1	1	1	1	1	1	528

Table 6-20: Summary confusion matrix of the SB-written results (actual count of tokens)

Green shade indicates response vowel=predicted vowel (match)
Red shade indicates response vowel≠predicted vowel (mismatch)
SV= stimulus vowel; Listgp= listener group, TQ= Turkish Quranic listeners

6.3.2 Generalizations on the vowel categorizations of listener groups in the SB- written conditions

1. **Mapping of vowel /a/:**

1.1. Both TA and TQ groups matched all the vowel /a/ categories as per the corpus except for the long plain variant which they mismatched to /e/ as in the PAT and SB-audio condition. This is reflected in table 6-21 and indicated with the grey shade.

Vowel length	Consonant context	TA group mapping of /a/	TQ group mapping of /a/
Short	Plain	/a/>/e/✓	/a/>/e/✓
	Emphatic	/a^ʕ/>/a/✓	/a^ʕ/>/a/✓
	Q	/a^ʕ/>/a/✓	/a^ʕ/>/a/✓
	ʕ	/a^ʕ/>/a/✓	/a^ʕ/>/a/✓
Long	Plain	/a:/> /e/×	/a:/>/e/×
	Emphatic	/a:/> /a/✓	/a:/>/a/✓
	Q	/a:^{ʕ}/>/a/✓	/a:^{ʕ}/>/a/✓
	ʕ	/a:^{ʕ}/>/a/✓	/a:^{ʕ}/>/a/✓

Table 6-21: vowel /a/ categorization in the SB-written condition

2. **Mapping of vowel /i/:**

Both TA and TQ groups mismatched the emphatic and q vowel /i/ variants (in the words s^{ʕ}i^{ʕ}ll>sill and qi^{ʕ}nn>qɯn) to the vowel /i/ as shown in table 6-22. This is comparable to their categorization of the q variant in the SB-audio condition but not the pharyngealized one which they mismatched to /e/. Additionally, it was only the TQ group who mismatched the pharyngealized /i^ʕ/to /ɯ/ instead of the predicted /i/ in the SB-audio-written condition (table 6-14).

Vowel length	Consonant context	TA group mapping of /i/	TQ group mapping of /i/
Short	Plain	/i/>/i/✓	/i/>/i/✓
	Emphatic	/i^ʕ/>/i/×	/i^ʕ/>/i/×
	Q	/i^ʕ/>/i/×	/i^ʕ/>/i/×
	ʕ	/i^ʕ/>/i/✓	/i^ʕ/>/i/✓
Long	Plain	/i:/>/i/✓	/i:/>/i/✓
	Emphatic	/i:^{ʕ}/>/i/✓	/i:^{ʕ}/>/i/✓
	Q	/i:^{ʕ}/>/i/✓	/i:^{ʕ}/>/i/✓
	ʕ	/i:^{ʕ}/>/i/✓	/i:^{ʕ}/>/i/✓

Table 6-22: vowel /i/ categorization in the SB-written condition

3. Mapping of vowel /u/:

Both TA and TQ groups mapped the categories of the vowel /u/ as predicted based on the corpus except for the plain short variant /u/ which they mismatched to /u/ as in table 6-23. This is the same pattern they followed in all three SB conditions.

Vowel length	Consonant context	TA group mapping of /u/	TQ group mapping of /u/
Short	Plain	/u/>/u/×	/u/>/u/×
	Emphatic	/u/>/u/✓	/u/>/u/✓
	q	/u/>/u/✓	/u/>/u/✓
	ʕ	/u/>/u/✓	/u/>/u/✓
Long	Plain	/u:/>/u/✓	/u:/>/u/✓
	Emphatic	/u:/>/u/✓	/u:/>/u/✓
	q	/u:/>/u/✓	/u:/>/u/✓
	ʕ	/u:/>/u/✓	/u:/>/u/✓

Table 6-23: vowel /u/ categorization in the SB-written condition

4. Match/mismatch across and within listener groups:

Table 6-24 and figure 6-4 below compare the match results of the three listener groups across the SB-audio and SB-written conditions.

4.1. The percentage of match for the TA group in both the audio condition and written condition is higher than the TQ group. It is 53.03% and 57.5758 for TA respectively, and it is 47.15% and 43.93% for the TQ respectively.

4.2. The average percentage of match in the written condition is very similar to the audio condition at 50.76% match in the written condition compared to 50.0945% in the audio.

		NA		Match		Mismatch	
Condition	**Listgp**	**%**	**Count**	**%**	**Count**	**%**	**Count**
Audio	TA	0	0	53.03	140	46.969	124
	TQ	0	0	47.159	249	52.841	279
Average		0	0	50.0945	194.5	49.905	201.5
		NA		**Match**		**mismatch**	
Condition	**Listgp**	**%**	**count**	**%**	**Count**	**%**	**count**
Written	TA	0.378	1	57.5758	152	42.0455	111
	TQ	0	0	43.9394	232	56.0606	296
Average		0.1893	0.5	50.7575	192	49.05303	203.5

Table 6-24: Crosstabulation of match~Listgp per stimulus presentation condition

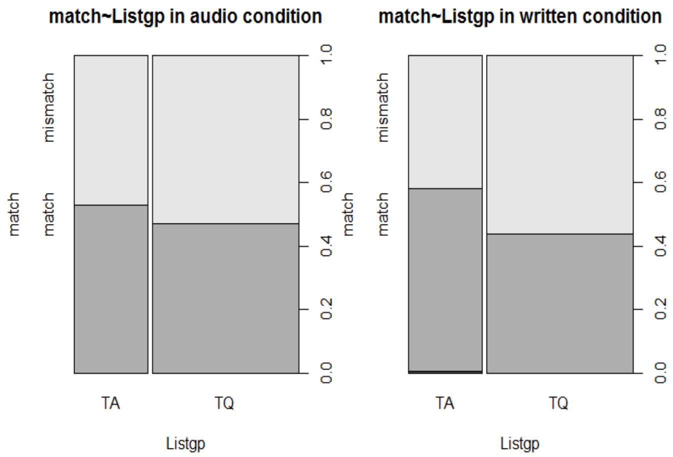


Figure 6-4:Barplot of match~Listgp in audio and written conditions

5. Match and listener group per vowel:

Tables 6-25 and 6-26 and figure 6-5 illustrate the percentage and count of match/mismatch results across and within the TA and TQ listener groups in the SB-audio and SB written conditions split by the vowel quality. The following observations can be made.

			Match		Mismatch	
Condition	**Vowel**	**Listgp**	**%**	**count**	**%**	**count**
Audio	**A**	TA	56.818	50	43.182	38
		TQ	57.954	102	42.045	74
		average	57.386	76	42.613	56
	I	TA	40.909	36	59.09	52
		TQ	22.727	40	77.272	136
		average	31.818	38	68.181	94
	U	TA	61.363	54	38.636	34
		TQ	60.795	107	39.204	69
		average	61.079	80.5	38.92	51.5

Table 6-25: Crosstabulation of match~Listgp per vowel in the audio condition

i. The percentage of match for the vowel /a/ in the audio condition is higher than mismatch at 57.386% (≅ 60%) across the three groups. The results of each group in descending order are TQ: 57.954% and TA: 56.818% which are close to each other.

ii. The percentage of match for the vowel /i/ in the audio condition is less than mismatch at 31.82% across the three groups. The match results of the groups in descending order are as follows; TA: 40.91% and TQ: 22.73%.

iii. The percentage of match for the vowel /u/ in the audio condition is higher than mismatch at 61.08% across the three groups. The results of the two groups in descending order is as follows; TA: 61.36% and TQ: 60.79% which are close to each other.

iv. Thus, the percentage of match for the vowels /a/ and /u/ in the audio condition is higher than mismatch across the TA and TQ groups at almost 60% match. On the other hand, the percentage of match for the vowel /i/ in the audio condition is lower than mismatch across the two groups, almost 30%.

			NA		Match		mismatch	
Condition	**Vowel**	**Listgp**	**%**	**count**	**%**	**Count**	**%**	**count**
Written	**a**	TA	1.1363	1	50	44	48.863	43
		TQ	0	0	52.840	93	47.159	83
		average	0.5681	0.5	51.420	68.5	48.011	63
	i	TA	0	0	60.227	53	39.772	35
		TQ	0	0	30.113	53	69.886	123
		average	0	0	45.170	53	54.829	79
	u	TA	0	0	62.5	55	37.5	33
		TQ	0	0	48.863	86	51.136	90
		average	0	0	55.681	70.5	44.318	61.5

Table 6-26: Crosstabulation of match~Listgp per vowel in the written condition

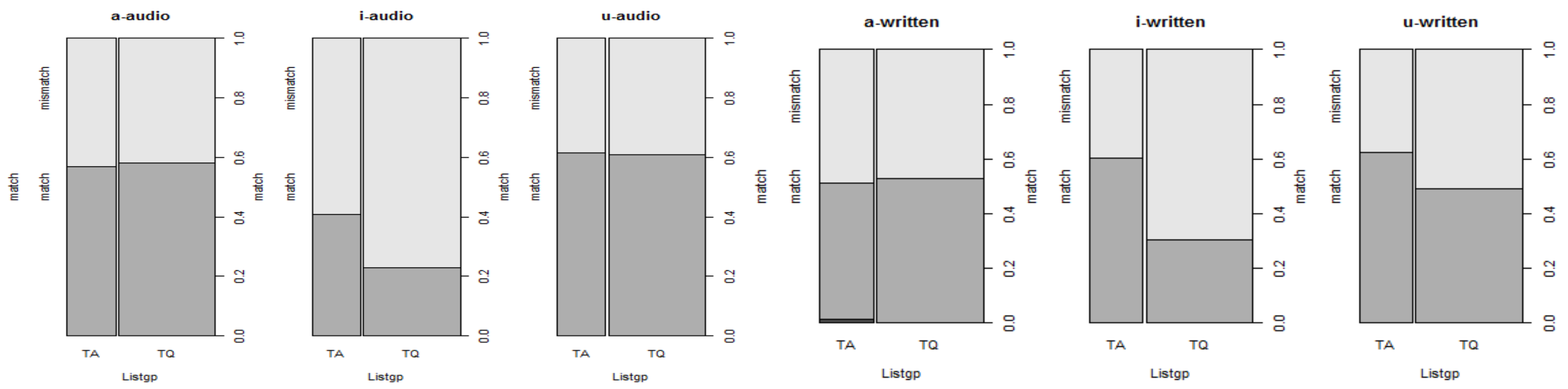
i. The percentage of match for the vowel /a/ in the written condition is higher than mismatch and the NA level (missing value) at 51.42%. The match results of the two groups in descending order is TQ: 52.8% and TA: 50%.

ii. The percentage of match for the vowel /i/ in the written condition is less than mismatch at 45.17%. The order of the match results for the two groups is TA: 60.23% and TQ: 30.11%.

iii. The percentage of match for the vowel /u/ in the written condition is higher than mismatch at 55.68%. The order of the match results for each group is as follows; TA: 62.5% and TQ: 48.86%.

iv. Thus, the percentage of match for the vowels /a/ and /u/ in the written condition across the TA and TQ groups is higher than mismatch at 51.42% and 55.68% respectively. However, match is less than mismatch for the vowel /i/ at 45.17% in the written condition across the two groups.

v. The bwplot (box-and-whisker plot) shows that the TQ groups displays more variation for the three vowels both in the audio condition and written condition than the TA group.



Barplots of match~Listgp per vowel in audio condition and written condition

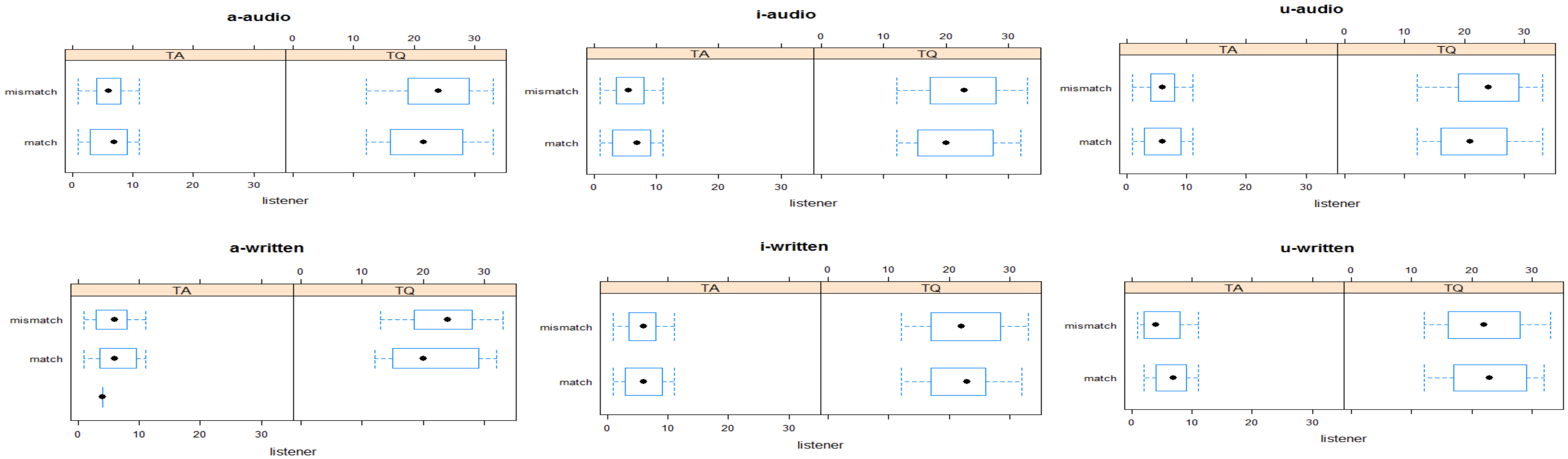


Figure 6-5: Barplots and bwplots of match~Listgp per vowel in the audio condition and written condition

In the last part of this section, the perceptual maps of the TA and TQ groups in the written condition of the Simulated Borrowing (henceforth, SB-W) are given in tables 6-27 and 6-28.

[a] --> (7/11) /e/$_T$ (63.63%) {a:3, e:7, i:1} [a:] --> (7/11) /e/$_T$ (63.63%) {a:4, e:6, ee:1}	[a]$_{dˤ}$ --> (5/11) /a/$_T$ (45.45%) {a:5, e:4m i:1, nill:1} [a]$_q$ --> (8/11) /a/$_T$ (72.72%) {a:7, a-e:1, e:2, i:1} [a]$_ʕ$ --> (7/11) /a/$_T$ (63.63%) {a:4, a-a:1, a-ı:1, e:3, ea:1, i:1} [a:]$_{dˤ}$ --> (7/11) /a/$_T$ (63.63%) {a: 4, aa:1, a-a:2, e:3, e-e:1} [a:]$_q$ --> (8/11) /a/$_T$ (72.72%) {a:8, e:2, ee:1} [a:]$_ʕ$ --> (8/11) /a/$_T$ (72.72%) {a:5, aa:1, a-a:2, e:2, ee:1}	[u]→ (9/11) /u/$_T$ (81.81%) {ü:1, u:9, ö:1} [u]$_{dˤ}$ →(8/11) /u/$_T$ (72.72%) {u:7, a-u:1, ö:1, ü:2} [u]$_q$ → (10/11) /u/$_T$ (90.90%) {u:10, ö:1} [u]$_ʕ$ → (10/11) /u/$_T$ (90.90%) {u:10, ö:1} [u:] --> (7/11) /u/$_T$ (63.63%) {u:6, u-a:1, ö:2, ü:2} [u:]$_{dˤ}$ → (10/11) /u/$_T$ (90.90%) {u:7, u-a:2, uu:1, ö:1} [u:]$_q$ → (10/11) /u/$_T$ (90.90%) {u:8, u-a:1, uu:1, ö:1} [u:]$_ʕ$ → (9/11) /u/$_T$ (81.81%) {e:1, u:7, uu:1, u-e:1, ö:1}	[i] → (11/11) /i/$_T$ (100%) {i:11} [i]$_{dˤ}$→ (9/11) /i/$_T$ (81.81%) {a:2, i:9} [i]$_q$ → (10/11) /i/$_T$ (90.90%) {a:1, i:10} [i]$_ʕ$ → (9/11) /i/$_T$ (81.81%) {a:1, i:8, ı:1, ıi:1} [i:] → (11/11) /i/$_T$ (100%) {i:11} [i:]$_{dˤ}$→ (9/11) /i/$_T$ (81.81%) {i:7, i-i:1, ı:2, ıi:1} [i:]$_q$→ (10/11) /i/$_T$ (90.90%) {a:1, i:9, i-i:1} [i:]$_ʕ$ --> (8/11) /i/$_T$ (72.72%) {a:1, i:8, ı:2}

Table 6-27: Perceptual maps of the TA group in the Written condition in the Simulated Borrowing experiment

[a]→ (18/22) /e/$_T$ (81.81%) {e:16, e-e:2, b:1, i:1, i-a:1, nill:1} [a:] --> (14/22) /e/$_T$ (63.63%) {e:10, ee:2, e-e:1, ee-e:1, nill:1, a:2, aa: 1}	[a]$_{d^ʕ}$ --> (16/22) /a/$_T$ (72.72%) {a:13, a-a:1, a-e:2, e:2, e-e:2, i-e: 1, ü:1} [a]$_q$ → (20/22) /a/$_T$ (90.90%) {a:16, a-a:1, a-e:3, b:1, e:1} [a]$_ʕ$ --> (17/22) /a/$_T$ (77.27%) {a:12, aa:1, a-a:1, a-e:1, a-i:2, e:1, i:1, i-e:2, u-e:1} [a:]$_{d^ʕ}$ --> (17/22) /a/$_T$ (77.27%) {a: 10, aa:5, a-a:1, a-e:1, b:1, e:1, ee:1, eu:1, i-e:1} [a:]$_q$ → (21/22) /a/$_T$ (95.45%) {a:10, aa:4, aa-e:1, a-e:4, ae-e:1, ai:1, ee:1} [a:]$_ʕ$ --> (20/22) /a/$_T$ (90.90%) {a:11, aa:3, a-a: 4, a-e:1, a-i:1, i:1, i-e:1}	[u]--> (13/22) /u/$_T$ (59.09%) {u:9, u-e:4, a:1, a-a:1, ü:6, ü-e:1} [u]$_{d^ʕ}$ --> (14/22) /u/$_T$ (63.63%) {u:10, uu:2, u-e:1, eu:1, a:1, a-a:1, i:1, ü:4,ü-e:1} [u]$_q$ --> (17/22) /u/$_T$ (77.27%) {u:15, u-e:2, e:1, i:1, ü:3, } [u]$_ʕ$ --> (17/22) /u/$_T$ (77.27%) {u:16, u-e:1} [u:] --> (10/22) /u/$_T$ or /ü/$_T$ (45.45%) {u:6, u-e:3, uu:1, a-e:1, ü:5, üü-e:1, o:1, üü:2, ü-a-e:1, ü-ü:1} [u:]$_{d^ʕ}$ --> (16/22) /u/$_T$ (72.72%) {u:10, uu:4,u-a-e:1, u-e:1, e-e:1, ü:3, ü-e:1, o-a-a:1} [u:]$_q$ --> (15/22) /u/$_T$ (68.18%) {u:10, uu:4, u-e:1, a:1, a-a:1, ü:3, ɪ:e-a:1, ü -a-a:1 } [u]$_ʕ$ --> (17/22) /u/$_T$ (77.27%) {u:11, uu:5, u-a-e:1, a:2, a-e:1, ü:2}	[i] → (20/22) /i/$_T$ (90.90%) {a-a:1, e:1, i:17, i-a:1, i-e:2} [i]$_{d^ʕ}$→ (20/22) /i/$_T$ (90.90%) {i:15, a-i:1, i-e:3, ɪi:1, a:1, b:1} [i]$_q$ →(21/22) /i/$_T$ (95.45%) {i:18, i-e:3, ü-e:1} [i]$_ʕ$ --> (16/22) /i/$_T$ (72.72%) {a:2, i:10, i-e:3, ɪ:4, ɪi:1, i-i:1, ia:1} [i:] → (22/22) /i/$_T$ (100%) {i:10, i-e:4, ɪi:3, i-i:1, i-a-e:1, ɪi:3, i-e-a:1} [i:]$_{d^ʕ}$ → (19/22) /i/$_T$ (86.36%) {i:13, i-e:3, a-i:2, ɪi:1, ɪ:1, ɪi:1, ɪ-e:1} [i:]$_q$ --> (14/22) /i/$_T$ (63.63%) {a:1, a-ɪ:1, e-e:1, i:8, i-a:1, i-e:1, ɪ:5, ɪi:2, i-ɪ:1, a-i-a:1} [i:]$_ʕ$ --> (16/22) /i/$_T$ (72.72%) { i:10, e-e:1, i-a:1, i-e:3, ɪ:3, ɪi:1, ɪ-ɪ:1, ɪi-e:1, ɪ-e-e:1}

Table 6-28: Perceptual maps of the TQ group in the Written condition in the Simulated Borrowing experiment

In the next section, other variables besides listener group are tested to measure whether these variables and/or their interactions have any effect on match. This is done through logistic regression GLMM in R (Team, 2015) on SB-audio and SB-audio-written data and next separately on SB-written data in section 6.5.

6.4 *Logistic regression of SB-audio and audio-written stimuli*

Following the exploration of the raw data, 2448 observations from 51 listeners; 18 T, 11 TA and 22 TQ participants were collected for the SB-audio and audio-written conditions (24 tokens per condition) and analyzed using Generalized Linear Mixed Effects Modelling (GLMM) in R (Team, 2015)using the Laplace Approximation method in the lme4 package (2015). A summary of the descriptive statistics of the variables involved is given in table 6-29.

Listener	Listgp	Gender	age	stimulus.presentation	vowel.quality
T1 : 48	T : 864	Female:1104	Min. :17	audio :1224	a:816
T10 : 48	TA: 528	Male :1344	1st Qu.:24	audio+written:1224	i:816
T11 : 48	TQ:1056		Median :29		u:816
T12 : 48			Mean :32		
T13 : 48			3rd Qu.:37		
T14 : 48			Max. :62		
(Other):2160					
st.vowel	**stimulus**	**Match**	**length**	**Context**	**freq.**
a :408	?al : 51	match :1336	long :1224	emphatic :612	Min. : 0.00
aa:408	?an : 51	mismatch:1111	short:1224	pharyngeal:612	1st Qu.: 0.44
i :408	?iih : 51	NA's : 1		plain :612	Median : 1.49
ii:408	?iirr : 51			q :612	Mean :10.38
u :408	?iyy : 51				3rd Qu.:13.09
uu:408	?izz : 51				Max. :75.67
	(Other):2142				

Table 6-29: descriptive statistics of the SB audio and audio-written data

6.4.1 Objective

A mixed effects logistic regression was performed to test whether knowledge of Arabic spelling will have an effect on the degree of match. The participants' groups consisted of monolinguals (T), bilinguals (TA) and Turkish speakers with knowledge of Arabic through Quranic recitation (TQ) who responded to two tasks, the first of which had audio only stimuli followed by audio-written stimuli.

6.4.2 Hypotheses and predictions:

Prior to defining the hypotheses in operational terms, the variables involved were highlighted.

Fixed effects structure based on hypotheses.

The same fixed effects defined in the audio-only condition (chapter 5) were examined in the audio-written condition of the simulated borrowing experiment in addition to the new variable of ***stimulus presentation*** which included the two levels of audio and audio-written.

Therefore, seven variables in total were included in the fixed effects structure as follows where the response variable is **match** and the two random effects are **listener** and **stimulus**.

- *Response variable:* **match**
- *Random effect:* **listener**
- *Random effect:* **stimulus**
- Listener group; T, TA and TQ: **Listgp**
- Stimulus length; long or short: **length**
- Stimulus context; emphatic, plain, pharyngeal, q: **context**
- Stimulus frequency (a continuous variable): **freq.**
- Age of the participant at the time of the experiment (a continuous predictor): **age**
- Stimulus presentation condition; either audio or audio-written: **stimulus.presentation**
- Vowel quality: either a, i or u, **vowel.quality.**

The two hypotheses derived from the RQ2 in 1. which was introduced in the objective section in 6.4.1. include the null hypothesis and the experimental/alternative hypothesis both given below.

- H_0: the TA and TQ groups (being the groups with knowledge of Arabic) will not exhibit closer degrees of match to the corpus in the audio-written condition compared to the audio condition.
- H1: the TA and TQ groups will incur closer degrees of match to the corpus in the audio-written stimulus presentation condition than they will in the audio only condition.

In the next subsection, the same protocol of model selection that was used in chapters 4 and 5 is applied to the SBAAW data; i.e. the Simulated Borrowing data with audio and audio-written data.

6.4.3 Protocol for model selection

6.4.3.1 Defining the fixed and random effects structures

i. Fixed effects structure: (data exploration)

As mentioned **in section** 6.4.1. seven exploratory variables were included in the fixed effects structure. Two of these are by-listener variables including Listgp and age whereas the remaining five are by-stimulus including stimulus.presentation, length, context, freq. and vowel.quality.

Some of these variables are variables of interest (VOI) such as ***Listgp*** (since it relates to knowledge of Arabic), ***stimulus.presentation*** (since it relates to the two conditions of stimulus presentation), ***length*** (as it relates to the length of the stimulus vowel being long or short), ***context*** (which relates to the consonants surrounding the stimulus vowels be them an emphatic, plain, pharyngeal or q), ***freq.*** (which denotes frequencies of real and nonsense words) and ***vowel.quality*** (which relates to the vowel type being either a, i or u). On the other hand, ***age*** is chosen as a control variable.

Graphical data exploration

Each of the seven fixed effects was plotted against the response variable *match* to check which of them reflected variability. All of them showed variability as illustrated by figure 6-6 below which ascertains that they should be part of the maximal model.

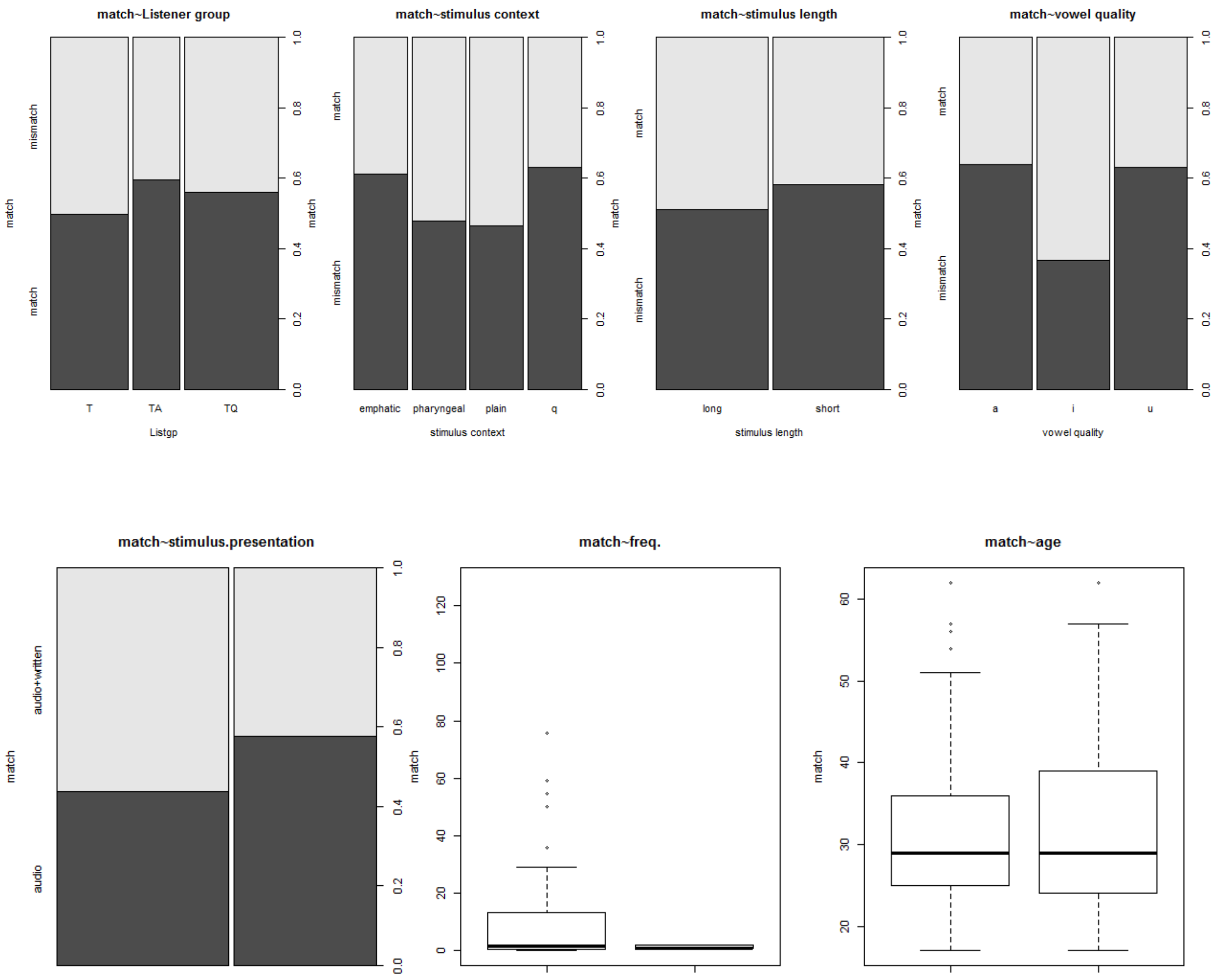


Figure 6-6: Fixed effects and the response variable 'match'

Next, eleven interactions in addition to the seven fixed variables were included in the fixed effects structure of the maximal model, with the interaction between ***Listgp*** and ***stimulus.presentation*** being the interaction of main interest. These include the following where the (:) indicates an interaction: ***Listgp:context***, ***Listgp:length***, ***Listgp:freq.***, ***Listgp:vowel.quality***, ***Listgp:stimulus.presentation***, ***context:length***, ***context:freq.***, ***context:vowel.quality***, ***length:freq.***, ***freq.:vowel.quality*** and ***age:vowel.quality***. Although the main interaction ***Listgp:stimulus.presentation***[31] did not reflect variability when plotted (figure 6-7), it was retained since it is the interaction of main interest.

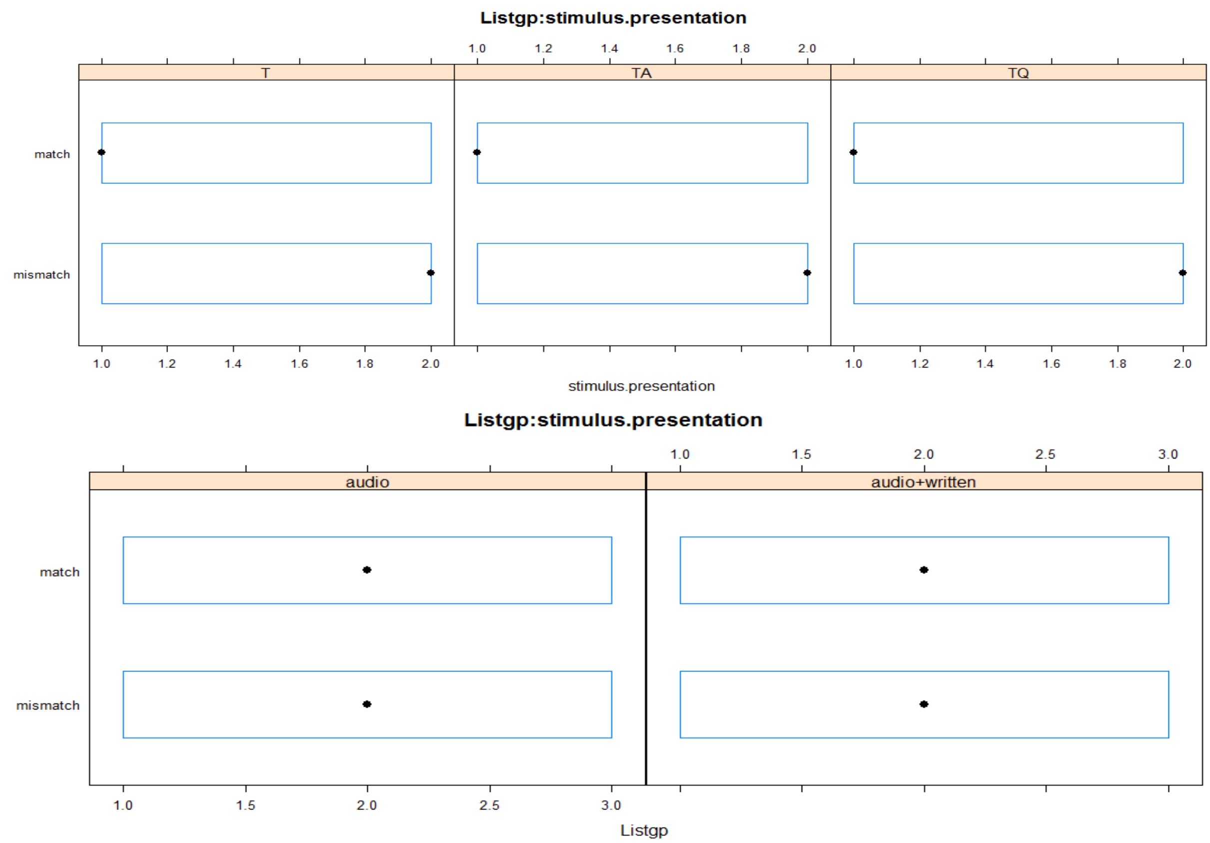


*Figure 6-7: bwplot of Listgp*stimulus presentation*

[31] The interaction Listgp:stimulus.presentation was plotted using the Lattice package(Sarkar, 2008). The codes below yield two different ways of visualizing this interaction, however with Listgp as the grouping factor in the first code and stimulus presentation as a grouping factor in the second.
bwplot(match ~ stimulus.presentation| Listgp, main= "Listgp:stimulus.presentation", data = SBAAW)
bwplot(match ~ Listgp| stimulus.presentation, main= "Listgp:stimulus.presentation", data = SBAAW)

Some other interactions such as ***Listgp:age***, ***context:age***, ***freq.:age*** were dropped from the maximal model for model simplification purposes and because they involved the control variable age which is not a variable of interest. Other interactions were not included in the final model because they did not display statistically significant variation such as ***context:stimulus.presentation***, ***length:stimulus.presentation***, ***length:age***, ***vowel.quality:stimulus.presentation*** and ***age:stimulus.presentation***. Two more interactions were dropped for model simpllification purposes and since they were not part of the research hypotheses despite reflecting variability. These are ***length:vowel.quality*** and ***freq.:stimulus.presentation***, however, ***length:vowel.quality*** was added at a later stage and reflected in the final model. Box and whisker plots of the interactions that were included in the maximal model appear in appendix 6-1.

ii. Random effects structure:

In this part, first the variation coefficient partition VCP (Steele, 2008b) is calculated to determine which random effects contribute most to the variation in the outcome variable. Next, the structure of the random effects is determined to check whether it would consist of slopes only, intercepts only or both (i.e., maximal; Barr et al, (2013)).

Variation inspection is done via plotting the random effects against the response variable and by examining the summary table of the null model. Figures 6-8 and 6-9 below demonstrate the relationships between each of the two random effects, i.e., listener and stimulus with match respectively whereas table 6-30 represent in order the random effects summary table of the null model.

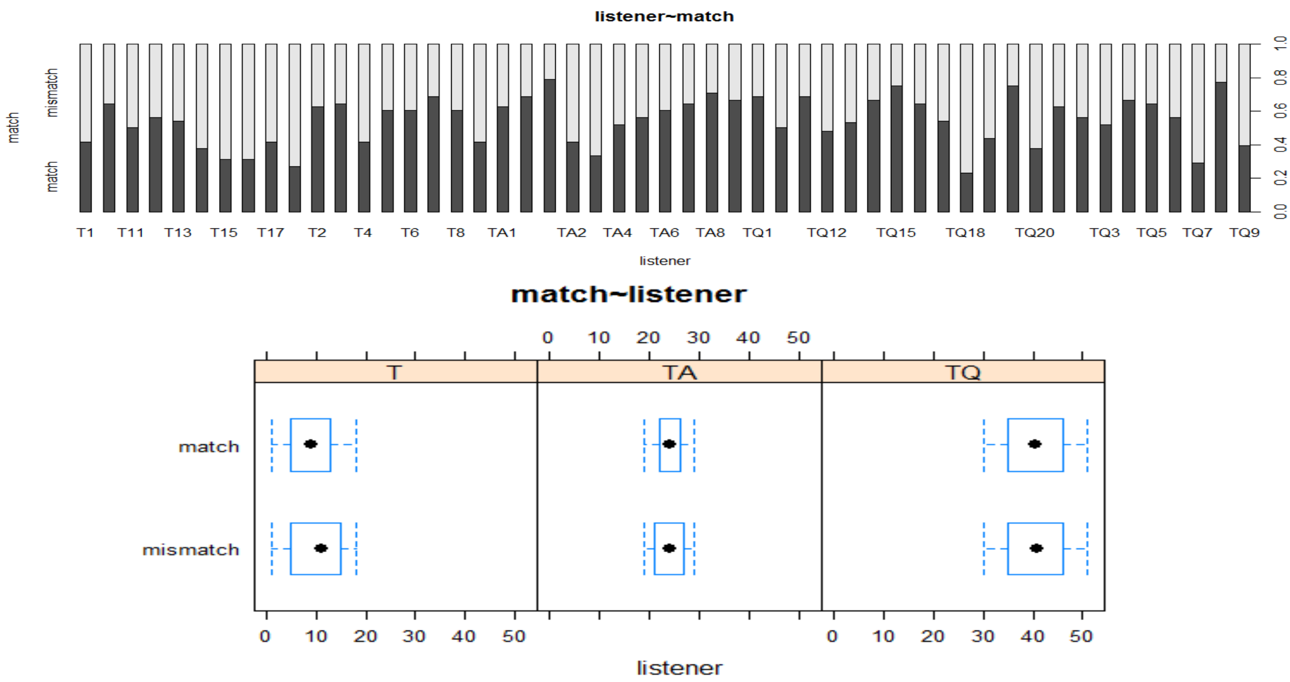


Figure 6-8: Listener~match

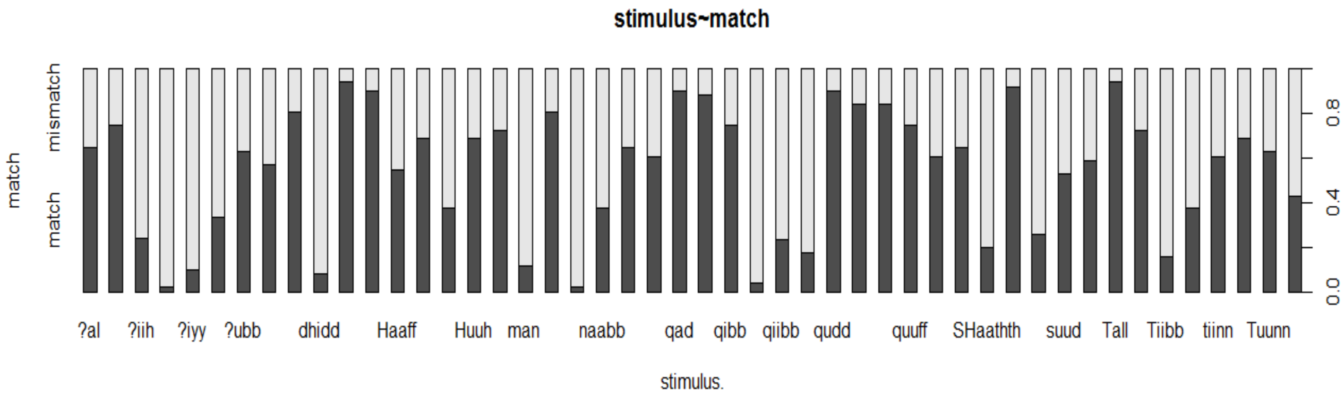


Figure 6-9: Stimulus~match

1. Null model:

m0.null<- glmer(match~ 1 + (1|listener)+ (1|stimulus), data = SBAAW, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1) (where SBAAW refers to the SB audio and audio-written data)

Random effects:

Groups	Name	Variance	Std.Dev.
Listener	(Intercept)	***0.6914***	0.8315
Stimulus	(Intercept)	***2.9405***	1.7148
Number of obs: 2447, groups: listener, 51; stimulus, 48			

Table 6-30: Random effects table of null model

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.1617	0.3704	0.436	0.6625
ListgpTA	-0.6757	0.3916	-1.726	**0.0844 .**
ListgpTQ	-0.4538	0.3118	-1.455	0.1456
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Table 6-31: Listgp basic model output summary table

Observations based on figures 6-8 , 6-9 and table 6-30:

1. Since the variance values of listener and stimulus are >0, both random effects must be part of the maximal model, and the overall variance of the outcome variable would be attributed to both effects.
2. The between-listener (within-stimulus) variance intercept in match is estimated as 0.6917, and the between-stimulus (within-listener) variance intercept is estimated as 2.490. Hence, the total variance is 0.6914+2.9405= 3.6319. The variance partition coefficient, VPC for listener is 0.6917/3.6319 = 0.1904, which indicates that 19.04% of the variance in match can be attributed to differences among listeners (i.,e slopes, namely Listgp and age). On the other hand, the VPC for stimulus is 2.490/3.6319 = 0.8096 indicating that almost 80.96% of the variance in match can be attributed to differences among stimulus tokens (length, context, frequency, vowel quality and stimulus presentation).
3. The variability value of stimulus is larger than that of listener; 2.490 compared to 0.6917. This might be probably due to the fact that the number of stimulus units (48 items) is less than those of the listener units (51 listeners of 2448 observations).

Now that we have established that the model will have slopes, the last step before constructing the maximal model would be to determine which slopes to be included. Baayen (2008, p. 290) states that “in general, predictors tied to subjects (age, sex, handedness, education level, etc.) may require by-item random slopes, and predictors related to items (frequency, length, number of neighbors, etc) may require by-subject random slopes.”

Based on the above definition by Baayen (2008), by listener variables such as Listgp and age are ***between listener*** and ***within-stimulus***. This means that they are slopes in the stimulus random structure (Listgp+age|stimulus) and do not vary across stimulus. However, a listener can belong to one Listgp but not the other or more than one, can be of a certain age but not another at the same time.

In the same vein, by-stimulus variables, namely context, length, frequency (freq.), stimulus presentation and vowel quality are ***between stimulus*** and ***within-listener***. That is, they are slopes in the listener random structure (context+length+freq.+vowel.quality+ stimulus.presentation|listener) and do not vary across listeners. At the same time, a stimulus token can belong to a certain level but not both at the same time. For instance, a token can belong to the emphatic, plain, pharyngeal or q context but not more than one at the same time. Similarly, a token can be either short or long but not both, a token can have a certain frequency reading but not more at the same time, a token can belong to either the audio condition or audio-written condition but not both at the same time and a stimulus token can include as its nucleus either one of the three vowel types a, i or u but not more at the same time. This is the same procedure that was followed in chapters 4 and 5.

One point to be mentioned before moving to the maximal model in the next section is the significance values of the basic model Listgp in table 6-31. We can already see that the TA group has a near significant value (0.08) compared to the T group (0.6) and the TQ group (0.14). This is an early indication of the listener groups' effects on the categorizations.

In the next subsection, the maximal model is constructed and the model simplification points are discussed so as to avoid non-convergence issues (Barr et al, (2013).

6.4.3.2 Regression models

A. <u>Maximal model</u>

The maximal model in light of Barr et al (2013) can be defined as a derived model that has the variables of experimental interest (according to the hypotheses), control variables and all possible interactions and a maximal random effects structure that includes both random effects units and their slopes. The maximal model of the SB- audio and audio-written data is presented in 2. below along with its summary output table in Table 6-32.

4. Maximal model:

stpmodel<-glmer(match~Listgp + context + length + freq. + vowel.quality + stimulus.present ation + age +Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality + Listgp:stimu lus.presentation+ context:length + context:freq. + context:vowel.quality + length:freq. + age :vowel.quality + freq.:vowel.quality + (Listgp|stimulus) + (stimulus.presentation|listener) , d ata = SBAAW , family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ = 1)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.023286	0.703470	-0.033	0.973593	
ListgpTA	0.387843	0.526329	0.737	0.461193	
ListgpTQ	0.272108	0.505923	0.538	0.590683	
contextpharyngeal	-0.227506	0.696490	-0.327	0.743935	
contextplain	0.459979	0.635811	0.723	0.469402	
contextq	-0.261577	0.677829	-0.386	0.699568	
lengthshort	-2.680630	0.626953	-4.276	1.91e-05	***
freq.	0.067751	0.063522	1.067	0.286163	
vowel.qualityi	0.540797	0.851753	0.635	0.525479	
vowel.qualityu	-0.254307	0.772653	-0.329	0.742054	
stimulus.presentationaudio+written	-0.716525	0.264663	-2.707	0.006783	**
age	0.002028	0.013494	0.150	0.880511	
ListgpTA:lengthshort	0.704622	0.392403	1.796	0.072549	.
ListgpTQ:lengthshort	-0.142910	0.385067	-0.371	0.710541	
ListgpTA:contextpharyngeal	-0.436963	0.457631	-0.955	0.339661	
ListgpTQ:contextpharyngeal	0.031859	0.463020	0.069	0.945143	
ListgpTA:contextplain	0.448176	0.468256	0.957	0.338508	
ListgpTQ:contextplain	0.340396	0.451120	0.755	0.450514	
ListgpTA:contextq	0.355460	0.462224	0.769	0.441880	
ListgpTQ:contextq	0.066240	0.467190	0.142	0.887250	
ListgpTA:freq.	-0.017687	0.012823	-1.379	0.167812	
ListgpTQ:freq.	-0.016326	0.012645	-1.291	0.196681	
ListgpTA:vowel.qualityi	-2.446858	0.410670	-5.958	2.55e-09	***
ListgpTQ:vowel.qualityi	-0.511971	0.417485	-1.226	0.220078	
ListgpTA:vowel.qualityu	-1.020602	0.396408	-2.575	0.010035	*
ListgpTQ:vowel.qualityu	-0.048221	0.392811	-0.123	0.902298	
ListgpTA:stimulus.presentationaudio+written	-0.255395	0.370459	-0.689	0.490571	
ListgpTQ:stimulus.presentationaudio+written	-0.784534	0.358953	-2.186	0.028844	*
contextpharyngeal:lengthshort	3.611735	1.241036	2.910	0.003611	**
contextplain:lengthshort	4.598890	0.789268	5.827	5.65e-09	***
contextq:lengthshort	1.057497	0.866090	1.221	0.222085	
contextpharyngeal:freq.	-0.123480	0.067098	-1.840	0.065726	.
contextplain:freq.	-0.163716	0.028057	-5.835	5.37e-09	***
contextq:freq.	-0.041509	0.026367	-1.574	0.115427	

contextpharyngeal:vowel.qualityi	1.637780	0.954088	1.717	0.086054	.
contextplain:vowel.qualityi	-1.518703	0.851683	-1.783	0.074557	.
contextq:vowel.qualityi	1.227316	0.957468	1.282	0.199900	
contextpharyngeal:vowel.qualityu	-0.309369	0.927714	-0.333	0.738776	
contextplain:vowel.qualityu	-0.655663	0.847096	-0.774	0.438924	
contextq:vowel.qualityu	-0.760498	0.924589	-0.823	0.410777	
lengthshort:freq.	-0.035142	0.064587	-0.544	0.586373	
vowel.qualityi:age	0.040325	0.012794	3.152	0.001623	**
vowel.qualityu:age	0.007222	0.011335	0.637	0.524041	
freq.:vowel.qualityi	0.050492	0.025581	1.974	0.048406	*
freq.:vowel.qualityu	0.154024	0.043596	3.533	0.000411	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 6-32: maximal model's output summary table

As can be seen from table 6-32, two variables were found significant, namely length (short) and stimulus.presentation (audio-written) in addition to eight interactions of which ***Listgp:stimulus.presentation*** was found significant with p-value≅0.03 for the TQ group in the audio-written condition.

As we did in chapters 4 and 5, we examine below the reduced model baseListgp in table 6-33 where the coefficients of listener groups are inspected for significance. We notice that the TA group exhibits a near significant p-value of 0.08 compared to an almost near significant p-value of 0.14 for the TQ group and a non-significant value for the T group at p≅0.7 in line with the observations in the raw data in 6.3.1.

baseListgp<- glmer(match~ Listgp + (Listgp|stimulus) + (1|listener), data = SBAAW, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.1617	0.3704	0.437	0.6624
ListgpTA	-0.6757	0.3916	-1.726	0.0844 .
ListgpTQ	-0.4538	0.3119	-1.455	0.1456
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				

Table 6-33: summary table of reduced table baseListgp

B. <u>Maximal model simplification techniques:</u>

A number of techniques were followed in order for the maximal model to converge. These include the following points.

1. Only variables of interest were derived from the research hypotheses along with the control variable and all possible interactions whereas other variables were eliminated if not within the scope of the work or if not reflecting variability.

2. The random effects structure was simplified where the two variables of the Listgp*stimulus.presentation only were included as slopes. This was done to overcome non-converge issues in R (Barr et al. (2013)), hence (Listgp|stimulus) and (stimulus.presentation|listener).
3. Theoretical assumptions such as collinearity were adhered to. For example, the variable ***nature*** (real or nonsense) was dropped since it is collinear with ***freq.***
4. The number of iterations of the model was raised to 2e5; i.e., control=glmerControl(optCtrl=list(maxfun=2e5)) and nAGQ =1).
5. The continuous predictors of age and freq. were scaled and centered[32] and integrated into the model.

C. Fitting regression models

Three regression models were fitted using stepwise backward logistic regression using *dropterm* and *update* features in the MASS package (Venables & Ripley, 2003). This entails deleting a single interaction or variable at a time when that interaction or variable did not reach the threshold of significance (p=<0.05). In addition, a fourth regression model was fitted in step_4 after adding the interaction ***length:vowel.quality***. The summary output tables of the fitted regression models are provided in Appendix 6-2.

D. Model's results' interpretation

The final model's output summary is given in table 6-34 where significant variables and interactions are indicated with an asterisk (*) in the P-value column.

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.116582	0.680685	-0.171	0.864011
ListgpTA	0.325316	0.456759	0.712	0.476325
ListgpTQ	0.250002	0.411839	0.607	0.543825
contextpharyngeal	-0.126394	0.641918	-0.197	0.843906
contextplain	0.791864	0.631386	1.254	0.209781
contextq	-0.069078	0.645544	-0.107	0.914784
lengthshort	-3.867144	1.008391	-3.835	0.000126 ***
freq.	0.060855	0.028433	2.140	0.032333 *
vowel.qualityi	0.987745	0.803557	1.229	0.218990
vowel.qualityu	-0.390851	0.747542	-0.523	0.601080
stimulus.presentationaudio+written	-0.641654	0.255015	-2.516	0.011865 *
age	0.002404	0.013349	0.180	0.857095

[32] Centering a continuous variable entails selecting a number at which interpreting the intercept is meaningful. Often in R, users are prompted to scale and center continuous variables with warning messages.

ListgpTA:lengthshort	0.383944	0.345320	1.112	0.266203	
ListgpTQ:lengthshort	-0.351230	0.325415	-1.079	0.280442	
ListgpTA:vowel.qualityi	**-2.325697**	**0.431281**	**-5.393**	**6.95e-08**	*******
ListgpTQ:vowel.qualityi	-0.488148	0.415100	-1.176	0.239604	
ListgpTA:vowel.qualityu	**-0.920474**	**0.411007**	**-2.240**	**0.025119**	*****
ListgpTQ:vowel.qualityu	0.008727	0.385647	0.023	0.981947	
ListgpTA:stimulus.presentationaudio+written	-0.177953	0.383631	-0.464	0.642744	
ListgpTQ:stimulus.presentationaudio+written	**-0.669892**	**0.352482**	**-1.901**	**0.057367**	**.**
contextpharyngeal:lengthshort	**4.180992**	**1.164568**	**3.590**	**0.000330**	*******
contextplain:lengthshort	**4.738237**	**0.775841**	**6.107**	**1.01e-09**	*******
contextq:lengthshort	**1.419348**	**0.857238**	**1.656**	**0.097778**	**.**
contextpharyngeal:freq.	**-0.131919**	**0.061037**	**-2.161**	**0.030672**	*****
contextplain:freq.	**-0.145273**	**0.025420**	**-5.715**	**1.10e-08**	*******
contextq:freq.	**-0.061919**	**0.027582**	**-2.245**	**0.024777**	*****
contextpharyngeal:vowel.qualityi	1.327294	0.914976	1.451	0.146882	
contextplain:vowel.qualityi	**-2.103533**	**0.866679**	**-2.427**	**0.015219**	*****
contextq:vowel.qualityi	1.022597	0.929646	1.100	0.271338	
contextpharyngeal:vowel.qualityu	-0.553982	0.865729	-0.640	0.522236	
contextplain:vowel.qualityu	-0.907525	0.847620	-1.071	0.284316	
contextq:vowel.qualityu	-0.973517	0.884063	-1.101	0.270816	
lengthshort:vowel.qualityi	0.288984	0.864819	0.334	0.738263	
lengthshort:vowel.qualityu	**1.707455**	**0.849639**	**2.010**	**0.044471**	*****
vowel.qualityi:age	**0.040462**	**0.012811**	**3.158**	**0.001586**	******
vowel.qualityu:age	0.007409	0.011328	0.654	0.513063	
freq.:vowel.qualityi	0.019460	0.033504	0.581	0.561345	
freq.:vowel.qualityu	**0.108146**	**0.046910**	**2.305**	**0.021144**	*****
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Table 6-34: Output summary table of step_4

As can be seen from table 6-34, eight interactions are reported as being significant whereas the ***Listgp*** interaction with ***stimulus.presentation*** is shown as having a near significant value. Thus, we inspect the allEffects plot created using the effects package (Fox, 2003) in figure 6-10 so as to display all interactions of effect on match. Combining the results from table 6-34 and figure 6-10, we end up with three significant variables along with a total of eight interactions. The significant variables are ***stimulus.presentation*** at the audio-written level, ***length*** at the short level and frequency whereas the interactions are ***Listgp:stimulus.presentation*** (near significant), ***Listgp:vowel.quality***, ***context:length***, ***context:freq.,*** ***context:vowel.quality***, ***length:vowel.quality***, ***vowel.quality:age*** and ***freq.:vowel.quality***. For simplicity purposes, only significant or near significant interactions are interpreted but not variables (***length***, f*requency* and ***stimulus.presentation***) since they are already reflected in the interactions.

In the effect displays below, the x-axis represents one of the the two independent variables of an interaction and the levels of the second variable shown in two panels. Moreover, the y-axis represents the probability of mismatched responses on a scale of 0 to 1 with 0 signifying a match (of the observed responses to the corpus mappings) and 1 signifying a mismatch. Degrees of match are thus said to be higher if being between 0 and 0.5 and lower if being between 0.5 and 1 given the contrasts coding of match (0) and mismatch (1). This is not to be confused with the contrasts coding used in chapters four and five where match=1 and mismatch=0.

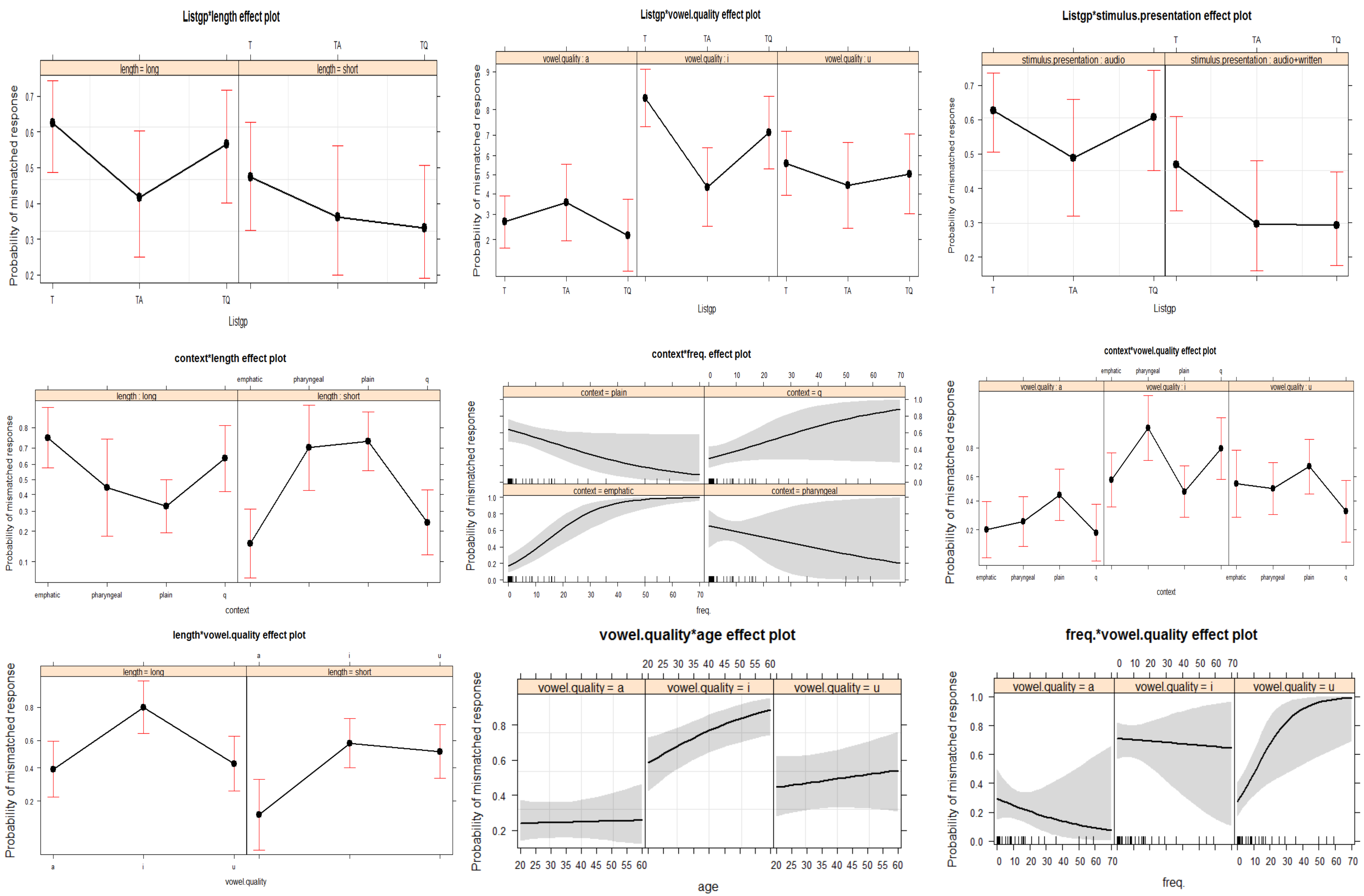


Figure 6-10: Figure 6-10: effect displays for all the interactions in model step_4

Figure 6-10 mainly shows that seeing the orthography 'helps' i.e. leads to mappings that more closely match the corpus - for both TA and TQ, though somewhat more for TA in some conditions (e.g. with [i]). The first interaction to be interpreted is the one of main theoretical interest, the ***Listgp*** interaction with ***stimulus.presentation*** in their effects on match. This is displayed in figures 6-10 and figure 6-11 where the simple effects of ***Listgp*** are not the same at the different levels of ***stimulus.presentation***. Noteworthy to mention here is that the mismatched responses are compared to observed patterns in the loanword corpus in chapter three.

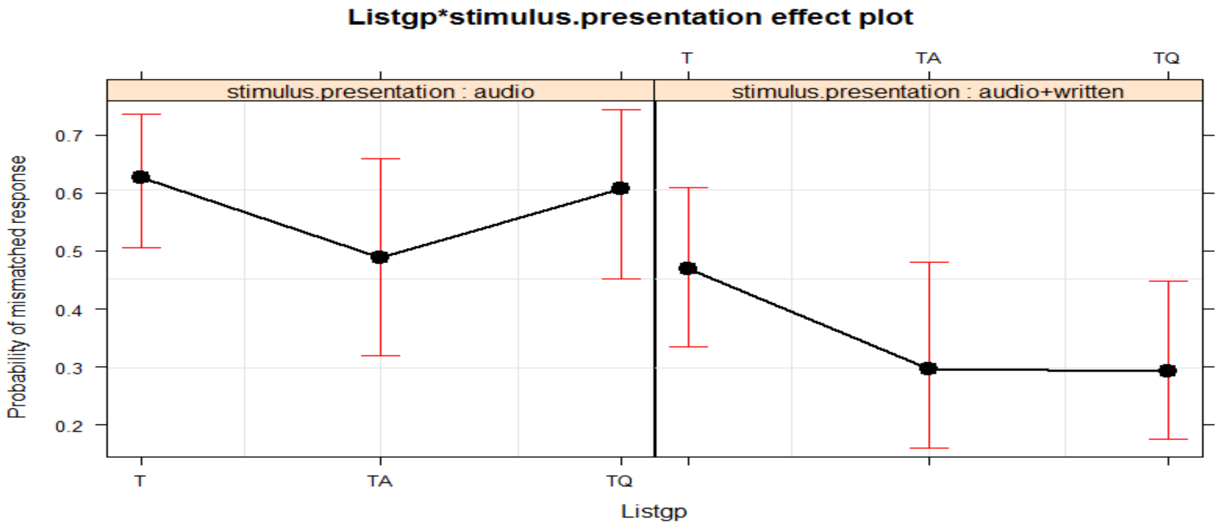


*Figure 6-11:Listgp*stimulus.presentation effect display, match=0, mismatch=1*

Figure 6-11 clearly shows that the probability of mismatched response for the three groups is lower in the audio-written condition than it is in the audio condition. This is especially significant for the TA and TQ groups (ones with knowledge of Arabic) in the audio-written condition whose results are shown in figure 6-11 as close to each other. In the audio condition, the TA group yielded the highest degrees of match among the three, followed by the TQ and finally the T group.

Moreover, we perceive from table 6-34 that the interaction ***Listgp: stimulus.presentation*** is near significant at a p-value of 0.05 for the TQ group in the audio-written level. Thus, as shown in figure 6-10, the TA and TQ achieved higher degrees of match than the T group when the stimulus material was both heard and written. This confirms the hypothesis that orthography plays a role in matching responses to corpus mappings.

The second interaction of ***Listgp:vowel.quality*** reveals how listener groups matched the three Arabic vowels as in figure 6-12 below. As can be visualized, the three listener groups tend to match monosyllabic words whose nucleus is the vowel /a/ with higher degrees of match (below 0.5 of mismatch), followed by the vowel /u/ (in the 50-50 range) and finally the vowel /i/ (0.5 and above mismatch). This is the same result derived from the raw data in 6.3.1. As for the simple effects of the ***Listgp*** variable, when the stimulus vowel is /a/, the TQ group displayed the highest degrees of match, followed by the T (with closer values to the TQ group's) and finally the TA group. This is reflected in the fact that the TA group was the only one that assimilated the vowel [a:] as /e/$_{T}$ instead of the predicted /a/$_{T}$. When the vowel is /u/, the TA reflected the highest degrees of match followed by the TQ and finally the T group that mismatched [u]$_{ʕ}$ to /o/, with the results of the TQ and T groups being close to each other. When the vowel is /i/, the TA group manifested the highest degrees of match (less than 0.5) (mismatched only one /i/ category), the TQ group reflected lower degrees of match (around 0.7) (mismatched [i:]$_{q}$>/ɯ/$_{T}$ and half the TQ participants mismatched [i:]$_{ʕ}$>/ɯ/$_{T}$) and the T group the lowest (around 0.85) (mismatched [i] to/ɯ/$_{T}$ and [i]$_{ʕ}$ to /e/$_{T}$).

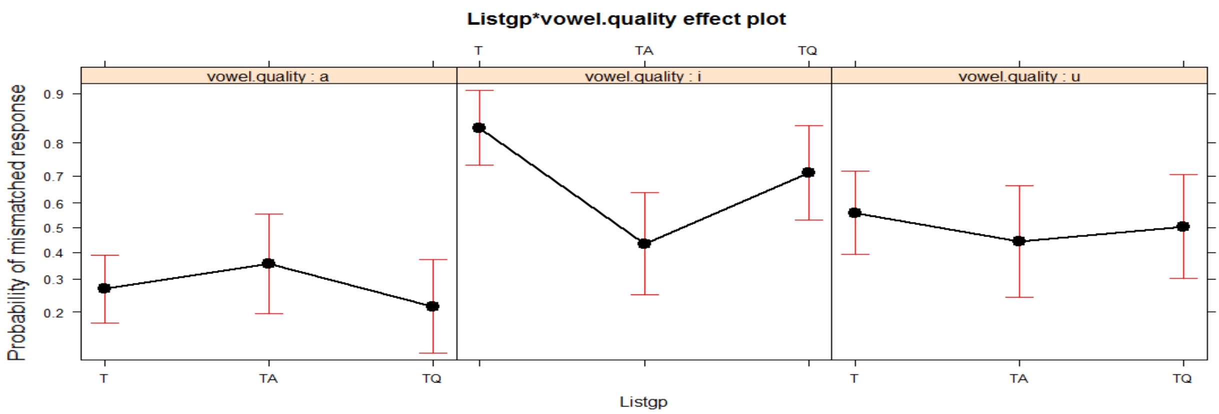


*Figure 6-12: Listgp*vowel.quality effect display, match=0 and mismatch=1*

The next three interactions pertain to the effects of context when interacting with ***vowel.quality***, ***length*** and ***frequency*** in their effects on match in figures 6-13, 6-14 and 6-15 respectively. Regarding the interaction ***context:vowel.quality***, overall, listeners tend to achieve the highest degrees of match for the vowel /a/, followed by /u/ and finally /i/ (figure 6-13) regardless of the different levels of context. However, we should not downplay the simple effects of context since listeners tend to yield different matching orders of the context levels for the three vowel qualities. Hence, when the stimulus vowel is /a/, listeners match the q level the highest, followed by emphatic, pharyngeal and finally the plain level. This means that they incur higher degrees of match for gutturals than for plain consonants. This is probably due to the mismatched mapping of [a:] to /e/$_{T}$ by the TA group with an average goodness of fit of 54.54% (see table 6-9).

As for the vowel /i/, listeners tend to achieve higher degrees of match in the order of plain, emphatic, q (e.g. TQ group perceiving [i:]$_{q}$ as /ɯ/$_{T}$) and pharyngeal (e.g. T group mismatching [i]$_{\text{ʕ}}$ to /e/$_{T}$ and half TQ participants assimilating [i:]$_{\text{ʕ}}$ as /ɯ/$_{T}$; higher degrees of match for plain environment (≅0.5) than guttural (>0.5). The matching patterns for the vowel /u/ seem to be closer to those of the vowel /a/ since listeners tend to reflect higher degrees of match in the order q, pharyngeal, emphatic and plain (since all listeners mismatched [u] to /u/$_{T}$; higher degrees of match for gutturals environment than plain.

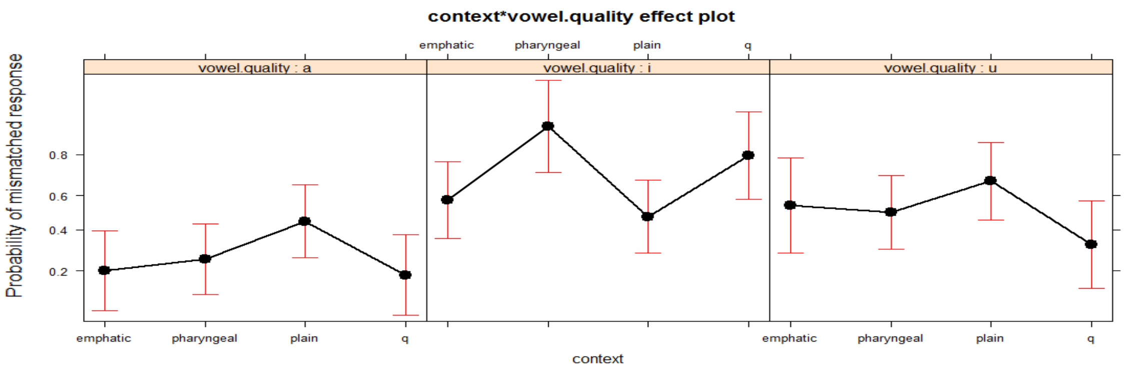


*Figure 6-13: context*vowel.quality effect display; match=0, mismatch=1*

As for the interaction of ***context and length***, it would be misleading to report the main effect of the variable length. Figure 6-14 shows that when the stimulus vowel is short, listeners tend to exhibit different degrees of match for the stimulus context. The order of match contexts from highest to lowest is emphatic, q, pharyngeal and then plain. In other words, listeners tend to exhibit higher degrees of match for the guttural stimulus when the nucleus vowel is short (e.g. mismatching [i] to/ɯ/$_T$ by the T and TA groups and [u] to /u/$_T$ by all three groups). On the other hand, when the stimulus nucleus is long, the order of the matched context is reversed; i.e., plain, pharyngeal, q and emphatic. In other words, listeners tend to achieve higher degrees of match for the plain consonants than for the guttural ones when the stimulus vowel is long (e.g. mismatching one long plain vowel in [a:] to /e/$_T$ by the TA group compared to two long guttural vowels in [i:]$_q$ as /ɯ/$_T$ by the TQ group and [i:]$_ʕ$ as /ɯ/$_T$ by half TQ participants).

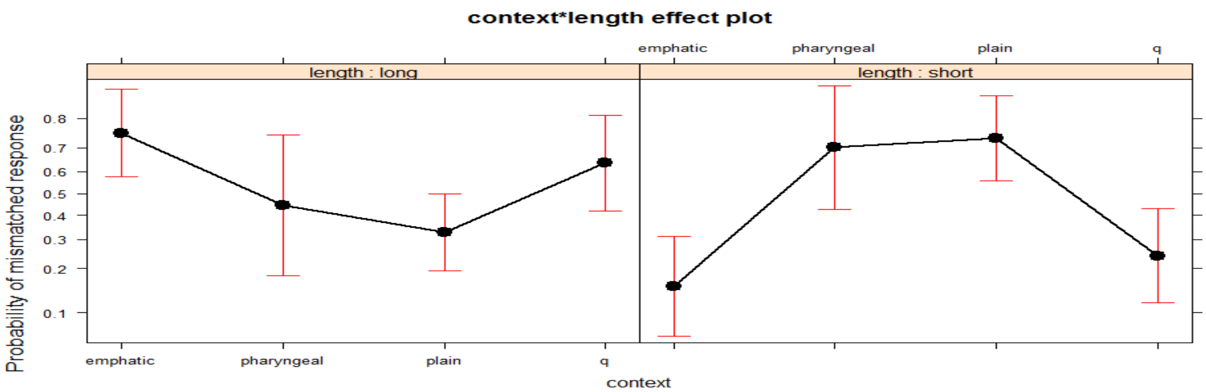


*Figure 6-14: context*length effect display; match=0, mismatch=1*

Regarding the interaction ***context:freq.***, two opposite patterns emerge as shown in figure 6-15. For high-frequency words, listeners tend to display higher degrees of match for stimulus words with plain (e.g. *suud* 'black': 16.07/100,000 words) and pharyngeal consonants as their onsets (with plain as highest followed by pharyngeals) (e.g. *ʔal* 'maybe': 20.98). On the other hand, they tend to achieve lower degrees of match for real words with emphatic consonants (e.g. Tall 'dew': 35.89) and q (in order) (e.g. qatt 'fodder'). As for nonsense words (words with 0 frequency) or words with low frequency, the scenario is reversed, i.e., listeners trigger higher match rates for words with emphatic (e.g. dhurr 'harm': 0.98) and q consonants (qaatt 'qat/kat'; plant: 0.16) and lower rates for pharyngeal (e.g. Huuh:0/100,000 words) and plain ones (zirr 'button': 1.15). This may suggest that listeners in this context do not perceive the emphatic, q and pharyngeal as a group versus plain consonants.

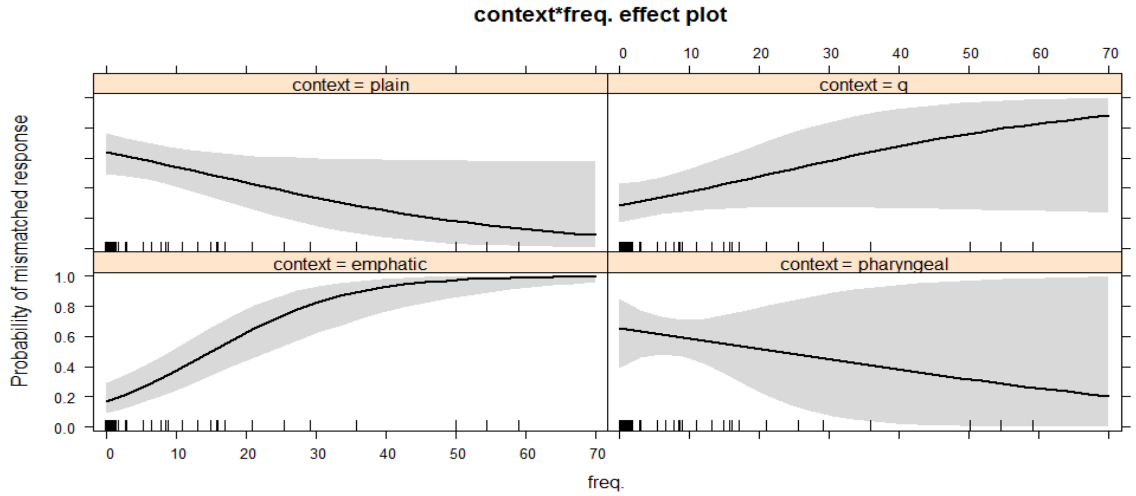


*Figure 6-15: context*frequency effect display, mismatch=1, match=0*

Figure 6-16 illustrates the interaction of ***vowel.quality*** with ***age*** in their effect on match which is reflected in the non-parallel lines across the panes. Regardless of age, listeners tend to achieve the highest match rates for words with the vowel /a/, followed by vowel /u/ and finally vowel /i/. This is a scenario where the interaction is overriding the main effect of a variable, age, and is more meaningful (Martin, 2014). Figure 6-16 also shows that there is a clear decrease of match for the vowel /i/ as age increases and a slight decrease for /u/. This is reflected in the significant interaction of **vowel.qualityi:age** in table 6-34.

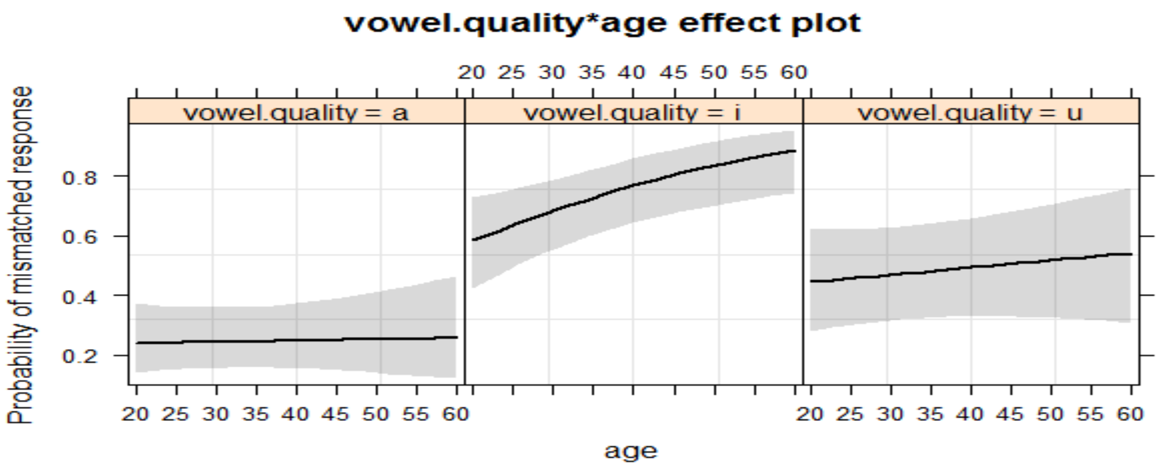


*Figure 6-16: age*vowel.quality effect display, match=0, mismatch=1*

Similar to the interaction effect in figure 6-16, the interaction ***freq.:vowel.quality*** is more meaningful than the main effect of freq. (figure 6-17). Listeners tend to demonstrate the highest degrees of match (<0.5) for high-frequency words when the stimulus vowel is /a/, lower degrees of match when the vowel is /i/ (>0.5) and the lowest when the vowel is /u/. On the other hand, when the words are nonsense (0 frequency) or are of low frequency, they tend to reflect higher degrees of match for words with the vowels /a/ and /u/ and lower degrees for the vowel /i/. In short, in the case of the vowels [a]/[i] there is a (very) slight increase of match as freq. increases, whereas for [u] there is a clear decrease of match as freq. increases. This is reflected in the significant interaction **frequency:vowel.quality** at the level u in table 6-34. An example of this is [u]$_{ʕ}$>/o/$_{T}$ (predicted /u/$_{T}$) which is mismatched by the T group.

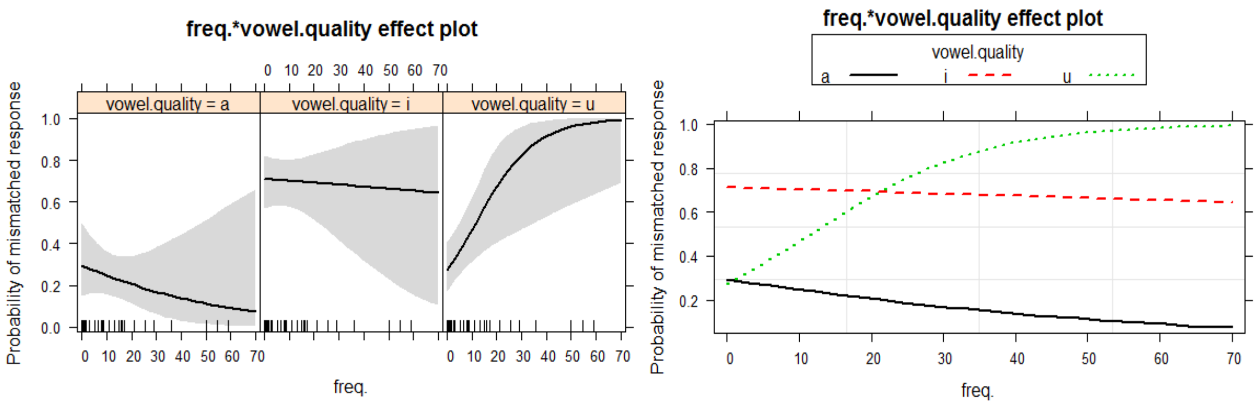


*Figure 6-17: frequency*vowel.quality effect plot, match=0, mismatch=1*

The final interaction to be interpreted is the one between ***length and vowel quality*** as presented in figure 6-18. According to the figure, when the stimulus vowel is long, listeners tend to display higher degrees of match (approximately 0.4) for the long vowels /a/ and /u/ (approximately 0.4) but lower degrees of match (approximately 0.8) for the long vowel /i/. When the stimulus vowel is short, listeners tend to reflect the highest degrees of match for the vowel /a/ and lower degrees of match for the vowels /i/ and /u/, respectively. An example of this is mismatching [u] as /u/$_{T}$ instead of /y/$_{T}$ by all listener groups which is reflected in the significant interaction **lengthshort:vowel.qualityu** in table 6-34.

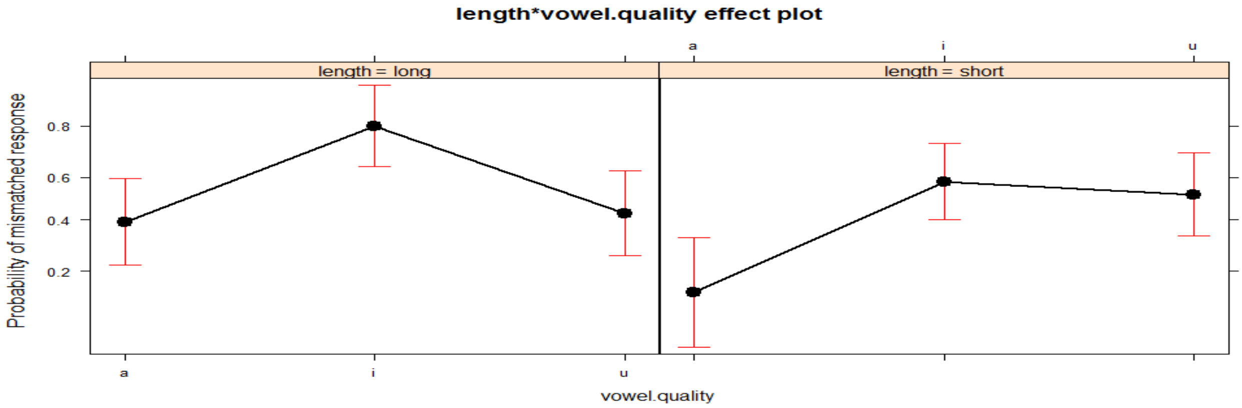


*Figure 6-18: length*vowel.quality effect display, match=0, mismatch=1*

6.1.1.1. Model validation

In this part, we validate the goodness of fit of the final model by examining the residuals plots for homoscedasticity and normality in the same way we did in chapters four and five. In the first plot in figure 6-20, the residuals can be seen as not forming patterns on the positive and negative areas which indicates that they are heteroscedastic, i.e. there is variability. Furthermore, the second plot in figure 6-20 illustrates an almost normal distribution of the best fitting model's residuals, with few outliers at both ends.

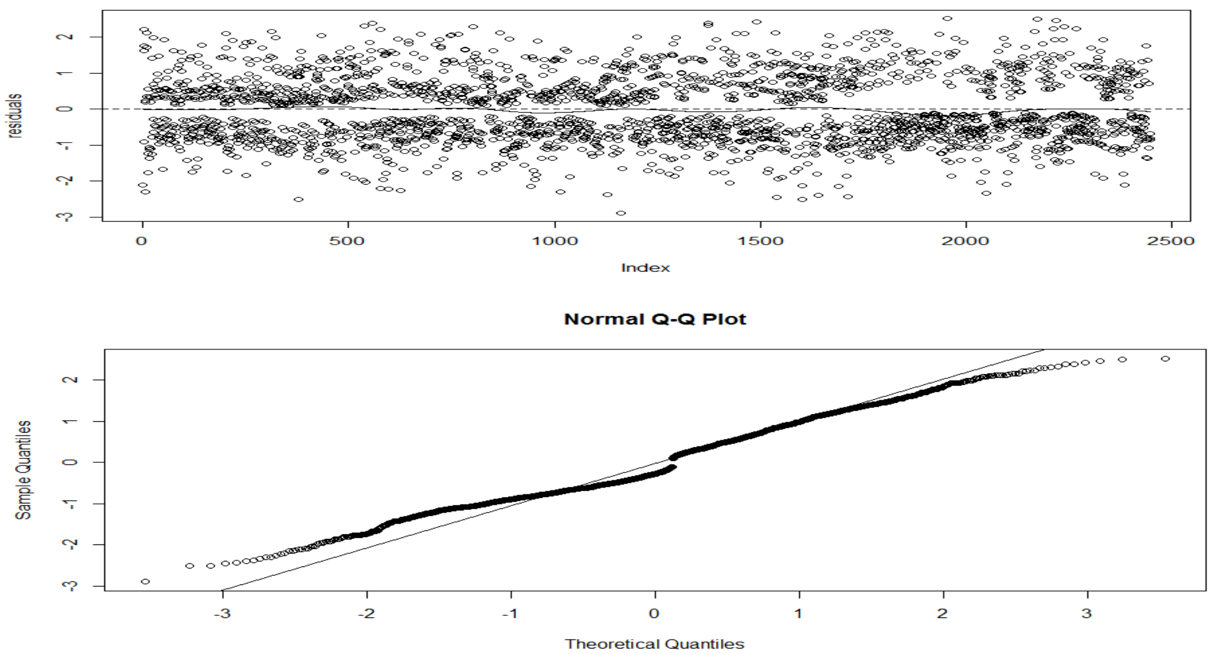


Figure 6-19: model validation test

A maximal generalized linear mixed effects modelling (GLMM) analysis was performed using R (Team, 2015) and lme4 (Bates et al., 2015) to predict the relationship between match (DV) and the six independent variables of listener group ***Listgp***, consonants' ***context***, stimulus ***length***, stimulus frequency ***freq., stimulus.presentation*** *and* ***vowel.quality***. The fixed effects structure included all six exploratory variables, one control variable, ***age*** as well as interactions reflecting variation with ***Listgp:stimulus.presentation*** as the interaction of main theoretical interest. The random effects structure included both intercepts and slopes both for listener and stimulus.

The theoretical assumptions of homoscedasticity and normality were observed and checked using residual plotting. The significance level (using chi-square test) with α-level of $p>0.05$ was adopted in the model selection step. In addition, confidence intervals at 95% (Barr et al. (2013)) were reported and mirrored in plots.

It was found that the probability of Turkish listeners matching assimilation patterns to ones predicted from the ALT corpus is dependent on both variables and interaction. The significant variables include ***length*** and ***stimulus presentation*** whereas significant interaction effects include ***Listgp:stimulus.presentation***, ***Listgp:vowel.quality*, context:freq.**, **context:length**, ***context:vowel.quality***, ***length:vowel.quality***, ***vowel.quality:age*** and ***vowel.quality*:*freq***.

The findings thus far reveal that the TA and TQ groups tend to trigger higher degrees of match than the T group both in the audio and audio-written condition. They also manifested the highest degree of match in the audio-written condition than they did in the audio condition.

Regarding ***context*** and its interaction with ***vowel.quality***, ***length*** and ***frequency***, listeners' performance was gauged in guttural and plain environments. The results suggest that listeners tend to reflect higher degrees of match for the vowel /a/, followed by /u/ and then /i/. Moreover, when the vowel quality is either /a/ or /u/, listeners perceived words with gutturals with higher degrees of match than they did for the plain consonants but they responded with higher degrees of match with plain consonants when the vowel involved was /i/. This result suggests that the listeners were sensitive to the residual effects of gutturals (context) and vowel quality.

Regarding the interaction of ***context and length***, listeners incurred higher degrees of match when the stimulus vowel was short in the guttural context whereas they exhibited higher degrees of match when the stimulus vowel was long in the plain context compared to the guttural one. This interaction too suggests that the listeners were sensitive to the residual effects of gutturals. The interaction ***context and freq.*** was not as clear-cut as the two previous interactions of ***context*** with ***vowel.quality*** and ***length***. The findings suggest that listeners perceive real and nonsense words differently depending on the stimulus consonant context.

For real words with high frequency, listeners incurred higher degrees of match when the stimulus consonant was either plain or pharyngeal but lower degrees of match with emphatic consonants and q. Conversely, for nonsense words, i.e., words with zero frequency or low frequency, listeners rendered higher match probability when the stimulus context included either an emphatic or a q consonant. They, however, yielded higher mismatched responses for nonsense words when the stimulus consonant was either a pharyngeal or a plain consonant. Thus, we may assume that the Turkish listeners in the experiment treated the members of the guttural class differently depending on the ***frequency*** of the stimulus words.

An alternative explanation is that the Turkish speakers (TA and TQ in this case) are not sensitive to frequency of words in Arabic, but rather that the speaker (who is a native speaker of Arabic) is sensitive to frequency. Therefore, the stimulus words are produced differently depending on whether they are real and frequent or infrequent or nonwords. Regardless of the explanation, frequency effect is detected in the responses of the TA and TQ groups.

Length was found to vary across vowel.quality which had an effect on the degrees of matched responses. When the stimulus vowel was short, listeners displayed the highest degrees of match for the vowel /a/ but low degrees of match for the vowels /i/ and /u/, in order. On the other hand, when the stimulus vowel was long, they exhibited higher degrees of match for the vowels /a/ and /u/; however, they incurred lower degrees of match for the vowel /i/. The significant interaction **lengthshort:vowelqualityu** is important in understanding the mismatched pattern [u]>/u/$_T$.

The last two interactions reveal listeners behaviour when ***vowel quality*** interacts with ***age*** and ***frequency***. First, listeners showed higher degrees of match for the vowel /a/, followed by /u/ and finally /i/ regardless of their age. As for the interaction of ***freq.*** with ***vowel.quality***, listeners' perception of the three vowels differed in real words with high frequency compared to nonsense words. They tended to render more accurate responses for the vowel /a/ in high-frequency words, followed by /i/ but not /u/. As for nonsense words, listeners tended to yield more accurate responses for the vowel/u/, then /a/ but not /i/.

All in all, the take home message here is that the TA do better, but there is an effect of orthography for both TA and TQ. In the next section, a logistic regression analysis of the Simulated Borrowing written data is presented in order to compare how the TA and TQ groups perceive Arabic orthography.

6.5 Logistic regression of SB-written stimuli

As was mentioned in section 6.1, the rationale for running the logistic regression on the SB-written dataset is to model the type of Arabic knowledge the Ottomans had be it spoken, written and/or for religious purposes. Thus, two hypotheses would be driving the analysis; either the TA group would trigger higher degrees of match to the corpus and thus we could make the claim that the Ottomans most probably possessed both written and spoken knowledge of Arabic. Alternatively, if the TQ group render higher degrees of match, then we could assume that the Arabic knowledge of the Ottomans could have been mostly in the written form and for religious purposes. It is expected that the TA group would yield more accurate responses based on their performance in the SB audio and audio-written dataset.

In the SB-written dataset, the sample consisted of 792 observations of 33 male and female participants: 11 TA and 22 TQ who were presented with 24 Arabic written words with long and short words embedded in four different consonantal contexts; emphatic, pharyngeal, plain and q. The frequency of the stimuli words ranged from 0 hz to 72.25 hz. The respondents ranged in age from 17 to 62. Table 6-35 summaries these variables.

Hence, the variables involved as shown in table 6-30 include the following (*italicized* and **bold-faced**) which are the same ones used in analyzing the SB-audio dataset.

Dependent variable: ***match*** with the levels match, mismatch (and NA[33])

By *listener* variables: ***Listgp*** (T, TA and TQ) and ***age*** (17:62)

By-*stimulus* variables: ***context*** (emphatic, pharyngeal, plain and q), ***length*** (long, short), ***freq.*** (0:72.25), and ***vowel.quality*** (a, i and u)

Listener	Listgp	Age	vowel.quality	st.vowel
TA1 : 24	TA:264	Min. :17.00	a:264	a :132
TA10 : 24	TQ:528	1st Qu.:24.00	i:264	aa:132
TA11 : 24		Median :28.00	u:264	i :132
TA2 : 24		Mean :31.18		ii:132
TA3 : 24		3rd Qu.:35.00		u :132
TA4 : 24		Max. :62.00		uu:132
(Other):648				
Stimulus	**match**	**Length**	**Context**	**freq.**
?aall : 33	match :384	long :396	emphatic :198	Min. : 0.0000
?ayy : 33	mismatch:407	short:396	pharyngeal:198	1st Qu.: 0.5375
?ihh : 33	NA's : 1		plain :198	Median : 2.7950
?iis : 33			q :198	Mean : 8.9817
?uth : 33				3rd Qu.: 6.9300
barr : 33				Max. :72.2500
(Other):594				

Table 6-35: descriptive statistics of the SB written dataset

6.5.1 Protocol for model selection

The protocol used here is the same one followed in chapters 4, 5 and in section 6.4.3. The steps have been summarized so as to avoid repetition.

6.5.1.1 Defining the fixed and random effects structures

Six fixed effects were included, namely ***Listgp***, ***context***, ***length***, ***freq.***, ***vowel.quality*** and age as a control variable (figure 6-20) along with ten interactions which show variability (appendix 6-3), all of which appear in the SBAAW dataset logistic regression analysis. These include

[33] In rare cases, NA is retained when the other responses of the participants who did not supply an answer were coherent; otherwise null responses were discarded.

Listgp:context, ***Listgp:length***, ***Listgp:vowel.quality***, ***Listgp:freq.***, ***context:length***, ***context:vowel.quality***, ***context:freq.***, ***length:freq***., ***freq.:vowel.quality*** and ***age:vowel.quality***. The interaction ***length:vowel.quality*** was dropped from the maximal model since it did not show variability, however, it was later fitted in the final model as was done in the previous chapters. Moreover, all interactions with the control variable age were eliminated so as to simplify the maximal model except for ***age:vowel.quality*** which was rendered since it was used in the SB –audio and audio-written dataset analysis. The maximal model formula is presented in A. below along with the output summary in table 6-36.

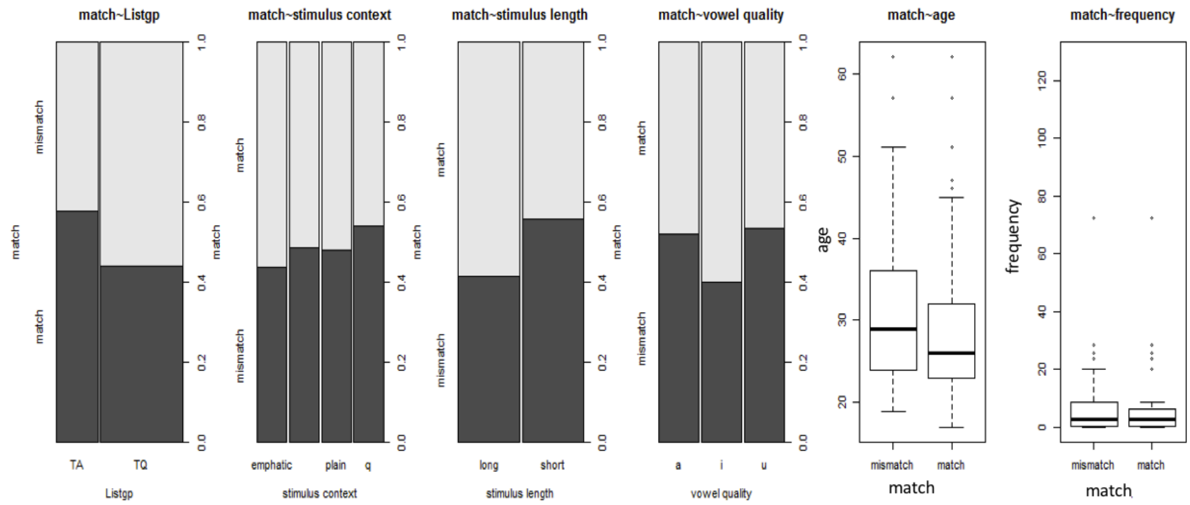


Figure 6-20: Fixed effects variables correlations with the response variable match

A. Maximal model

wdatadriven<-glmer(match~Listgp+context+length+vowel.quality+age+freq.+ Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality +context:length + context:freq.+ context:vowel.quality + length:freq.+ age:vowel.quality +freq.:vowel.quality + (Listgp|stimulus) + (length+context|listener) , data = SBwritten, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z)
(Intercept)	1.045025	1.190330	0.878	0.3800
ListgpTQ	0.668211	1.006180	0.664	0.5066
contextpharyngeal	-1.388187	1.363369	-1.018	0.3086
contextplain	-0.267354	0.924636	-0.289	0.7725
contextq	-1.037242	0.785232	-1.321	0.1865
lengthshort	0.720075	0.889892	0.809	0.4184
vowel.qualityi	**-2.545126**	**1.071763**	**-2.375**	**0.0176***
vowel.qualityu	-1.665911	1.042678	-1.598	0.1101
age	-0.020167	0.025923	-0.778	0.4366

freq.	0.021804	0.157871	0.138	0.8902
ListgpTQ:lengthshort	**-1.667604**	**0.741496**	**-2.249**	**0.0245***
ListgpTQ:contextpharyngeal	0.325962	0.831605	0.392	0.6951
ListgpTQ:contextplain	0.422577	0.785801	0.538	0.5907
ListgpTQ:contextq	-0.003625	0.669515	-0.005	0.9957
ListgpTQ:freq.	-0.004153	0.019714	-0.211	0.8331
ListgpTQ:vowel.qualityi	**2.583188**	**0.637461**	**4.052**	**5.07e-05*****
ListgpTQ:vowel.qualityu	0.886332	0.620258	1.429	0.1530
contextpharyngeal:lengthshort	-0.611310	0.882524	-0.693	0.4885
contextplain:lengthshort	-0.488142	1.021300	-0.478	0.6327
contextq:lengthshort	-1.191944	0.966373	-1.233	0.2174
contextpharyngeal:freq.	0.085364	0.066690	1.280	0.2005
contextplain:freq.	0.001137	0.070666	0.016	0.9872
contextq:freq.	0.041131	0.035804	1.149	0.2506
contextpharyngeal:vowel.qualityi	1.624351	1.543446	1.052	0.2926
contextplain:vowel.qualityi	**-1.980249**	**0.954106**	**-2.076**	**0.0379***
contextq:vowel.qualityi	1.336321	1.115198	1.198	0.2308
contextpharyngeal:vowel.qualityu	0.089599	1.330630	0.067	0.9463
contextplain:vowel.qualityu	1.149734	0.945783	1.216	0.2241
contextq:vowel.qualityu	0.490593	0.900277	0.545	0.5858
lengthshort:freq.	-0.056380	0.153173	-0.368	0.7128
vowel.qualityi:age	**0.043442**	**0.022611**	**1.921**	**0.0547.**
vowel.qualityu:age	0.017837	0.022599	0.789	0.4300
vowel.qualityi:freq.	**0.088991**	**0.042157**	**2.111**	**0.0348***
vowel.qualityu:freq.	0.070684	0.051105	1.383	0.1666

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 6-36:Maximal model output summary table of the SB written dataset

B. Maximal model simplification techniques:

A number of techniques were followed in order for the maximal model to converge. These include the following points.

1. Variables of interest and their pertinent interactions were included in the maximal model. However, one variable was not included in the model as it did not show variability and was not part of the SB audio and audio-written dataset logistic regression analysis. Moreover, interactions related to the control variable were not integrated in the model for simplification purposes except for age:vowel.quality which was in the maximal model of the SBAAW dataset.
2. The random effects structure was also simplified and included only the three variables Listgp, context and length as slopes in order to avoid any non-converges issues with the model (Barr et al. (2013)), thus the random structure (Listgp|stimulus) and (context+length|listener).
3. Theoretical assumptions were abided including collinearity as was done with the previous datasets.
4. The optimization of the model was maximized to 2e5; i.e., control=glmerControl(optCtrl=list(maxfun=2e5)) and nAGQ =1).

5. Continuous predictors were centered and scale in R prior to running the logistic models.

C. Model selection:

The model selection processes involved automatic logistic regression using the dropterm and update commands in the MASS package (Venables & Ripley, 2003) as was done in the analysis of the PAT data, SB audio, and SB audio+audio-written datasets. Five models were derived in nine steps using backward algorithm where a single interaction was deleted at a time when its p-value did not reach significance 5%. The interaction length:vowel.quality was added to the final fitted model in step_6. The dropterm application steps are provided in appendix 6-4 with the final model's output table given below in table 6-37.

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.833896	1.113904	0.749	0.4541	
ListgpTQ	1.002307	0.726003	1.381	0.1674	
contextpharyngeal	0.003590	0.594173	0.006	0.9952	
contextplain	-0.076145	0.563917	-0.135	0.8926	
contextq	**-1.228073**	**0.684598**	**-1.794**	**0.0728**	**.**
lengthshort	-0.466140	0.815756	-0.571	0.5677	
vowel.qualityi	**-2.655220**	**1.070751**	**-2.480**	**0.0131**	*****
vowel.qualityu	-1.278548	0.985371	-1.298	0.1944	
age	-0.019823	0.026858	-0.738	0.4605	
freq.	0.007846	0.017837	0.440	0.6600	
ListgpTQ:lengthshort	**-1.726584**	**0.674180**	**-2.561**	**0.0104**	*****
ListgpTQ:vowel.qualityi	**2.718546**	**0.609496**	**4.460**	**8.18e-06**	*******
ListgpTQ:vowel.qualityu	0.802186	0.572841	1.400	0.1614	
contextpharyngeal:vowel.qualityi	0.444133	0.890653	0.499	0.6180	
contextplain:vowel.qualityi	**-1.785754**	**0.833343**	**-2.143**	**0.0321**	*****
contextq:vowel.qualityi	**1.705846**	**0.971802**	**1.755**	**0.0792**	**.**
contextpharyngeal:vowel.qualityu	-1.199898	0.745276	-1.610	0.1074	
contextplain:vowel.qualityu	0.692812	0.731887	0.947	0.3438	
contextq:vowel.qualityu	0.069507	0.868078	0.080	0.9362	
lengthshort:vowel.qualityi	0.581692	0.839912	0.693	0.4886	
lengthshort:vowel.qualityu	0.336401	0.821025	0.410	0.6820	
vowel.qualityi:age	**0.042666**	**0.022842**	**1.868**	**0.0618**	**.**
vowel.qualityu:age	0.017526	0.022812	0.768	0.4423	
vowel.qualityi:freq.	**0.085704**	**0.042396**	**2.021**	**0.0432**	*****
vowel.qualityu:freq.	0.073253	0.059680	1.227	0.2197	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Table 6-37: simple effects of significant variables and interactions in the final model of the SB written dataset

As can be seen in table 6-37, two variables and five interactions are significant. These include ***context***, ***vowel.quality*** and ***Listgp:length***, ***Listgp:vowel.quality***, ***context:vowel.quality, age:vowel.quality*** and ***freq.:vowel.quality***.

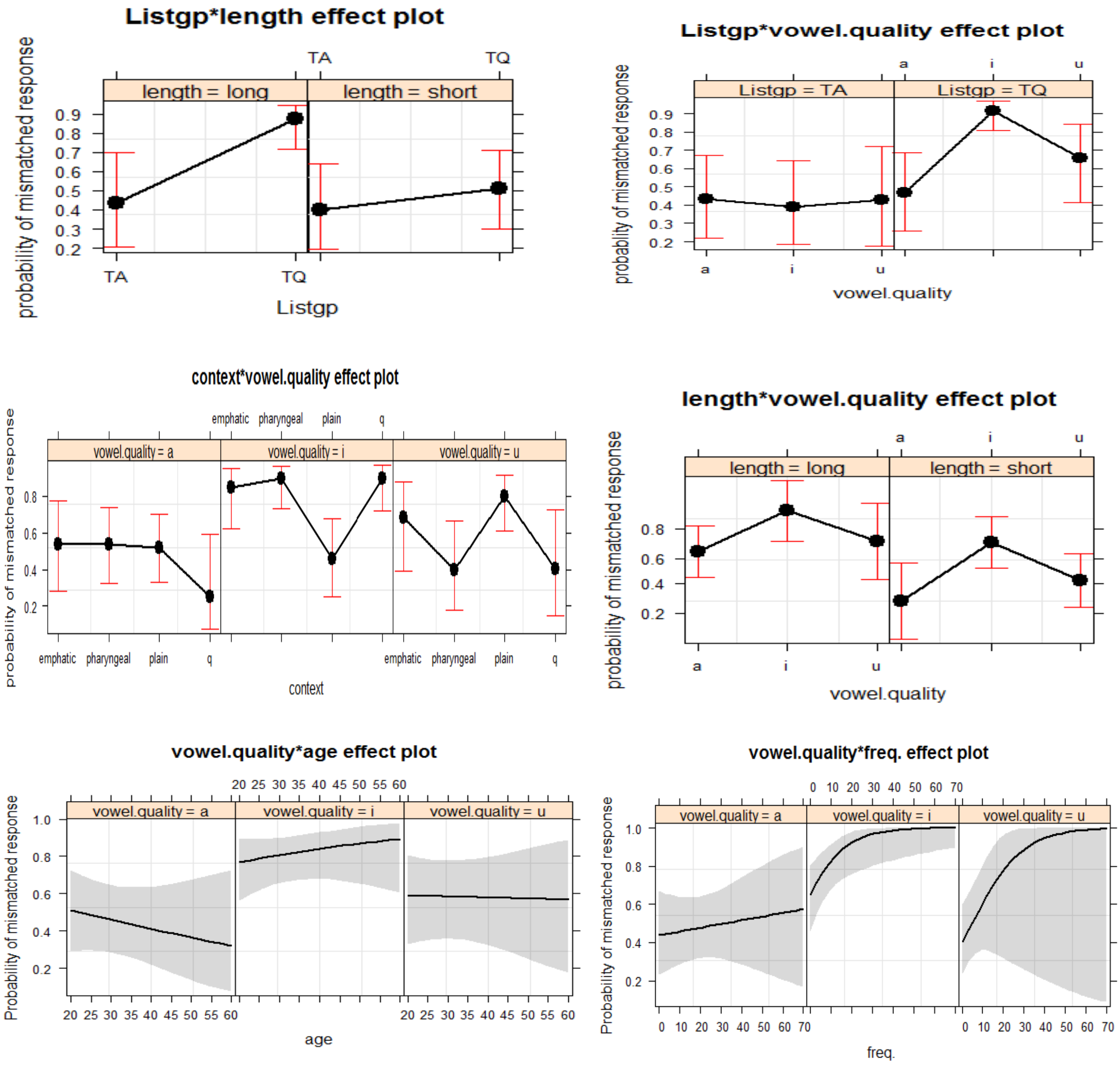


Figure 6-21: effect displays for all the interactions in model step_6

Next, the results are interpreted where again the contrasts coding match=0 and mismatch=1 is assumed and match echoes the observed mappings in the ALT corpus.

D. Model interpretation

The interaction of **Listgp** with **length** was found significant. First, the TA group exhibited higher degrees of match (<0.5) than the TQ group both when the stimulus vowel was short and long as shown in figure 6-22 of ***Listgp:lengh*** effect. Moreover, the TQ group manifested a clear increase in degrees of match when the stimulus vowel was short compared to when it was long. This is reflected in the significant interaction of **ListgpTQ:lengthshort** in table 6-37 as exemplified by $[i]_q$>$/i/_T$ (21/22; i:18, i-e:3, ü-e:1) compared to $[i:]_q$>$/i/_T$ (14/22; a:1, a-ı:1, e-e:1, i:8, i-a:1, i-e:1, ı:5, ıi:2, i-i:1, a-i-a:1).

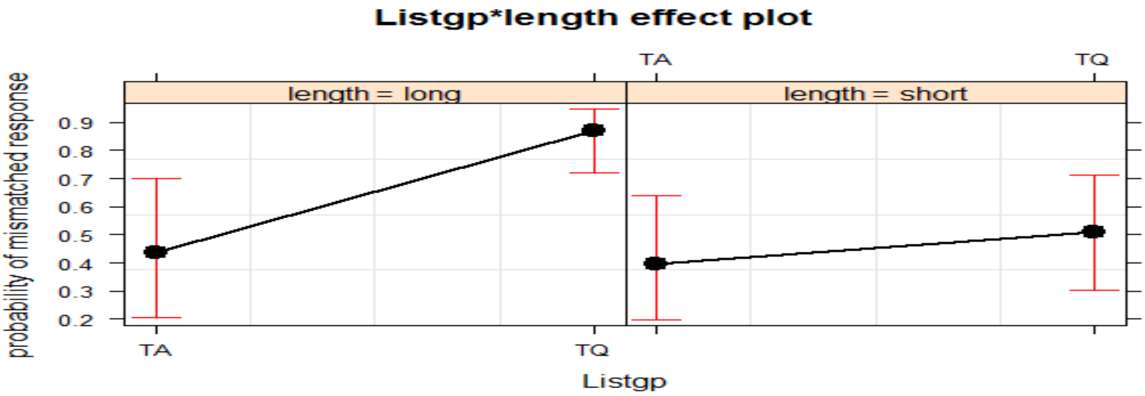


*Figure 6-22: Listgp*length effect display, mismatch-=1, match=0*

Moreover, the TA also manifested higher degrees of match across the three vowels /a/, /i/ and /u/ than the TQ group in the interaction ***Listgp:vowel.quality*** in figure 6-23. Nevertheless, each group reflected variation in their mapping of the three vowels. That is, the TA group yielded higher match score for the vowel /i/, followed by /u/ and /a/ whereas the TQ group /a/, /u/ and /i/. The /i/ vowel is displayed with the least degrees of match; almost 0.95. This is also reflected in table 6-35 in the interaction **ListgpTQ:vowelqualityi,** for example, $[i:]_q$>$/i/_T$ (14/22; a:1, a-ı:1, e-e:1, i:8, i-a:1, i-e:1, ı:5, ıi:2, i-i:1, a-i-a:1).

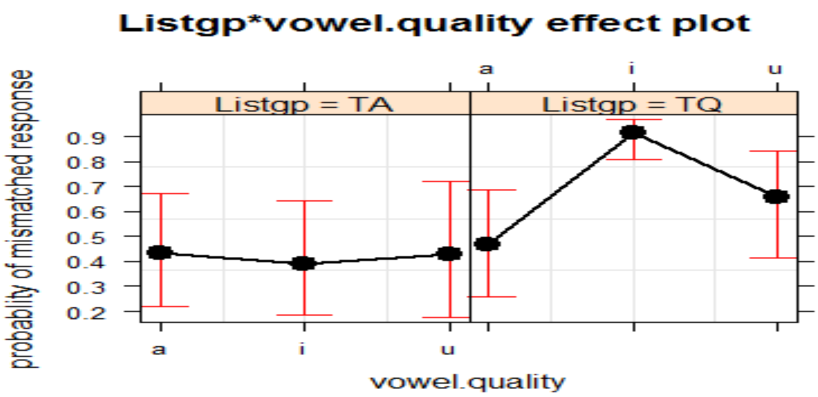


Figure 6-23:Listgp:vowel.quality effect dispolay, mismatch=1, match=0

The third interaction of ***context with vowel quality*** in their effect on match is depicted in figure 6-24. For the vowel /a/, listeners reflected higher degrees of match in the q environment and lower degrees of match in the plain, pharyngeal and emphatic, meaning they did not treat plain and guttural environments as dichotomous. As for the vowel /i/, listeners showed higher degrees of match in the plain environment and lower degrees of match in the emphatic, pharyngeal and q, i.e., lower for the guttural environment. As for the vowel /u/, listeners manifested higher degrees of match in the pharyngeal and q environment and lower degrees of match for the emphatic followed by the plain environment.

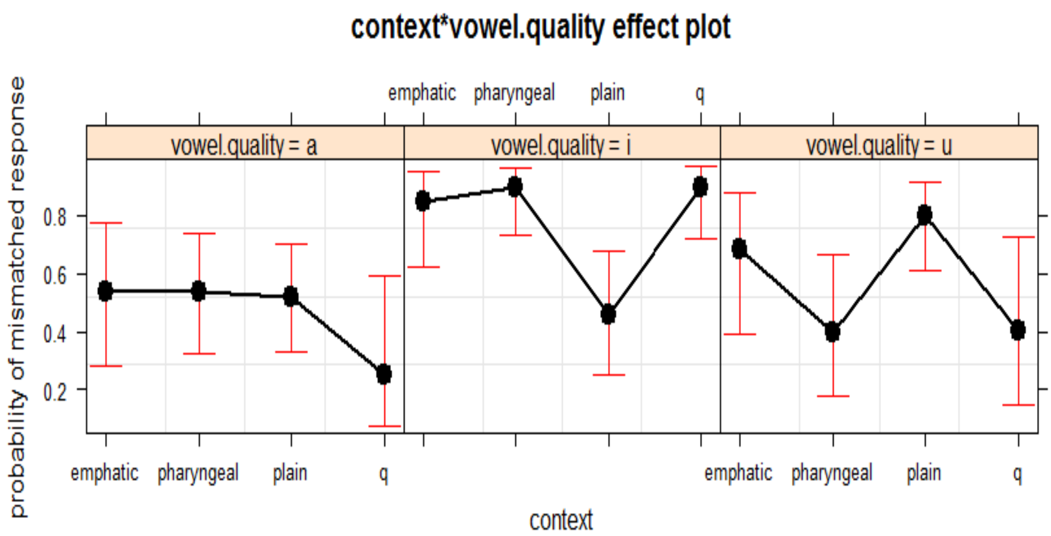


*Figure 6-24: context*vowel.quality effect display, mismatch=1, match=0*

As for the interaction of ***vowel quality and age*** (figure 6-25), listeners demonstrated different patterns across various age points. Older listeners exhibited higher degrees of match for the vowel /a/ compared to younger ones; slightly higher degrees of match for the vowel /u/ and lower degrees of match for the vowel /i/ than younger listeners. However, younger listeners reflected even lower degrees of match for the three vowels. They incurred 50-50 degrees of match for the vowel /a/, almost 40% for the vowel /u/ and about 20% for the vowel /i/.

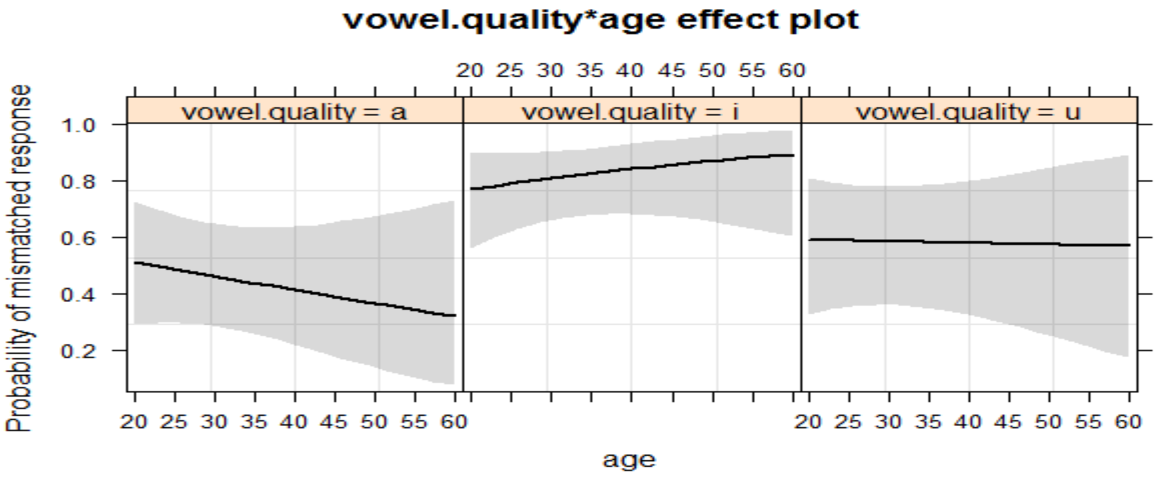


*Figure 6-25: vowel.quality*age effect display, match=0, mismatch=1*

The last interaction of ***vowel quality and freq.*** in their effect on match reveals main effects of the interaction and effect of the vowel.quality variable. As elucidated in figure 6-26, at a frequency of 0 hz (nonsense words or low frequency words), listeners tend to reflect the highest degrees of match for the vowel /u/ followed by the vowel /a/ and lower degrees of match for the vowel /i/. However, at a maximum frequency (72.250 hz), they tended to demonstrate the highest degrees of match for the vowel /a/ and the lowest for the vowels /u/ and /i/. As for the effects of the vowel quality variable, the assimilation patterns of the vowel /a/ slightly deteriorates whereas listeners' perception of the vowels /u/ and /i/ considerably deteriorates when the stimulus word is of high frequency.

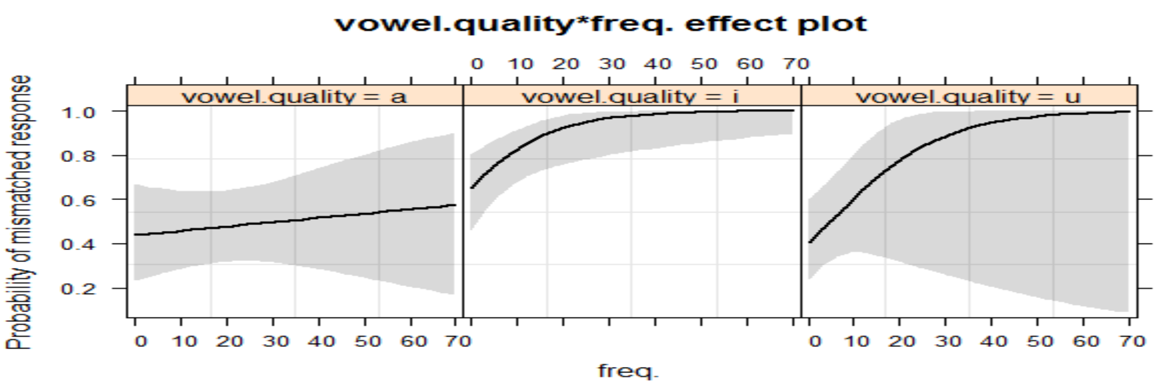


*Figure 6-26: vowel.quality*freq. effect display; match=0, mismatch=1*

E. Model validation

As done in the previous logistic regression analyses in section 6.4.3.2 and chapters 4 and 5, the results of the final model were verified by plotting the model's residuals for goodness of fit. Figure 6-27 indicates that the final model is of good fit since the residuals are heteroscedastic, not forming stochastic patterns, and of an almost normal distribution.

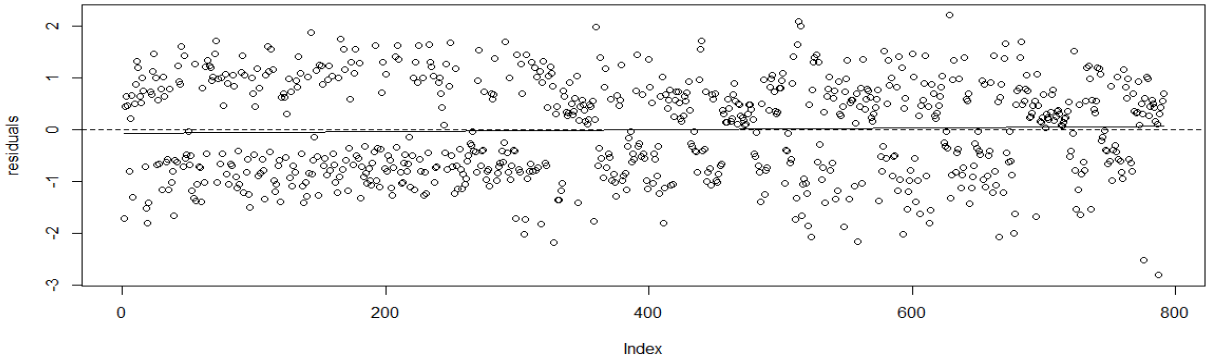


Figure 6-27: Checking for homoscedasticity

6.5.1.2 Reporting results

A GLMM regression analysis was conducted in R to forecast the correlation between the dependent variable match and six variables along with ten interactions using a data-driven approach based on the research hypotheses. The fixed effects structure included ***Listgp***, ***length***, ***context***, ***freq.***, ***vowel.quality*** and **age**, which was used as a control variable. The interactions involved ***Listgp:length***, ***Listgp:context***, ***Listgp:freq.***, ***Listgp:vowel.quality***, ***context:length***, ***context:freq.***, ***context:vowel.quality*** , ***length:freq.***, ***age:vowel.quality*** and ***freq.:vowel.quality***. The variable ***length:vowel.quality*** was added to final model after running the dropterm applications. A non-maximal random effects structure with slopes and intercepts was assumed which consisted of the variables of main theoretical interest; i.e., ***Listgp***, ***context*** and ***vowel quality***. The theoretical assumptions of collinearity, homoscedasticity and normality were examined and adhered to throughout the analysis.

Two variables and five interactions were found to have effects on the response variable match. These are ***context***, ***vowel.quality***; ***Listgp:length***, ***Listgp:vowel.quality***, ***context:vowel.quality, age:vowel.quality*** and ***freq.:vowel.quality***. The main finding of the analysis is that the TA group displayed higher degrees of match in both the ***Listgp:length*** and ***Listgp:vowel.quality*** interactions than the TQ group confirming the results of the raw data in section 6.3.1. As for the variables context and vowel.quality, their main effects were evident in the interaction ***context:vowel.quality*** where listeners mapped the vowel /a/ with higher degrees of match in the /q/ environment and lower degrees of match in the plain and two other guttural environments, the vowel /i/ with higher degrees surrounding plain consonants and lower degrees surrounding the gutturals (emphatic, pharyngeal and q), and the vowel /u/ with higher degrees of match in the pharyngeal and q settings and lower degrees of match surrounding emphatic and plain consonants.

In addition, listeners tended to incur higher degrees of match for short vowels regardless of the ***vowel quality***. However, they also reflected simple effects of the interaction of ***length*** and ***vowel.quality*** since they exhibited high degrees of match for the short vowel /a/ compared to low degrees for its long counterpart, low degrees of match for the short vowel /i/ compared to even lower degrees of match for its long counterpart and high degrees of match for the short vowel /u/ compared to lower degrees of match for its long counterpart.

Furthermore, the main effects of vowel quality was reported in the interaction ***vowel.quality:age*.** Older listeners tended to display dramatically higher degrees of match for the vowel /a/ than younger listeners, dramatically lower degrees of match for the vowel /i/ and slightly improved performance for the vowel /u/ compared to younger listeners. Moreover, vowel quality varied across nonsense and real words (***vowel.quality:freq.***). It was found that listeners tended to exhibit high degrees of match for the vowel /a/ both in nonsense and high-frequency real words, the highest degree of match for the vowel /u/ in nonsense words but the lowest in high-frequency words and low degrees of match for the vowel /i/ in nonsense words but even the lowest in high-frequent words.

In a nutshell, in the SB-written task we see that both the TA and TQ reflected similar patterns; however, the TA do better, but there is an effect of orthography for both TA and TQ.

6.6 *Summary and discussion*

To recap, listener groups' performance in the audio and audio-written conditions was compared, on the one hand, and within the written condition on the other. Both raw data and results from the GLMM analysis showed that the groups with knowledge of Arabic (TA and TQ) exhibited higher degrees of match in the audio-written condition when compared with the audio one. The degree of match was higher in the audio-written condition at a percentage of 61.9% compared to 48.44% in the audio-only task by the three groups as exemplified by the significant interaction **stimulus.presentationaudio+written** in table 6-34. This, in turn, corroborates the hypothesis that knowledge of Arabic writing system (orthography) enhances the degrees of match to the corpus patterns.

This finding supports the hypothesis that perception is not the only factor responsible for the Arabic loanword adaptation into Turkish as shown in the SB experiments thus far. The role of bilinguals is also accentuated since they manifested the highest degrees of match in the audio and audio-written experiment, on the one hand, and in the written one on the other.

In what follows, we first discuss the perceptual mapping patterns of the three groups in the audio-written condition compared to the corpus facts as in figure 6-28 and to the mappings of the same three groups in the audio-only task presented in chapter five and reproduced in figure 6-29. This is so as to review and discuss which patterns were mismatched before evaluating the role of perception and orthography in the adaptation process. These patterns were given in section 6.3 whereas the ones pertaining to the audio-only task were provided in section 5.3.1. in chapter five.

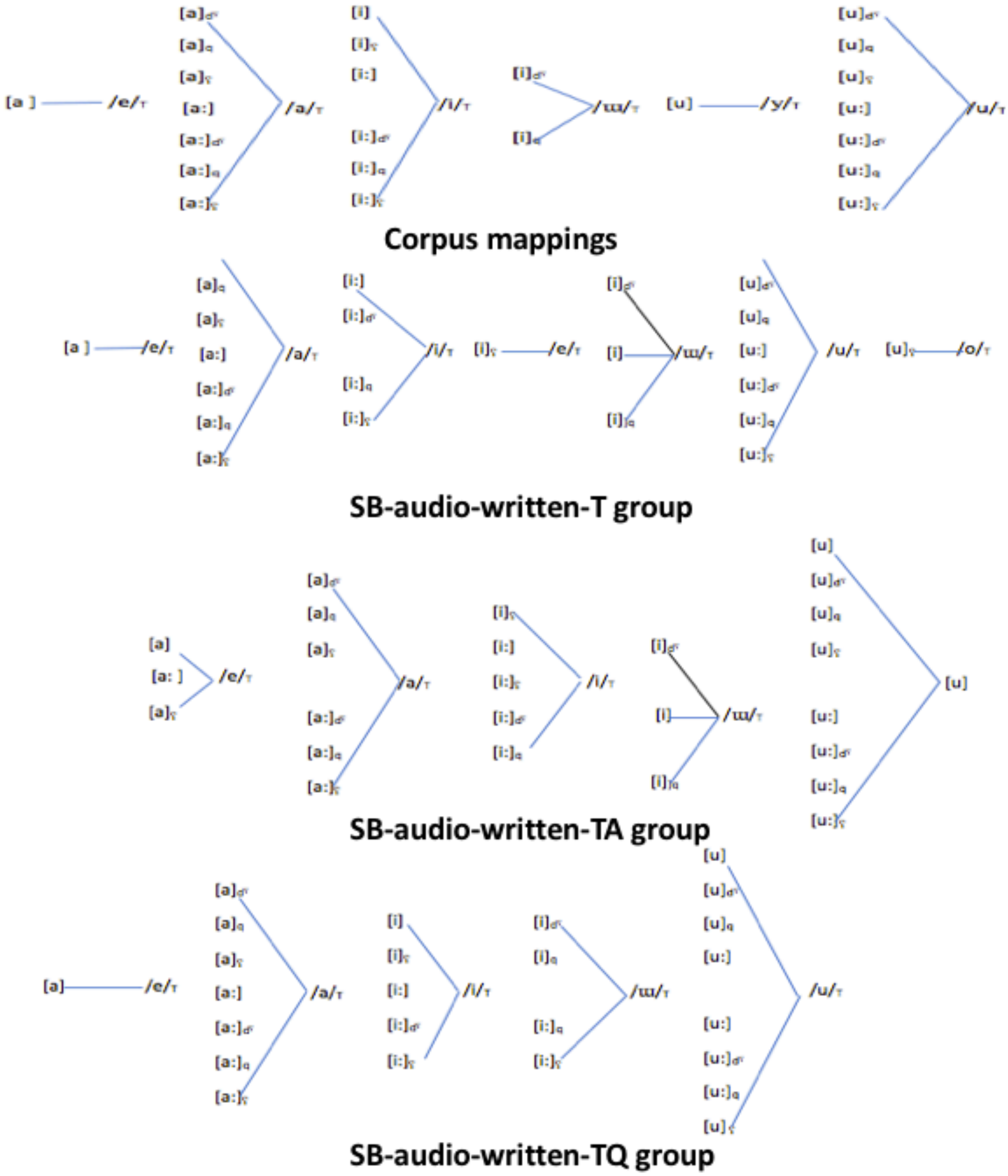


Figure 6-28: corpus and Turkish listener groups' mappings of Arabic loanwords into Turkish in the SB-audio-written task

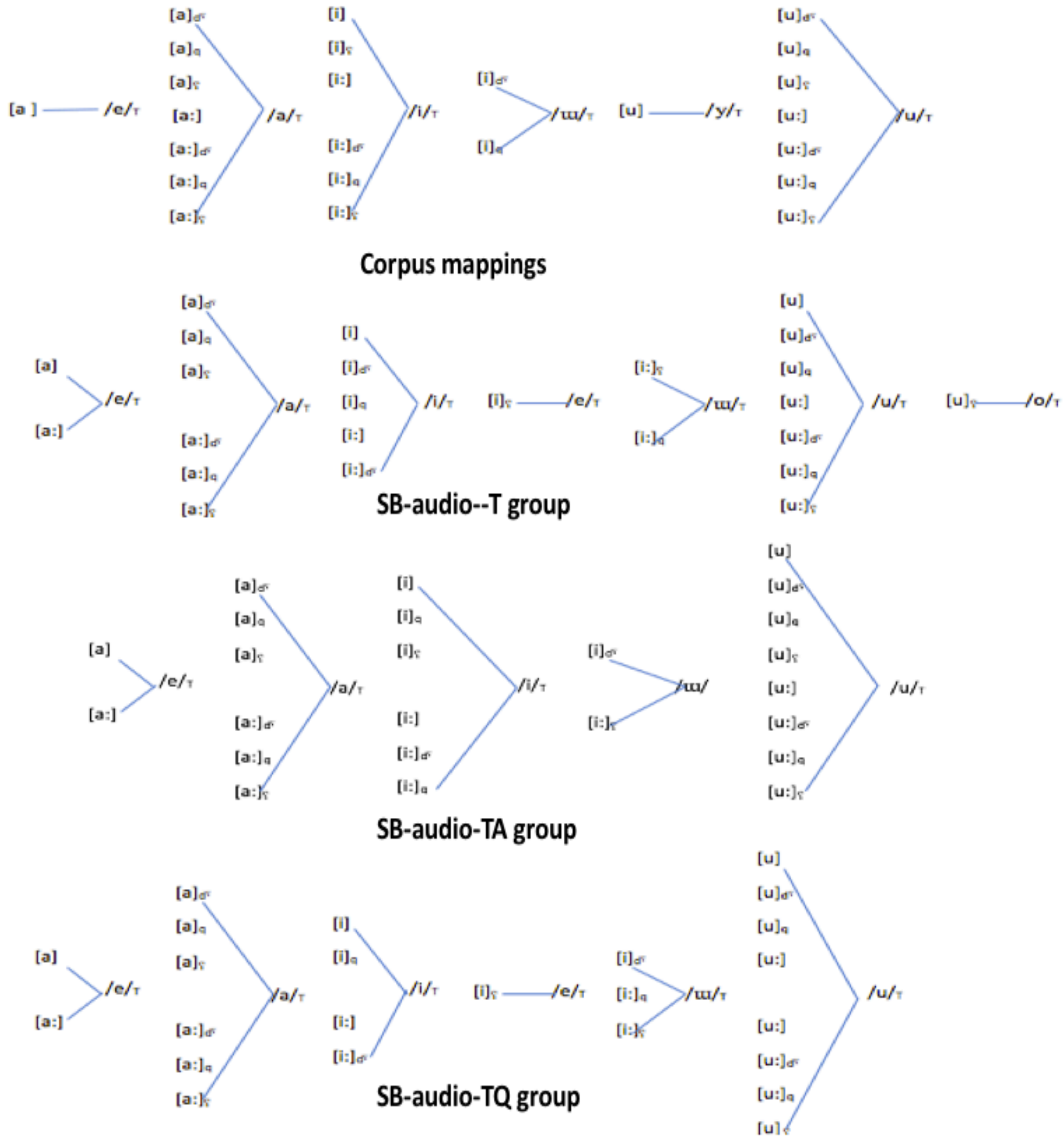


Figure 6-29: corpus and Turkish listener groups' mappings of Arabic loanwords into Turkish in the SB-audio task

In figure 6-28, we notice that the TA group reflected four mismatched response categories in the SB-audio-written task, namely [u]>/u/$_T$ (predicted /y/$_T$), [a:]>/e/$_T$ (predicted /a/$_T$) , [i]>/ɯ/$_T$ (predicted /i/) and half the TA participants mismatched [a]$_{ʕ}$ as /e/$_T$ (predicted /a/$_T$) whereas the other half matched it to /a/$_T$. TA had fewer mismatched categories than T group since the TA group displayed matched responses at a percentage of 66.28% in the audio-written condition. In the logistic regression, this difference between T and TA was exemplified by the significant interactions **ListgpTA:vowel.qualityi** and **ListgpTA:vowel.qualityu**. Similarly, in the SB-audio task the TA group also mismatched the two categories [a:] and [u] to /e/$_T$ (predicted as /a/$_T$) and /y/$_T$ respectively as. However, they mismatched [i]$_q$ and [i:]$_{ʕ}$ to /i/$_T$ (predicted as /ɯ/$_T$) and /ɯ/$_T$ (predicted /i/$_T$) respectively in the SB-audio but not the SB-audio-written condition.

The TQ group yielded four mismatched perceptual maps[34]: [u]>/u/$_T$ (predicted /y/$_T$), [i:]$_q$>/ɯ/$_T$ (predicted /i/$_T$) and [i:]$_{ʕ}$>/ɯ/$_T$ (predicted /i/$_T$) where half the TQ group mismatched the source vowel in this latter pattern and the other half matched it to /i/$_T$ which was mirrored in the significant interaction **ListgpTQ:stimulus.presentationaudio+written**. The overall percentage of their matched responses formed reached almost 65%. In the SB-audio task, not only did the TQ render the three mismatched perceptual maps [i:]$_q$>/ɯ/$_T$, [i:]$_{ʕ}$>/ɯ/$_T$ and [u]>/u/$_T$ as in the SB-audio-written but they also mismatched the three other perceptual maps [a:]>/e/$_T$, [i]$_q$>/i/$_T$, and [i]$_{ʕ}$>/e/$_T$.

The T group whose overall matched responses reached only, 54% reflected the highest number of mismatched maps of four in total, namely [u]>/u/$_T$ (predicted /y/$_T$), [i]>/ɯ/$_T$ (predicted /i/$_T$), [i]$_{ʕ}$>/e/$_T$ (predicted /i/$_T$) and [u]$_{ʕ}$>/o/$_T$ (predicted /u/$_T$). In the SB-audio task, the T group mismatched eight categories, three of which were the same in the SB-audio-written task namely [u]>/u/$_T$ (predicted /y/$_T$), [i]$_{ʕ}$>/e/$_T$ (predicted /i/$_T$) and [u]$_{ʕ}$>/o/$_T$ (predicted /u/$_T$). The five other mismatched perceptual maps are [a:]>/e/$_T$ (predicted /a/$_T$), [i]$_{d^ʕ}$>/i/$_T$ (predicted /ɯ/$_T$), [i]$_q$>/i/$_T$ (predicted /ɯ/$_T$), [i:]$_q$>/ɯ/$_T$ (predicted /i/$_T$ and [i:]$_{ʕ}$>/ɯ/$_T$ (predicted /i/$_T$). The mismatched patterns by the three groups in the SB-audio-written task are summarized in table 6-38 and are discussed below.

---

[34] A perceptual map refers to the source category being mapped onto the native langueg category, such as in [i]$_q$>/i/$_T$ where [i]$_q$ is the source language category, /i/$_T$ is the borrowing language category and (>) means 'mapped onto'.

S.no	Mismatched category	Listgp
1.	$[u]_A$>/u/$_T$ (predicted /y/$_T$)	T, TA, TQ
2.	/i/$_A$ >/ɯ/$_T$ (predicted /i/$_T$).	T, TA
3.	/a:/>/e/$_T$ (predicted /a/$_T$)	TA
4.	$[i:]_q$>/ɯ/$_T$ (predicted /i/$_T$)	TQ
5.	$[i:]_ʕ$>/i/$_T$ or /ɯ/ (predicted /i/$_T$)	TQ
6.	$[i]_ʕ$>/e/$_T$ (predicted /i/$_T$)	T
7.	$[u]_ʕ$>/o/$_T$ (predicted /u/$_T$)	T

Table 6-38: mismatched vowel categories in the SB-audio-written task

In the first pattern in table 6-38, all three groups perceive the back rounded Arabic vowel [u] as its Turkish counterpart /u/$_T$ which is the same mapping they all manifested in the audio-only task. This assimilation can be justified on perceptual grounds since the two vowel categories $[u]_A$ and /u/$_T$ are close to each other in F1 [height] and F2 [backness] on the vowel space in figure 6-30. In addition, $[u]_A$ and /u/$_T$ share the same phonological vowel quality (agree in height and backness). In the logistic regression analysis, this was reflected in the interaction ***context:length*** in figure 6-14 where listeners displayed the lowest degrees of mismatch in the plain context.

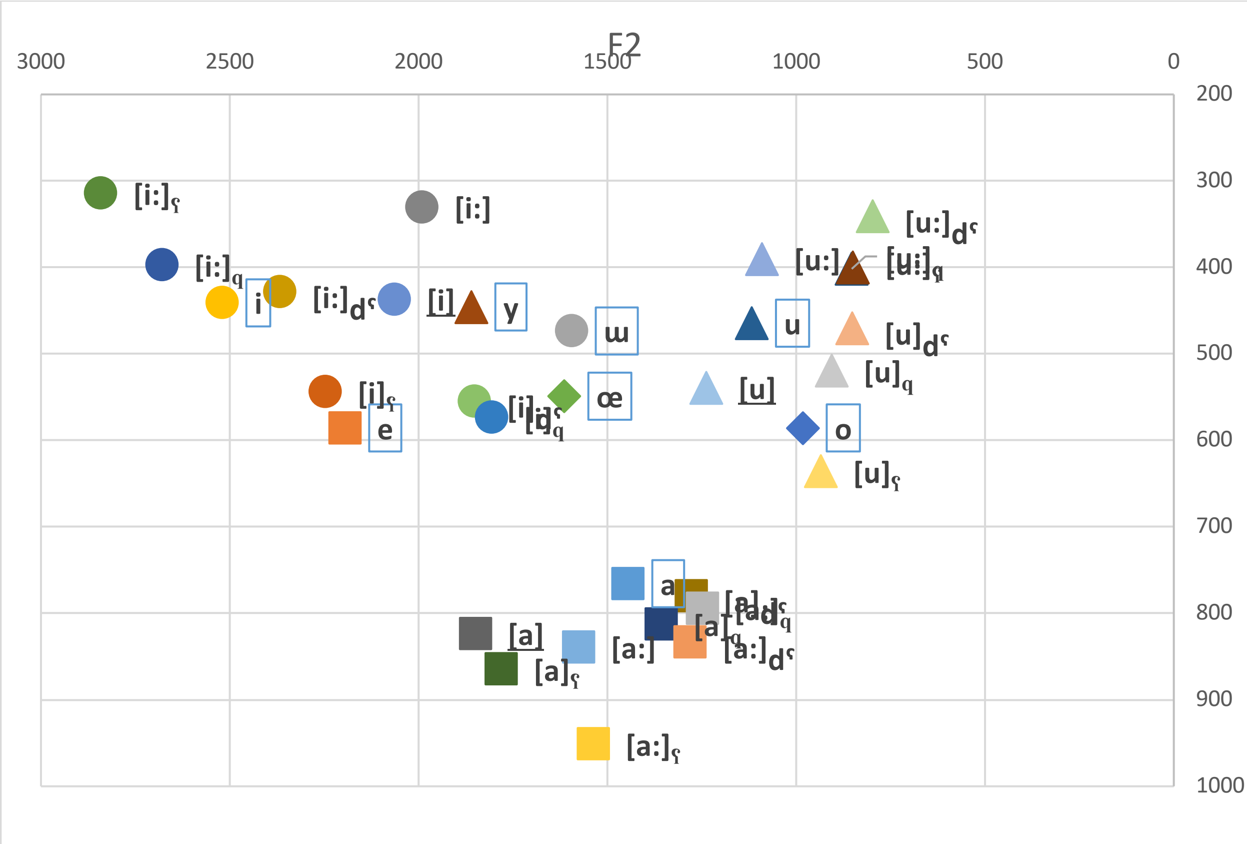


Figure 6-30: Mean frequency values of 2 Turkish speakers and 1 Arabic speaker's vowel formants plot (real words) in the SB-audio-written task

Vowels in boxes are Turkish vowels and the rest are Arabic ones; underlined /a/, /i/and /u/ are Arabic vowels

In the second pattern in table 6-38, only the T and TA groups mapped the vowel $[i]_A$ as $/ɯ/_T$ where $/i/_T$ is predicted but not the TQ group, so knowledge of Arabic (phonology) can be safely ruled out as the source of the categorization. As for the position of the $[i]_A$ and $/ɯ/_T$ on the vowel space, we can see that other Turkish vowel categories are closer than $/ɯ/_T$ such as $/e/_T$ and $/i/_T$, the latter of which is correctly mapped to by the TQ group. In the logistic regression performed, this was reflected in the significant interaction **contextplain:vowelqualityi** in figure 6-13.

In the third mismatched mapping in table 6-38, the TA group only mismatched the Arabic category [a:] as $/e/_T$ as shown in table 6-38 where $/a/_T$ is predicted. In the logistic regression run, this pattern was reflected in the interaction length:vowel.qulity in figure 6-18 where listeners reflected higher degrees of match for short $[a]_A$ compared to $[a:]_A$. This pattern can be explained on phonological grounds since $[a:]_A$ and $/e/_T$ are both [-high] and [+front] vowels whereas $/a/_T$ is [-high] and [+back]. Hence, $[a:]_A$ and $/e/_T$ are phonologically more similar. In the logistic regression analysis, this pattern was reflected in the significant interaction **contextplain:lengthshort** in figure 6-14.

The assimilation by the T and TQ groups of the perceptual map $[a:]>/a/_T$ can be said to be perceptually grounded. As illustrated in figure 6-30, $[a:]_A$ and $/a/_T$ are acoustically closer to each other in both height and distance than $[a:]_A$ and $/e/_T$. Therefore, the assimilation of the T and TQ of $[a:]_A$ is phonetically grounded whereas it is phonologically driven by the TA group.

In the fourth and fifth mismatched perceptual maps in table 6-38, all the listeners of the TQ group mismatched $[i:]_q$ and half of them mismatched $[i:]_ʕ$ as $/ɯ/_T$ instead of $/i/_T$. Studying the vowel space in figure 6-30 reveals that the closest Turkish vowel category to the two Arabic categories $[i:]_q$ and $[i:]_ʕ$ is in fact $/i/_T$ and not $/ɯ/_T$. This suggests that these two maps are not perceptually motivated. Moreover, $[i:]_q$ and $[i:]_ʕ$ are phonologically more similar to $/i/_T$ in height and frontness than to $/ɯ/_T$ with which they only share height but not backness. This indicates that these perceptual maps are not phonologically supported.

In the sixth pattern in table 6-38, only the T group mismatched $[i]_ʕ$ as $/e/_T$. In terms of the position of the two vowel categories in figure 6-30, we notice that $/e/_T$ is closer to $[i]_ʕ$ closer than $/i/_T$ or even $/ɯ/_T$ in both height and frontness. This means that the mapping of the T group of the perceptual map $[i]_ʕ> /e/_T$ is phonetically based. On the other hand, the mapping of the TA and TQ groups might be influenced by their knowledge of Arabic (phonology) since they both matched [i] as $/i/_T$. This latter map is indeed phonologically sustained since [i] and $/i/_T$ share the two phonemic features [+high] and [+front]. In the SB-audio task, only the T and TQ mismatched $[i]_ʕ$ as $/e/_T$. This suggests that seeing the script in addition to hearing it has improved the degrees of match for the TQ group. This is reflected in the significant interaction of **ListgpTQ:stimulus.presentationaudio+written** in table 6-34.

In the last pattern in table 6-38, similar to what was seen in the SB-audio only task, the T group assimilated the $[u]_{ʕ}$ category as $/o/_{T}$. These two vowel categories are shown in the vowel space (figure 6-32) as being closer in F1 and F2 to each other than $/u/_{T}$ is to $[u]_{ʕ}$. This means that this perceptual map of the T group is phonetically grounded. On the other hand, it is phonologically motivated for the TA and TQ groups who both matched $[u]_{ʕ}$ to the predicted $/u/_{T}$. That is, $[u]_{ʕ}$ and $/u/_{T}$ are similar in height and backness whereas $[u]_{ʕ}$ and $/o/_{T}$ are similar only in backness. In the logistic regression analysis, the interaction ***context:vowel.quality*** in figure 6-13 encompasses the $[u]_{ʕ}$> $/o/_{T}$ pattern since listeners displayed higher degrees of match for the vowel /u/ in the guttural environment including the pharyngeal context than the plain one. The assumption here is that this is true since both the TA and TQ groups matched the predicted vowel quality $/u/_{T}$ in the pharyngeal context.

All in all, the T group acted as naïve listeners as expected, the TA group were closer to the corpus patterns, however, with some interference from Arabic phonology while the TQ group taking a medial position. The two groups with Arabic knowledge displayed higher degrees of match compared to the monolingual Turkish group which supports the hypothesis that knowledge of Arabic writing affects the degree of match of the observed vowel mappings in the SB-audio-written task to those observed in the corpus. This result was expressed in the interaction ***Listgp:stimulus.presentation***. In addition, seven other interactions were also found to have a role on the matching as summarized in table 6-39.

S.no	***Variable/interaction***	***SBaudio+Audio-written (2 V+ 8I)***
1	**Listgp**	
2	**Context**	
3	**Freq.**	✓
4	**Age**	
5	**Length**	✓
6	**Vowel.quality**	
7	**Stimulus.presentation**	✓
9	**Listgp: stimulus.presentation**	✓
10	**Listgp:context**	
11	**Listgp:length**	
12	**Listgp:vowel.quality**	✓
13	**Context:length**	✓
14	**Context:freq.**	✓
15	**Context:vowel.quality**	✓

16	**Length:vowel.quality**	✓
17	**Vowel.quality:age**	✓
18	**Vowel.quality:freq.**	✓

Table 6-39: A summary table of the significant variables and interactions in the SB-audio-written task

As for the results of the SB-written task, it was found that the TA and TQ groups both yielded similar categorization patterns as in figure 6-39 which suggests that ***Listgp*** as a variable alone does not have a significant effect on the matching. However, the interaction of ***Listgp*** with ***length*** and ***vowel quality*** respectively plays a role in the mapping since the TA group exhibited higher degrees of match in both interactions than the TQ group.

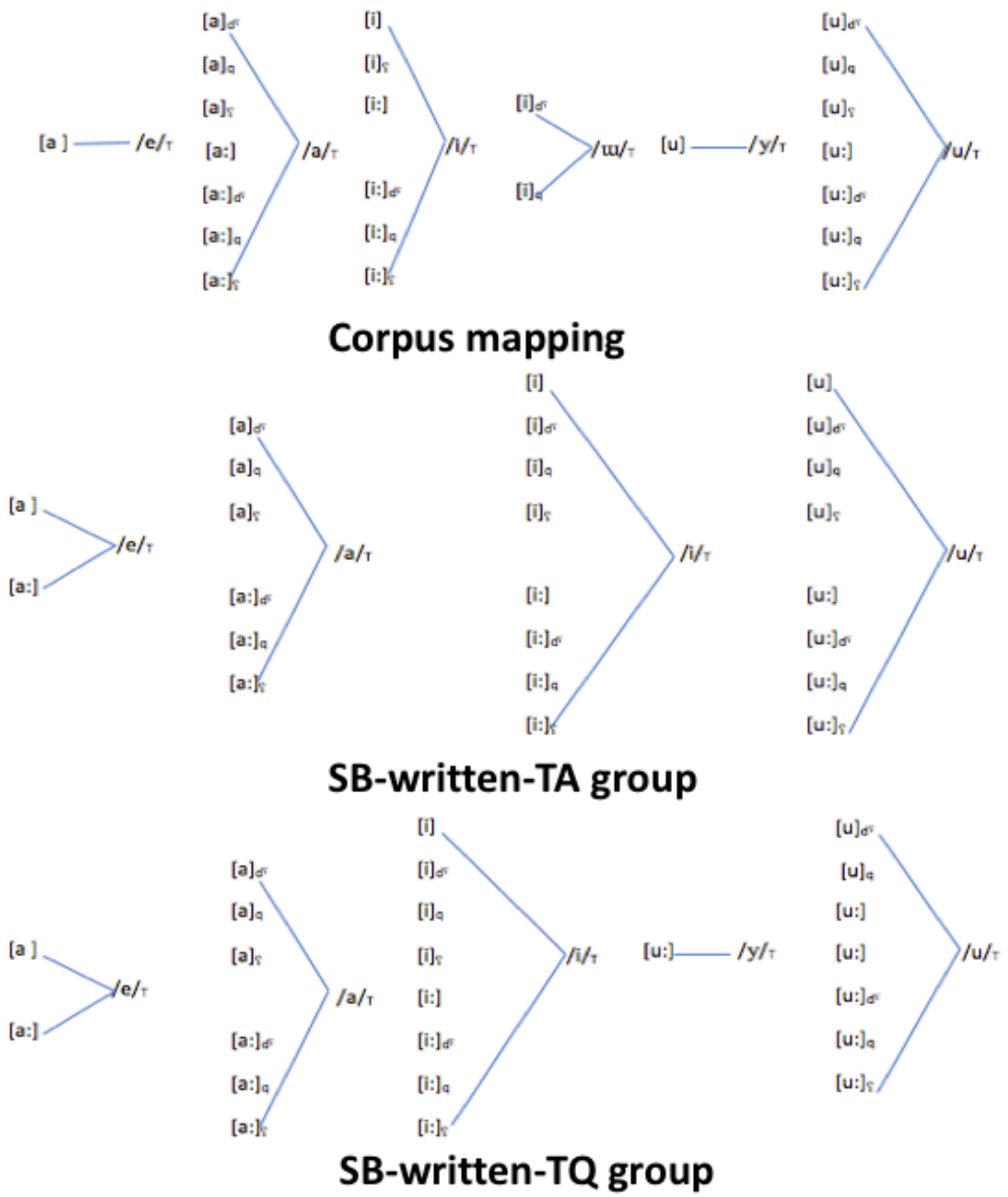


Figure 6-31: corpus and Turkish listener groups' mappings of Arabic loanwords into Turkish in the SB-written task

When comparing the TA and TQ mappings in the SB-written condition to that of the corpus (figure 6-31), we can see that the number of mismatched categories is not as many as those in the audio condition (tables 6-40 and 6-41). This indicates that seeing the spelling of Arabic improves the percentage of match.

S.no	Mismatched category	Listgp
1.	[a:]>/e/$_{T}$ (predicted /a/$_{T}$)	T, TA, TQ
2.	[i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$)	T, TA, TQ
3.	[u]>/u/$_{T}$ (predicted /y/$_{T}$)	T, TA, TQ
4.	[i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted /i/$_{T}$)	T, TA, TQ
5.	[i]$_{ʕ}$>/e/$_{T}$ (predicted /i/$_{T}$)[35]	T and TQ
6.	[i:]$_{q}$>/ɯ/$_{T}$ (predicted /i/$_{T}$)	T and TQ

Table 6-40: mismatched vowel categories in the SB-audio task

S.no	Mismatched category	Listgp
1.	[a:]>/e/$_{T}$ (predicted /a/$_{T}$)	TA & TQ
2.	[i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$)	TA & TQ
3.	[u]>/u/$_{T}$ (predicted /y/$_{T}$)	TA & TQ
4.	[i]$_{d^{ʕ}}$>/i/$_{T}$ (predicted /ɯ/$_{T}$)	TA &TQ
5.	[u:]>/y/$_{T}$ (predicted /i/$_{T}$)	Half the TQ group

Table 6-41: mismatched vowel categories in the SB-written task

Crucially, we notice in table 6-41 that the listeners in the TA and TQ groups still mismatched the two patterns [a:]>/e/$_{T}$ (predicted /a/$_{T}$) and [u]>/u/$_{T}$ (predicted /y/$_{T}$) in the written condition as they did in the two other conditions. This suggests that the source of the mismatch is Arabic phonology since /y/ is not part of the Arabic inventory and [a:]>/e/$_{T}$ are phonologically similar in being [+front] and [-high].

In addition, both groups mismatched the two patterns [i]$_{d^{ʕ}}$>/i/$_{T}$ (predicted /ɯ/$_{T}$) and [i]$_{q}$>/i/$_{T}$ (predicted /ɯ/$_{T}$) probably being influenced by Arabic phonology since [i]$_{d^{ʕ}}$ and [i]$_{q}$ share the two distinctive features of [+high] and [+front] with [i]$_{A}$. These two mismatched mappings were reflected in the finding that the ***context*** and ***vowel quality*** of the stimulus and their interaction ***context:vowel.quality*** had effects on the matched responses. The last mismatched pattern was that of [u:]>/y/$_{T}$ (predicted /i/$_{T}$) by half the TQ group only who might have been influenced by their Turkish phonology of front-back vowel harmony distinction. That is, half the TQ participants might have interpreted the long vowel in the written word as a front vowel as a result of the absence of guttural/emphatic consonants from the Arabic

[35] In the mapping [i:]$_{ʕ}$>/a-ɯ/$_{T}$ (predicted /i/$_{T}$), the response vowels were considered as /ɯ/. Whenever the listeners responded with two vowels, the second vowel was only considered since long vowels are not allowed in Turkish.

word. Table 6-42 provides a summary of the variables and interactions shared in the SB-written and SB-audio conditions and the ones significant in each condition.

S.no	***Variable/interaction***	***SBaudio 4 variables+ 9 interactions***	***SBwritten 2 variables+ 6 interactions***
1.	**Context**	✓	✓
2.	**Freq.**	✓	
3.	**Length**	✓	
4.	**Vowel.quality**	✓	✓
5.	**Listgp:context**	✓ (near significant)	
6.	**Listgp:length**		✓
7.	**Listgp:vowel.quality**	✓	✓
8.	**Context:length**	✓	
9.	**Context:freq.**	✓	
10.	**Context:vowel.quality**	✓	✓
11.	**Length:freq.**	✓	
12.	**Length:vowel.quality**	✓	
13.	**Vowel.quality:age**	✓	✓
14.	**Vowel.quality:freq.**	✓	✓

Table 6-42: A summary table of the significant variables and interactions in the SB-written and SB-audio tasks

In conclusion, it was found in the simulated borrowing tasks that the groups with knowledge of Arabic orthography and phonology displayed closer degrees of match to the corpus patterns with the TA group exhibiting the highest degrees of match. This finding entails that perception alone is not responsible for the mismatched patterns as orthography as well was found to play a role. Nevertheless, the perceptual maps of the three groups together exemplify the effects of perception for the most part in addition to effects of Arabic phonology, Turkish phonology and morphology (to a limited extent), and Arabic orthography. By extension, this suggests that a hybrid model of both phonetics, phonology (Arabic and Turkish), orthography and Turkish morphology should be used to analyze the loanword corpus data. Moreover, the results of the SB-written task suggest that the knowledge of the Ottomans of the Arabic loanwords was of high proficiency as that of bilinguals; i.e., spoken and written.

7 General discussion, conclusions and implications

7.1 Introduction

This chapter discusses the main findings that addressed the main research questions of the thesis. These outcomes are interpreted vis-à-vis the ongoing debate of the three loanword adaptation positions; ***perception-only*** (Boersma, 2009; Peperkamp & Dupoux, 2003; Peperkamp et al., 2008; Silverman, 1992), ***phonology-only*** (LaCharité & Paradis, 2005; Paradis, 1995; Paradis & LaCharité, 1997, 2001, 2008; Peperkamp et al., 2008; Silverman, 1992) or a **medial hybrid** stance involving **both perception and phonology** (Kenstowicz and Suchato, 2006; Smith, 2006, Chang, 2008 and Dolus, 2013) and what these findings mean to loanword phonology. In addition, the concept of bilingualism is revisited and defined in relation to the findings of the thesis, and what this means for interpretation of the data shown here regarding Arabic loanwords in Turkish. Finally, the chapter concludes with the work's implications, limitations and contributions.

According to the uniformitarian principle, (Murray, 2015) present sound changes must have operated in the past under the same laws or principles. Thus, it is possible to model such sound changes in laboratories or, by extension, experimentally. In our context, this was the case with the residual effects of gutturals on neighbouring vowels (corpus data); the assumption driving the PAT and Simulated Borrowing tasks was that the adaptation process the speakers of Osmanlica shouldered in the past could be modeled with present day Turkish listeners in order to shed light on a) the mode of input of the borrowing, b) who performs the borrowing and c) factors influencing the borrowing.

7.2 Main findings and what they mean

A number of findings in relation to loanword phonology were borne out in this thesis. In chapter three, a new corpus was presented on the ALT data in modern day Turkish from which vocalic mappings were identified. This task was followed by stratifying the words by the time period they were in use. This led to identifying the residual effects of gutturals on neighbouring vowels in the corpus which were cited in Turcology books. In chapters four, five and six, the main research question was to explore the factors that might result in the residual effects. The main question in chapter four was to explore whether perceptual assimilation (Peperkamp et. Al, 2008) was the source of the residual effects of gutturals. In other words, do MST listeners categorize -/+ pharyngealized long and short Arabic vowels in nonce words into different native language categories (or not)? In chapter five, the main question was to test whether speakers of Turkish would yield similar results to those of the PAT experiment when presented with new non-borrowed Arabic words and non-words. In chapter six, the role of orthography in the matching of the responses to the observed patterns in the corpus was gauged. Hence, the hypotheses driving chapters four and five assumed that the Perceptual model would explain the corpus data whereas chapter six was based on the hypothesis that orthography played a role in the mappings. The answers to the three research questions are discussed below.

7.3 Results of chapter three: adaptation of the corpus data in Turkish

As was shown in chapter three, a new corpus was presented on the adaptation of ALT data into Turkish on the residual effects of gutturals on neighbouring vowels which conform to the findings reached in Turcology references including Tietze (1992), Stein (2006) and (Schaade) 1927. Twelve patterns were identified in the current work where six Arabic long vowels (plain and pharyngealized) are adapted as their short counterparts in Turkish in most cases, i.e., preserving phonological vowel quality where the presence or absence of the guttural in the source word does not affect the quality of the resulting Turkish vowel. That is, $[a:]_A$ and $[a:^ʕ]_A$>$/a/_T$ (where superscripted ^{ʕ} denotes pharyngealization or uvularization and > means adapted as); $[i:]_A$ and $[i:^ʕ]_A$>$/i/_T$; and $[u:]_A$ and $[u:^ʕ]_A$>$/u/_T$. Nevertheless, phonological vowel quality is not maintained in the adaptation of three short Arabic vowels since the presence and/or absence of gutturals in the Arabic source words determines the quality of the Turkish vowel. That is, $[a^ʕ]_A$>$/a/_T$, $[u^ʕ]_A$>$/u/_T$, $[i]_A$>$/i/_T$ but $[a]_A$>$/e/_T$, $[u]_A$>$/y/_T$ and $[i^ʕ]_A$>$/ɯ/_T$. Thus, in the adaptation of the three short vowels $[a^ʕ]_A$>$/a/_T$, $[u^ʕ]_A$>$/u/_T$ and $[i]_A$>$/i/_T$ where vowel quality does not change in the resulting Turkish word, we could establish that the preserved vowel quality is phonetic rather than phonemic since the resulting Turkish vowel is sensitive to the presence or absence of gutturals in the Arabic word.

The shortening of the Arabic long vowels in the loanwords can be explained by the Turkish phonotactic rule of banning long vowels, i.e., phonology (of the borrowing language) can account for the shortening of long vowels. This is in spite of the fact that some Arabic and Persian loanwords in modern Turkish have been cited to exhibit original vowel length or compensatory lengthening as explained in chapter 2 in 2.3.1. (Göksel & Kerslake, 2005). Some examples showing compensatory vowel lengthening are given in (1):

1) /a:/ /matˤbaʕah/$_A$>/matbaː/$_T$ *matbaa* 'press'
/u:/ /mawdˤu:ʕ/$_A$ >/mevzu:/$_T$ *mevzu* 'topic' (pronounced as *mevzuu*)
/i:/ /fiʕl/$_A$ >/fiil/$_T$ *fiil* 'verb'
/e:/ /taʔasuf/$_A$ >/teessyf/$_T$ *teessüf* 'sorrow'

We can argue that phonology is responsible for the adaptations of long vowels into Turkish since the presence or absence of gutturals in the source word, thus the resulting phonetic effects on the quality of the vowels themselves, does not affect these adaptations. Some examples reflecting this are shown in example (2):

2)	/dʒaːsuːs/$_A$	>/dʒasus/$_T$	*casus*	'spy'
	/ramadˤaːn/$_A$	>/ɾamazan/$_T$	r*amazan*/ramadan	'month of fasting for Muslims'
	/jatiːm/$_A$	>/jetim/$_T$	*yetim*	'orphan'
	/ðˤariːf/$_A$	>/zaɾif/$_T$	*zarif*	'grazeful'
	/ʔusluːb/$_A$	>/yslup/$_T$	*üslup*	'style'
	/sˤaːbuːn/$_A$	>/sabun/$_T$	*sabun*	'soap'

The results from the present corpus are the same as those mentioned in the Turcology references, where /a:/, /u:/ and /o:/ were adapted as /a/, /u/ and /o/, and /i:/ as /i/ regardless of the presence or absence of gutturals in the source word. This gives weight to the role of phonology in the adaptation of the Arabic loanwords since reference is to be made to phonemes rather than allophones. In other words, the influence of vowel quality in this case is phonological. In addition, the adaptation of the Arabic long vowels was shown to be also phonetically grounded in chapter three (see 3.5 in figure 3-2) based on the F1 and F2 properties of Arabic and Turkish vowels. This is because the Arabic long vowels (i.e. [a:]$_A$, [a:ˤ]$_A$, [i:]$_A$, [i:ˤ]$_A$, [u:]$_A$ and [u:ˤ]$_A$) perceptually appear closer to their short Turkish counterparts (i.e. /a/$_T$, /i/$_T$ and /u/$_T$) than to any other vowels on the vowel space. Figure 3-2 from chapter three is reproduced below as figure 7-1 for exposition purposes.

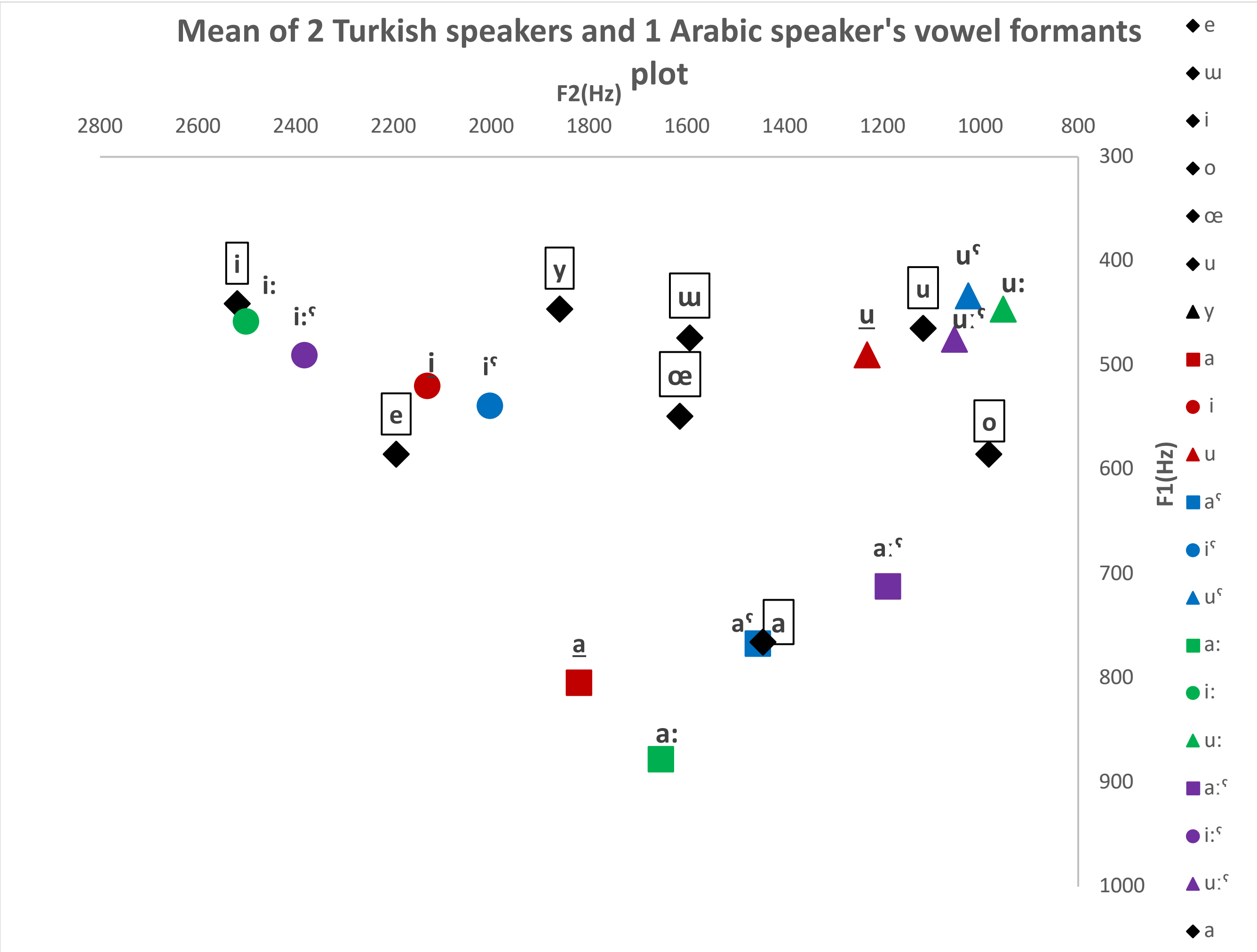


Figure 7-1: Vowel chart of Arabic and Turkish

Red = plain short Arabic vowels **green**= plain long Arabic vowels, **blue**= emphatic short Arabic vowels, **purple**= emphatic long Arabic vowels and **black** diamond=Turkish vowels; circles = [i], squares = [a], triangles = [u]

Regarding short vowels, the patterns found in the corpus also conform to those in the Turcology sources too (Tietze, 1992, Schaade, 1927 and Stein, 2006) in that the presence of emphatics/gutturals in the Arabic word affects the adaptation of vowels in Turkish. In the current work, where vowels were in the vicinity of gutturals in the Arabic cognate word, residual effects of the gutturals were detected in the resulting loanwords and the patterns were consistent. For instance, $[a^{ʕ}]_{A}$, $[u^{ʕ}]_{A}$ and $[i^{ʕ}]_{A}$ were adapted as /a/, /u/ and $/ɯ/_{T}$ respectively as in (3):

3) $/qurʕah/_{A}$>$/kuɾa/_{T}$ *kura* ‘drawing of lots’
$/rizq/_{A}$>$/ɾɯzk/_{T}$ *rizk* ‘earning a living’

/ruxsˤah/$_A$>/ɾuhsat/$_T$ *ruhsat* ‘licence’

The data in (3) suggest that vowel quality is not phonemic in the adaptation of the Arabic short vowels since the presence of gutturals in the source word affects the resulting Turkish vowel. Likewise, in the adaptation of the three other short vowels [a], [u] and [i] as /e/, /y/ and /i/, we find that the vowel quality of the resulting vowels is not all the same as those found in the source word when the guttural consonant is absent from the Arabic word. Some examples showing this are in given in (4).

4) /dars/$_A$>/ders/$_T$ *ders* ‘lesson’
/dʒumlah/$_A$>/dʒymle/$_T$ cümle ‘sentence’
/rasim/$_A$>/resim/$_T$ resim ‘picture’

From both (3) and (4), we can establish that the adaptation of short vowels is phonetic. This is due to the fact that the presence and/or absence of gutturals in the Arabic word determines the quality of the Turkish vowel. In this vein, we can refer to the F1/F2 values of both Arabic and Turkish in the vowel chart and check whether the adaptation of the corpus facts works. In chapter three (see 3.5), the relevant F1 (height) and F2 (backness) properties were explained based on the only data available to us nowadays, which is present day Turkish and Arabic.

We found from chapter three and we can see from figure 7-1 that most short vowel adaptations are phonetically grounded based on similarity in F1 and F2 values of Arabic and Turkish vowels, however, not all of them. Three cases diverge, namely the patterns [iˤ]$_A$>/ɯ/$_T$, [i]$_A$>/i/$_T$ and [u]>/y/$_T$ since they are closer to /e/$_T$, /e/$_T$ and /u/$_T$ respectively (figure 7-1). One way to account for these adaptations is by reference to the Turkish phonology, namely vowel harmony. Speakers of Turkish might have perceived plain and pharyngealized vowels as back-front contrasts. That is, they might have contrasted [a]$_A$>/e/$_T$ with [aˤ]$_A$>/a/$_T$, [i]$_A$>/i/$_T$ with [iˤ]$_A$>/ɯ/$_T$ and [u]$_A$>/y/$_T$ with [uˤ]$_A$>/u/$_T$. Nevertheless, we cannot assume the same theory for the adaptation of Arabic long vowels as no contrasts are evident.

The adaptation of another group of short vowels shows that the phonology of the source language (Arabic) might also be the source of the adaptation since the vowel category is preserved in the resulting Turkish word. These patterns include $[a^{ʕ}]_A$>$/a/_T$, $[u^{ʕ}]_A$>$/u/_T$ and $[i]_A$>$/i/_T$, the first two patterns of which were also found acoustically driven (figure 7-1).

7.3.1. Summary

Thus far, we argued that the adaptation of Arabic long vowels into Turkish is equally phonologically and phonetically motivated. That is, the presence or absence of gutturals in the Arabic source word does not affect the quality of the resulting Turkish vowels, and long Turkish and Arabic vowels have similar F1 and F2 properties. On the other hand, the adaptation of the short vowels is mostly phonetically driven since the presence or absence of gutturals in the Arabic source word influences the quality of the Turkish vowel. Three short Turkish vowels were found to be perceptually similar in their F1/F2 values to their Turkish counterparts (i.e. $[a]_A$>$/e/_T$, $[a^{ʕ}]_A$>$/a/_T$ and $[u^{ʕ}]$>$/u/_T$) whereas three other short vowels were found different from their counterparts in Turkish in either F1 or F2 values (i.e. $[i^{ʕ}]_A$>$/ɯ/_T$, $[i]_A$>$/i/_T$ and $[u]$>$/y/_T$). Moreover, phonological categories were preserved in three adapted patterns, namely $[a^{ʕ}]_A$>$/a/_T$, $[u^{ʕ}]_A$>$/u/_T$ and $[i]_A$>$/i/_T$ which means that the phonology of Arabic might also be the source of these adaptations.

These findings are important for loanword phonology and the ongoing debate of the two competing approaches; phonological and phonetic. We can say that in the adaptation of Arabic long vowels reference is to be made to both phonemes (i.e. phonological adaptation) and allophones (i.e. phonetic adaptation.) Conversely, in the adaptation of the Arabic short vowels reference is to be mainly made to allophones, i.e. phonetic approach but also to phonemes in certain cases. Therefore, the existence of phonological and perceptual explanations side by side equates to a hybrid model due to the existence of a mix of factors that interact in the loanword phonology.

Add to this that other factors such as orthography may have also played a role in the adaptation. This is due to the fact that in Arabic, long vowels are always reflected in the written form as ?alif (أ), yaa? (ي) and waaw (و). On the other hand, short vowels are only optionally represented (diacritics; vocalization) as fatHah (َ), Dhammah (ُ) and kasrah (ِ) which means that borrowers had a clear indication of what the vowel 'should' be, while for the short vowels they had to figure out the vowel quality from perception alone - allowing for the 'allophone'/phonetic influence here (see 2.2.3.). This might explain the mapping of $[u]_A$>$/y/_T$ and $[i]_A$>$/\text{ɯ}/_T$ compared to $[u:]_A$>$/u/_T$ and $[i:]_A$>$/i/_T$ in the corpus data especially that many of the words were borrowed via Persian and/or were probably in their written forms.

Since Osmanlica is no longer used and the adaptation process has ended, a correspondence theory such as Optimality Theory (OT), for example, is not used here, because we have no way to determine what type of input was used in the borrowing process. A number of questions remain regarding the mode of input of the borrowing (audio, written or both), who initiated the borrowing process (bilinguals or naïve listeners) and which factors affected the borrowing (language experience, age, context, vowel length, etc.). In order to answer these questions, the principle of uniformitarianism (Murray, 2015), as mentioned in chapter three (see 3.5.) was assumed and two perceptual experiments were conducted; Perceptual Assimilation Task and Simulated Borrowing.

Three research questions were formulated prior to conducting the experiments which are related to the mode of input, who performs the borrowing and which factors affect the borrowing process. First (RQ1), do MST listeners categorize -/+ emphatic Arabic vowels (long and short) into different Turkish vowel categories? Second, do MST listeners categorize -/+ emphatic vowels (long and short) into different Turkish vowel categories in real non-borrowed Arabic words and non-words? Third, does orthography play a role in the adaptation of Arabic vowels into Turkish?

In the next section, I review the results of the PAT experiment and compare them to those of the established loanwords.

7.4 *Results of chapter four: Perceptual mappings in the Perceptual Assimilation Task-Perception, phonology or both?*

It was found in chapter four that almost 70% of the perceptual assimilations *match* with the mappings in the corpus where ‘match’ is defined as perceptually assimilating to the same vowel as corpus when the stimulus material was nonsense words of the form hVd. Three patterns only were mismatched to the corpus data patterns; $[a{:}]_A$>$/e/_T$ (predicted $/a/_T$), $[i^ʕ]_A$>$/e/_T$ (predicted $/ɯ/_T$) and $[u]_A$>$/u/_T$ (predicted $/y/_T$). In addition, all listener groups shortened Arabic long vowels as was found in the corpus patterns in chapter three. However, they mapped one Arabic long vowel onto a Turkish vowel of a different quality $[a{:}]_A$>$/e/_T$.

Knowledge of Arabic phonology was not found to have any effect on the assimilation of 70% of the PAT patterns since all three groups of listeners reflected the same patterns. In fact, the monolingual T group manifested a slightly higher match percentage of 69.7% compared to 67.8% and 67.6% respectively for the TA (bilingual) and TQ (Quranic Turkish speakers) groups. It was also found from the logistic regression results that context (-/+emphasis), vowel quality and the interactions of *context* and *vowel quality*, on the one hand, and *length* and *vowel quality*, on the other, had significant effects on the vowel mappings. These significant effects together can be taken to answer RQ1 of how MST listeners classify -/+ pharyngealized long and short Arabic vowels in nonce words into different Turkish categories.

In what follows, I compare the PAT perceptual patterns to those found in the corpus. Column one in table 7-1 represents the source vowel categories, column two PAT perceptual maps and column three depicts predicted maps according to the corpus patterns. Moreover, column four indicates whether the most common perceptual assimilation pattern for each vowel is the same as in the corpus mapping, labeled as *as predicted*. The last column in the table refers to the average goodness of fit of the mappings in the PAT (since all listener groups yielded the same patterns) in percentile computed by dividing the average number of tokens mapped to the preferred vowel per individual in each group by the total number of listeners in each group multiplied by 100 and then averaging the resulting number across the three groups. For example, for the perceptual map $[a]_A$>$/e/_T$, the T group scored (98/123) 79.67%, TA scored (88/132) 66.66% and TQ group (335/429) 83.33%. The average score of the three groups is then taken which is 74.80% in the example.

Arabic vowel category	PAT perceptual maps	Predicted maps based on corpus patterns	As predicted	Average goodness of fit across listener groups in PAT
[i]$_A$	/i/$_T$	/i/$_T$	√	86.12%
[iˤ]$_A$	/e/$_T$	/ɯ/$_T$	X	56.23%
[i:]$_A$	/i/$_T$	/i/$_T$	√	91.17%
[i:ˤ]$_A$	/i/$_T$	/i/$_T$	√	89.26%
[a]$_A$	/e/$_T$	/e/$_T$	√	74.80%
[aˤ]$_A$	/a/$_T$	/a/$_T$	√	88.72%
[a:]$_A$	/e/$_T$	/a/$_T$	X	84.70%
[a:ˤ]$_A$	/a/$_T$	/a/$_T$	√	88.6%
[u]$_A$	/u/$_T$	/y/$_T$	X	78.16%
[uˤ]$_A$	/u/$_T$	/u/$_T$	√	90.87%
[u:]$_A$	/u/$_T$	/u/$_T$	√	90.50%
[u:ˤ]$_A$	/u/$_T$	/u/$_T$	√	87.79%

Table 7-1: PAT vowel patterns and their predicted categories based on corpus patterns in Turkish

As can be seen from table 7-1, only three patterns reflect variable mapping, namely [iˤ]$_A$>/e/$_T$, predicted as /ɯ/$_T$ (56.23%); [a]$_A$>/e/$_T$ predicted as /e/$_T$ (74.80%) and [u]$_A$>/u/$_T$ predicted as /y/$_T$ (78.16%). Of these, two patterns were perceptually assimilated to different vowels than those in the corpus data, by the listeners in the PAT experiment, [iˤ]$_A$>/e/ and [u]$_A$>/u/$_T$. Conversely, the pattern in [a]$_A$>/e/$_T$ accords with the predictions of the established words. Since mismatch occurs in two of these three variable mappings, this comes to no surprise. Nonetheless, most of the assimilated patterns were mapped consistently to their predicted categories in the corpus data and most of them mirror a resemblance to those categories. These include nine patterns; [a:]$_A$>/e/$_T$ (84.70%), [aˤ]$_A$>/a/$_T$ (88.72%), [a:ˤ]$_A$>/a/$_T$ (88.6%), [i:]$_A$>/i/$_T$ (91.17%), [i:ˤ]$_A$>/i/$_T$ (89.26%), [uˤ]$_A$>/u/$_T$ (90.87%), [u:ˤ]$_A$>/u/$_T$ (87.79%), /i/$_A$>/i/$_T$ (86.12%) and [u:]$_A$>/u/$_T$ (90.50%). The only pattern that stands out is [a:]$_A$>/e/$_T$, perceptually assimilated to the 'wrong' vowel near-categorically by the three groups with an average percentage of 84.70%. Two options can be entertained in this regard: phonology of the native language represented by vowel harmony (VH) in present day Turkish or a mix of perception, phonology and other factors.

Regarding the role of the native language, listeners might have been influenced by their Turkish VH rules in categorizing these variants into front and back vowels. If we assume this is the case, then the listeners might have treated Arabic plain $[a]_A$ and $[a:]_A$ as front vowels and pharyngelized $[a^ʕ]_A$ and $[a:^ʕ]_A$ as back vowels. The VH explanation may sound plausible; however, if the listeners really depended on their Turksh VH rules, then we may wonder why they did not uniformly use the same approach with the other Arabic vowel variants, i.e., $[i]_A$ and $[i:]_A$ versus $[i^ʕ]_A$ and $[i:^ʕ]_A$; and $[u]_A$ and $[u:]_A$ versus $[u^ʕ]_A$ and $[u:^ʕ]_A$. Instead, the three phonetic variants $[i]_A$, $[i:^ʕ]_A$, $[i:]_A$ were perceptually mapped onto $/i/_T$ but only $[i^ʕ]$ onto $/e/_T$ (predicted /ɯ/) and $[u]_A$, $[u:]_A$, $[u^ʕ]_A$ and $[u:^ʕ]_A$> $/u/_T$ (/u/ predicted to map to /y/). This clearly shows that the vowel harmony explanation, role of native language phonology, cannot account for the mismatched perceptual maps in the PAT experiment.

The alternative explanation and the one adopted here is that of a mix of perception and phonology. First, we need to revisit the results of the logistic regression from chapter four. It was found in table 4-16 and figure 4-8 in chapter 4 that two factors and two interactions were significant. These are *context*, *vowel quality* and the interactions of *context with vowel quality* and *vowel length with vowel quality*. These results are important in explaining the perceptual maps in the PAT itself and are key to understanding the findings within the current models of loanword phonology.

First of all, the only three perceptual maps that were *mismatched* by the listeners in the PAT experiment to their predicted vowel categories in the corpus were $[a:]_A$>$/e/_T$ (predicted as $/a/_T$), $[i^ʕ]_A$>$/e/_T$ (predicted as $/ɯ/_T$) and $[u]_A$>$/u/_T$ (predicted as $/y/_T$). The research question pursued in chapter four was how Turkish listeners categorize short and long -/+emphatic Arabic vowels into different Turkish categories.

The findings show that the listeners perceptually mapped the Arabic vowels onto vowels of the same quality in nine vowel categories (most cases) in Turkish, namely $[a^ʕ]_A$, $[a:^ʕ]_A$>$/a/_T$, $[i]_A$, $[i:]_A$ and $[i:^ʕ]_A$>$/i/_T$, and $[u]_A$, $[u^ʕ]_A$, $[u:]_A$ and $[u:^ʕ]_A$>$/u/_T$. This gives weight to the phonological model. However, listeners classify three short Arabic allophones as vowels of different categories in three cases, namely $[a:]_A$>$/e/_T$ (predicted as $/a/_T$), $[i^ʕ]_A$>$/e/_T$ (predicted as /ɯ/T) and [a]A>/e/T (as predicted), supporting the phonetic model. That is, Turkish listeners seem to treat the Arabic vowels as being phonetically sensitive to the presence and/or absence of gutturals (in other words the context) which, in turn, determines the quality of the resulting Turkish vowel.

In the logistic regression results, *context* was found very significant at the uvularized (guttural) level which probably refers to the perceptual mismatched map [iˤ]A>/e/T. This is also sustained by the significance of the interaction between context (uvularized) and vowel quality (vowel *i*) and the interaction between vowel length (short) and vowel quality (vowel *i*). Similarly, the mismatched perceptual map of [u]A>/u/T instead of /y/T might be the site referenced by the interaction between vowel length (short) and vowel quality (vowel *u*). The assimilation of the Arabic [a:]A as Turkish /e/T is not, however, clearly expressed in the results of the logistic regression.

The three vowel categories which diverge from those in the corpus mappings as they appear in table 7-1 can be tested against phonetic approximation and phonological approximation (see 2.2. in chapter 2). In order to do this, we need first to refer to the F1/F2 values of the Arabic and Turkish vowels on the vowel space as given in figure 7-2 below.

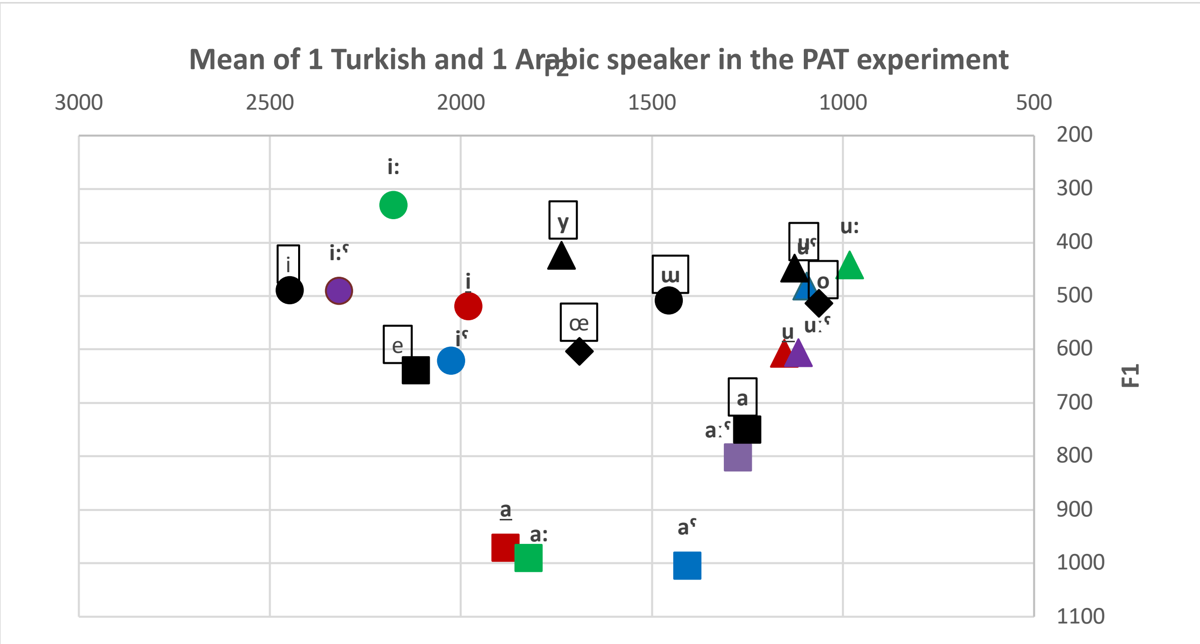


Figure 7-2: Vowel chart of Arabic and Turkish

Red = plain short Arabic vowels, **green**= plain long Arabic vowels, **blue**= emphatic short Arabic vowels, purple= emphatic long Arabic vowels and black diamond=Turkish vowels; circles = [i], squares = [a], triangles = [u]

Figure 7-2 illustrates that the two Arabic plain variants $[a]_A$ (in red square) and $[a:]_A$ (in green square) are positioned close to each other. The Arabic $[a:]_A$ appears (centrally) closer to Turkish $/e/_T$ than $/a/_T$ in terms of F2 (backness) although in height [F1] it is closer to $/a/_T$, meaning that this pattern is phonetically motivated as shown in 1) below.

1) Comparison of the formants of Arabic $[a]_A$ and $[a:]_A$ with Turkish $/a/_T$ and $/e/_T$

Arabic [a]	F2 1882	Turkish /e/	F2 2116	Turkish /a/	F2 1250
	F1 971		F1 640		F1 749
Arabic [a:]	F2 1821	Turkish /a/	F2 1250	Turkish /e/	F2 2116
	F1 990		F1 749		F1 640

From (1), we can see that the difference between Turkish $/e/_T$ {F2:2116} and Arabic $[a]_A$ {F2: 1882} is **234 hz** whereas the difference between Turkish /a/ {F2: 1250} and Arabic [a] {F2: 1882} is **632 hz**. This reveals that Turkish $/e/_T$ is closer to Arabic $[a]_A$ than Turkish $/a/_T$ is; thus, the percpetual mapping of $[a]_A$ to $/e/_T$ is based on phonetic proximity. As for the Arabic $[a:]_A$, the difference between Turkish $/e/_T$ {F2: 2116} and Arabic $[a:]_A$ {F2: 1821} is **295 hz** while the difference between Turkish $/a/_T$ {F2: 1250} and Arabic $[a:]_A$ {F2:1821} is **571 hz**, meaning that $[a:]_A$ is phonetically closer to Turkish $/e/_T$ than $/a/_T$. This means that present day Turkish speakers heard Arabic $[a:]_A$ as a front vowel in the PAT in contrast to the Arabic $[a:]_A$ variant in the corpus which seems to be a back vowel.

Phonologically, comparing the vowel categories in terms of their distinctive features (see chapter two), we can see as given in (2) that Arabic $[a:]_A$ and Turkish $/a/_T$ agree in height but not frontness/backness since $[a:]_A$ is a front vowel while Turkish $/a/_T$ is a back vowel. On the other hand, Arabic $[a:]_A$ and Turkish $/e/_T$ agree in both height and frontness, $[a:]_A$ being a low front vowel and Turkish $/e/_T$ being a mid-front vowel. Thus, the perceptual mapping $[a:]_A>/e/_T$ is phonologically (from 5)) as well as phonetically grounded (figure 7-2).

2) Comparison of feature specifications of Arabic $[a:]_A$ and Turkish $/a/_T$ and $/e/_T$

A. $[a:]_A>/a/_T$

Arabic [a:]	**>**	**Turkish /a/**
[-high]	√	[-high]

[+front]	X	[-front]

B. $[a:]_A > /e/_T$

Arabic [a:]	**>**	**Turkish /e/**
[-high]	√	[-high]
[+front]	√	[+front]

The mismatched pattern of $[i^ʕ]_A > /e/_T$ is phonetically grounded as shown in figure 7-2 and 3) below. This is clearly reflected by the F1 and F2 measurements given in 3) which show that $[i^ʕ]_A$ is perceptually closer in F2 and F1 to $/e/_T$ than $/i/_T$ in F2 and F1.

3) Comparison of the formants of Arabic $[i]_A$ and $[i^ʕ]$ and Turkish $/i/_T$ and $/e/_T$

Arabic [i] F2 1979 F1 519	Turkish /i/ F2 2448 F1 489	Turkish/e/ F2 2116 F1 640
Arabic $[i^ʕ]$F2 2024 F1 621	Turkish/e/ F2 2116 F1 640	Turkish /i/ F2 2448 F1 489

Comparing the vowel specifications of Arabic $[i^ʕ]_A$ to Turkish $/i/_T$ and $/e/_T$ in 4), we conclude the following. We can see that Arabic $[i^ʕ]_A$ and Turkish $/i/_T$ agree in height and frontness. On the other hand, Arabic $[i^ʕ]_A$ and Turkish $/e/_T$ are different in one feature, namely height since $[i^ʕ]_A$ is a high vowel while Turkish $/e/_T$ is a mid-vowel [-high]. Hence, we can establish that the perceptual map $[i^ʕ]_A > /e/_T$ is phonetically grounded (from 4 and figure 7-4)) but not phonologically motivated (from 7).

4) Comparison of Arabic $[i^ʕ]_A$ and Turkish $/i/_T$ and $/e/_T$ in terms of phonological features

A. $[i^ʕ]_A > /i/_T$

Arabic $[i^ʕ]$	**>**	**Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i^ʕ]_A > /e/_T$

Arabic $[i^ʕ]$	**>**	**Turkish /e/**
[+high]	x	[-high]

[+front]	√	[+front]

In the same line, the Arabic variant $[i]_A$, as shown in 3) is closer to Turkish $/e/_T$ than $/i/_T$ in F2 but to $/i/_T$ than $/e/_T$ in F1. However, it is perceptually mapped as Turkish $/i/_T$. From (5) below, we can see that $[i]_A$ is phonologically similar to the Turkish $/i/_T$ in the two features of height and frontness while it is similar to the Turkish $/e/_T$ only in height. This indicates that the perceptual map $[i]_A>/i/_T$ is phonologically and phonetically supported.

5) Comparison of Arabic $[i]_A$ and Turkish $/i/_T$ and $/e/_T$ in terms of phonological features

A. $[i]_A>/i/_T$

Arabic [i]	**>**	**Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i]_A>/e/_T$

Arabic [i]	**>**	**Turkish /e/**
[+high]	x	[-high]
[+front]	√	[+front]

The last perceptual map is $[u]_A>/u/_T$ predicted as $[u]_A>/y/_T$ according to the corpus mappings. Phonetically $/u/_T$ is acoustically closer than $/y/_T$ in F2 (backness) to $[u]_A$ whereas $/y/_T$ is closer in F1 (height) than $/u/_T$ to $[u]_A$ as illustrated in 6). In both cases, the mapping is phonetically grounded.

6) Comparison of the formants of Arabic $[u]_A$ and Turkish $/y/_T$ and $/u/_T$

Arabic [u] F2 1153 F1 607	Turkish/y / F2 1736 F1 423	Turkish /u/ F2 1126 F1 447

Next, we compare the three vowel categories in terms of their distinctive features. From 7), we can see that $[u]_A$ and $/u/_T$ agree in both height and frontness, meaning that the perceptual map $[u]_A>/u/_T$ is phonologically corroborated. As for the two vowels $[u]_A$ and $/y/_T$, we notice that they agree in height but not frontness which, in turn, means that this perceptual map (i.e. $[u]_A$>/y/$_T$) is not phonologically driven (from 7 below). Thus, we can establish that the perceptual map $[u]_A$>/u/$_T$ is both phonologically and phonetically sustained.

7) Comparison of feature specifications of $[u]_A$ and Turkish /u/$_T$ and /y/$_T$

A. $[u]_A$>/u/$_T$

Arabic [u]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. $[u]_A$>/y/$_T$

Arabic [u]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

7.4.1 Summary

Thus, we can construe that perception explains all three mismatched perceptual maps either in terms of F1 (height) or F2 (backness) or both. Additionally, in two cases Arabic phonology was also found to play a role in the assimilation since the Arabic vowel categories were preserved in the resulting Turkish words, i.e. $[a:]_A$>/e/$_T$ and $[u]_A$>/u/$_T$ but not in $[i^ʕ]_A$>/e/$_T$ which was only phonetically supported..

All in all, 70% of the PAT results match the corpus mappings, which already indicates that perception (alone) can account for a large proportion of the corpus mappings; the three mismatched cases confirms that these are cases where the PAT results mostly reflect the phonetics and in some cases also the phonology, so the corpus mappings must reflect something else. In the next two sections, we explore the potential role of lexical/phonological knowledge of the language (ch5) and of orthography (ch6).

7.5 Results of chapter five: Perceptual mappings in the Simulated Borrowing experiment- audio-only (SB-A)

In chapter five we saw that the percentage of perception dropped to almost 50% (48% to be precise). In the audio task of the Simulated Borrowing experiment (henceforth SB-A), real Arabic and nonsense words were used in contrast to the *hVd* stimuli in the PAT. Moreover, the listeners were not restricted to the eight Turkish vowels but rather wrote their responses in Turkish spelling, resulting in more mappings which at times introduced loan vowels (e.g. long vowels). The research question explored in chapter five was whether Turkish speakers would yield similar results to those in the PAT experiment and, in turn, to the corpus data when the stimuli were real non-borrowed Arabic and nonsense words. The main finding of the chapter is that the bilingual TA group exhibited the highest degrees of match (53%) to the corpus patterns making them the closest to the perception of the speakers of Osmanlica of Arabic loanwords, if we assume the uniformitarian principle. This, in turn, suggests that the speakers of Osmanlica too, as original borrowers of Arabic loan words, were bilingual.

Compared to the PAT experiment, the listener groups yielded different perceptual maps to each other which indicates that listener group plays a role in the mappings when the stimuli were real monosyllabic words. This result was also sustained in the findings of the logistic regression in chapter five where the interaction of the TQ group was found near significant at the plain level and at the vowel quality ***i*** level; **ListgpTQ:contextplain, contextplain:vowel.quality*i*** and **ListgpTQ:vowel.qualityi**. In addition, the interaction of the TA group at the vowel quality ***i*** level was also found very significant (**ListgpTA:vowel.quality*i***). These results combined show that i) listener group had a (near)/significant effect in the interactions with context and vowel quality albeit not a main effect and ii) the interactions involving groups that know Arabic were found significant compared to the monolingual T group.

In what follows, I adopt the same approach I used when discussing the results of the PAT experiment. In other words, I first compare the perceptual maps of each group to their predicted categories in the corpus. Table 7-2 compares the results of the T group to those of the corpus mappings. The average goodness of fit in the last column indicates the number of each listener with the correct responses in each listener group divided by the total number of the listeners in each particular group. For instance, in the table below 15 out of 18 T listeners mapped $[i]_A$ onto $/i/_T$. We divide 15 into 18 and multiply the product by 100 to get the goodness of fit, which in this case is 83.33%.

Arabic vowel category	SB-T perceptual maps	Predicted maps based on corpus patterns	As predicted	Average goodness of fit in the T group in SB-A
[i]	$/i/_T$	$/i/_T$	√	83.33%
$[i]_{d^ʕ}$	$/i/_T$	$/ɯ/_T$	X	72.22%
$[i]_q$	$/i/_T$	$/ɯ/_T$	X	77.77%
$[i]_ʕ$	$/e/_T$	$/i/_T$	X	72.22%
[i:]	$/i/_T$	$/i/_T$	√	100%
$[i:]_{d^ʕ}$	$/i/_T$	$/i/_T$	√	66.66%
$[i:]_q$	$/ɯ/_T$	$/i/_T$	X	77.77%
$[i:]_ʕ$	$/ɯ/_T$	$/i/_T$	X	88.88%
[a]	$/e/_T$	$/e/_T$	√	83.33%
$[a]_{d^ʕ}$	$/a/_T$	$/a/_T$	√	72.22%
$[a]_q$	/a/	/a/	√	88.88%
$[a]_ʕ$	/a/	/a/	√	83.33%
[a:]	$/e/_T$	$/a/_T$	X	61.11%
$[a:]_{d^ʕ}$	$/a/_T$	$/a/_T$	√	100%
$[a:]_q$	$/a/_T$	$/a/_T$	√	100%
$[a:]_ʕ$	$/a/_T$	$/a/_T$	√	100%
[u]	$/u/_T$	$/y/_T$	X	61.11%
$[u]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	83.33%
$[u]_q$	$/u/_T$	$/u/_T$	√	83.33%
$[u]_ʕ$	$/o/_T$	$/u/_T$	X	61.11%
[u:]	$/u/_T$	$/u/_T$	√	94.44%
$[u:]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	94.44%
$[u:]_q$	$/u/_T$	$/u/_T$	√	100%
$[u:]_ʕ$	$/u/_T$	$/u/_T$	√	94.44%

Table 7-2: SB vowel patterns and their predicted categories for the T group based on corpus patterns in Turkish

In table 7-2 above, we notice that **eight** categories out of **twenty-four** categories are not mapped by the T group onto their predicted categories in the corpus; however, the remaining **sixteen** categories are correctly perceived as their respective categories. The mismatched responses include two perceptual maps (out of eight) in the plain environment which were mapped ‘incorrectly’ by all listeners in the PAT experiment. These are [a:]>/e/$_{T}$ (predicted as /a/$_{T}$) and [u]>/u/$_{T}$ (predicted as /y/$_{T}$) both of which were assimilated variably with a percentage of 61.11%. In addition, six perceptual maps include uvularized vowels. Of these, three guttural consonants surround [i] in the Arabic word, i.e. [i]$_{d^{ʕ}}$>/i/$_{T}$ (predicted as /ɯ/$_{T}$), [i]$_{q}$>/i/$_{T}$ (predicted as /ɯ/$_{T}$) and [i]$_{ʕ}$>/e/$_{T}$ (predicted as /i/$_{T}$); two guttural consonants including q and pharyngeals but not emphatics surround [i:], i.e. [i:]$_{q}$>/ɯ/$_{T}$ (predicted as /i/$_{T}$) and [i:]$_{ʕ}$>/ɯ/$_{T}$ (predicted as /i/$_{T}$) in addition to the pharyngeal in the proximity of [u], i.e. [u]$_{ʕ}$ >/o/$_{T}$ (predicted as /u/$_{T}$). Of the above eight categories, listeners in the T group categorically mapped [i:]$_{ʕ}$ as /ɯ/$_{T}$ (predicted as /i/$_{T}$) with a goodness of fit of 88.88% whereas they perceived the remaining seven categories inconsistently.

As for the bilingual group, the vowel patterns and their predicted categories are provided in table 7-3 below.

Arabic vowel category	**SB-TA perceptual maps**	**Predicted maps based on corpus patterns**	**As predicted**	**Average goodness of fit in the TA group in SB-A**
[i]	/i/$_{T}$	/i/$_{T}$	√	90.90%
[i]$_{d^{ʕ}}$	/ɯ/$_{T}$	/ɯ/$_{T}$	√	54.54%[36]
[i]$_{q}$	/i/$_{T}$	/ɯ/$_{T}$	X	72.72%
[i]$_{ʕ}$	/i/$_{T}$	/i/$_{T}$	√	54.54%
[i:]	/i/$_{T}$	/i/$_{T}$	√	90.90%
[i:]$_{d^{ʕ}}$	/i/$_{T}$	/i/$_{T}$	√	54.54%
[i:]$_{q}$	/i/$_{T}$	/i/$_{T}$	√	63.63%
[i:]$_{ʕ}$	/ɯ/$_{T}$	/i/$_{T}$	X	63.63%
[a]	/e/$_{T}$	/e/$_{T}$	√	81.81%
[a]$_{d^{ʕ}}$	/a/$_{T}$	/a/$_{T}$	√	81.81%
[a]$_{q}$	/a/	/a/	√	81.81%
[a]$_{ʕ}$	/a/	/a/	√	54.54%
[a:]	/e/$_{T}$	/a/$_{T}$	X	72.72%
[a:]$_{d^{ʕ}}$	/a/$_{T}$	/a/$_{T}$	√	90.90%
[a:]$_{q}$	/a/$_{T}$	/a/$_{T}$	√	100%
[a:]$_{ʕ}$	/a/$_{T}$	/a/$_{T}$	√	100%

[36] The remaining 45.46% of the mappings include other vowel categories as shown in table 7-6. The same applies to the mappings of [i]$_{ʕ}$, [i:]$_{d^{ʕ}}$, [a]$_{ʕ}$, and [u]$_{d^{ʕ}}$.

[u]	/u/$_T$	/y/$_T$	X	81.81%
[u]$_{d^ʕ}$	/u/$_T$	/u/$_T$	√	54.54%
[u]$_q$	/u/$_T$	/u/$_T$	√	90.90%
[u]$_ʕ$	/u/$_T$	/u/$_T$	√	63.63%
[u:]	/u/$_T$	/u/$_T$	√	100%
[u:]$_{d^ʕ}$	/u/$_T$	/u/$_T$	√	100%
[u:]$_q$	/u/$_T$	/u/$_T$	√	100%
[u:]$_ʕ$	/u/$_T$	/u/$_T$	√	100%

Table 7-3: SB vowel patterns and their predicted categories for the TA group based on corpus patterns in Turkish

In the table above, we can see that **four** perceptual maps out of **twenty-four** are not perceived as predicted. These are [i]$_q$>/i/$_T$ (predicted as /ɯ/$_T$), [i:]$_ʕ$>/ɯ/$_T$ (predicted as /i/$_T$), [a:]>/e/$_T$ (predicted as /a/$_T$) and [u]>/u/$_T$ (predicted as /y/$_T$). Hence, both the T and TA group so far incorrectly perceived the two perceptual maps [a:]>/e/$_T$ and [u]>/u/$_T$ different than the corpus mappings yet similar to the PAT's. Regarding the uvularized vowels, the TA group categorized [i]$_q$ and [i:]$_ʕ$ onto the 'wrong' category, however, matched [i]$_{d^ʕ}$ (54.54%), [i:]$_{d^ʕ}$ (54.54%), [i:]$_q$ (63.63%) and [i]$_ʕ$ (54.54%) as predicted- although the tokens are small- making their perception the closest so far to the perception of Osmanlica speakers' of Arabic loanwords. Next, I compare the vowel patterns of the TQ group in the SB-A to those of the corpus in table 7-4.

Arabic vowel category	**SB-TQ perceptual maps**	**Predicted maps based on corpus patterns**	**As predicted**	**Average goodness of fit in the T group in SB-TQ**
[i]	/i/$_T$	/i/$_T$	√	86.36%
[i]$_{d^ʕ}$	/ɯ/$_T$	/ɯ/$_T$	√	50%[37]
[i]$_q$	/i/$_T$	/ɯ/$_T$	X	77.27%
[i]$_ʕ$	/e/$_T$	/i/$_T$	X	59.09%
[i:]	/i/$_T$	/i/$_T$	√	95.45%
[i:]$_{d^ʕ}$	/i/$_T$	/i/$_T$	√	59.09%
[i:]$_q$	/ɯ/$_T$	/i/$_T$	X	59%
[i:]$_ʕ$	/ɯ/$_T$	/i/$_T$	X	68.18%
[a]	/e/$_T$	/e/$_T$	√	86.36%
[a]$_{d^ʕ}$	/a/$_T$	/a/$_T$	√	86.36%

[37] The other 50% include other vowel categories. See table 7-7. The same applies to the mappings of [u]$_{d^ʕ}$>/u/$_T$ and [u]$_ʕ$>/u/$_T$.

$[a]_q$	/a/	/a/	√	95.45%
$[a]_ʕ$	/a/	/a/	√	90.90%
[a:]	$/e/_T$	$/a/_T$	X	90.90%
$[a:]_{d^ʕ}$	$/a/_T$	$/a/_T$	√	100%
$[a:]_q$	$/a/_T$	$/a/_T$	√	100%
$[a:]_ʕ$	$/a/_T$	$/a/_T$	√	95.45%
[u]	$/u/_T$	$/y/_T$	X	53.57%
$[u]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	50%
$[u]_q$	$/u/_T$	$/u/_T$	√	100%
$[u]_ʕ$	$/u/_T$	$/u/_T$	√	50%
[u:]	$/u/_T$	$/u/_T$	√	95.45%
$[u:]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	95.45%
$[u:]_q$	$/u/_T$	$/u/_T$	√	100%
$[u:]_ʕ$	$/u/_T$	$/u/_T$	√	95.45%

Table 7-4: SB vowel patterns and their predicted categories for the TQ group based on corpus patterns in Turkish

The TQ group yielded **six** mismatched perceptual patterns and eighteen matched perceptual maps to those in the corpus. The mismatched ones include the two perceptual maps of $[a:]>/e/_T$ (predicted as $/a/_T$) 90.90% and $[u]>/u/_T$ (predicted as $/y/_T$) 53.57% which are similar to the mappings in the PAT and in the T and TQ mappings in the SB-A. The other four maps include the uvularized mappings of the short and long [i] in the vicinity of q and pharyngeals but not emphatics, i.e. $[i]_q$, $[i:]_q$, $[i]_ʕ$ and $[i:]_ʕ$. Tables 7-2 to 7-4 show that the perceptions of the two categories that know Arabic are closer to that of the speakers of Osmanlica in adapting Arabic loanwords with the bilingual group being even closer.

As I did in the discussion of the PAT results, here too I resort to the phonetic and phonological approaches to determine the source of the mappings of the three listener groups as presented in table 7-2 to 7-4. In order to do this, I first examine the vowel space of Arabic and Turkish vowels in real words in figure 7-3 and compare the distinctive features of Arabic and Turkish phonemes. I focus my attention on all the perceptual maps that were mismatched (not as predicted) by the three groups collectively (eight perceptual maps), i.e. $[a:]_A>/e/_T$ (predicted as $/a/_T$), $[u]_A>/u/_T$ (predicted as $/y/_T$), $[i]_{d^ʕ}>/i/_T$ (predicted as $/ɯ/_T$), $[i]_q>/i/_T$ (predicted as $/ɯ/_T$), $[i]_ʕ>/e/_T$ (predicted as $/i/_T$), $[i:]_q>/ɯ/_T$ (predicted as $/i/_T$) and $[i:]_ʕ>/ɯ/_T$ (predicted as $/i/_T$) and $[u]_ʕ >/o/_T$ (predicted as $/u/_T$).

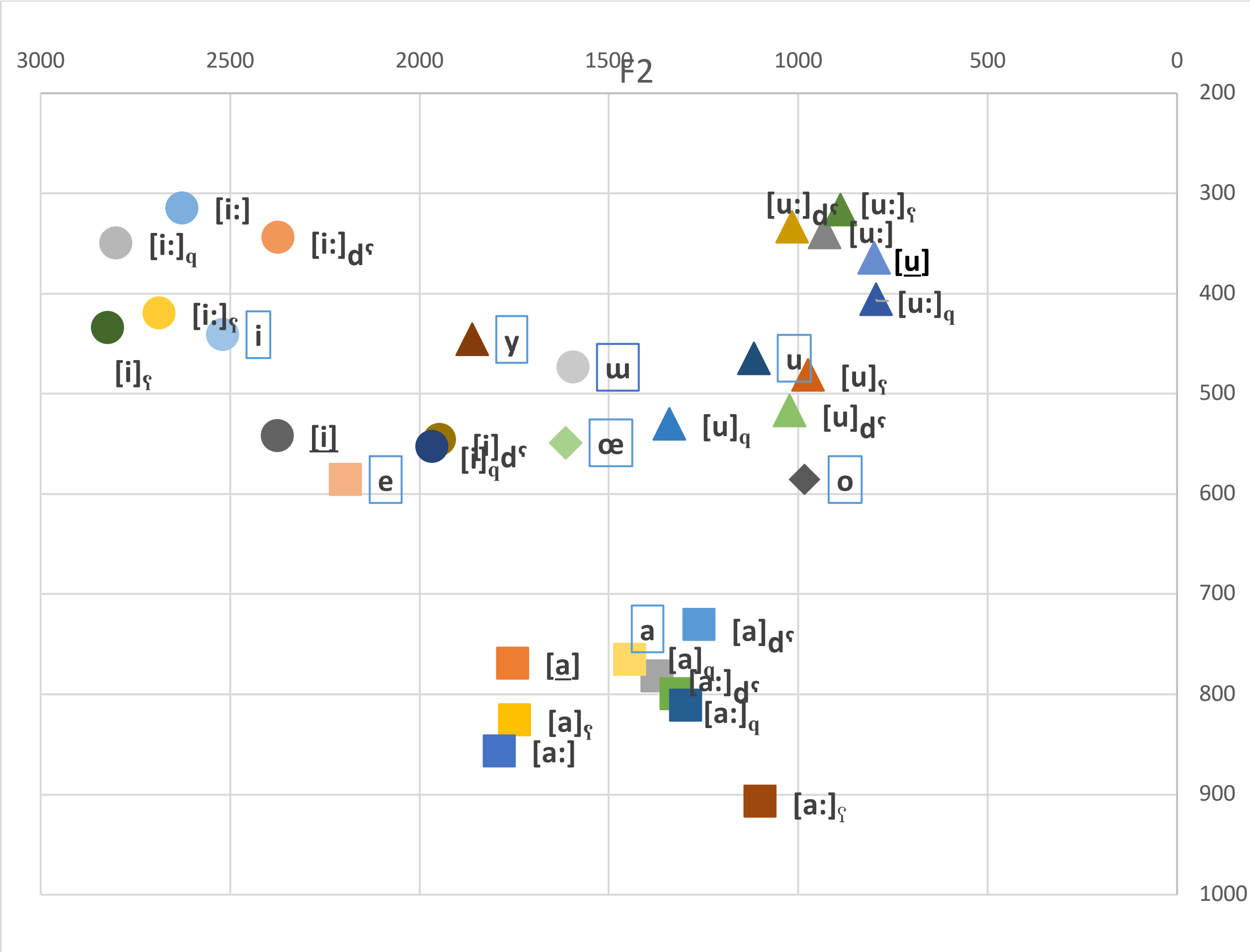


Figure 7-3: Mean frequency values of 2 Turkish speakers and 1 Arabic speaker's vowel formants plot (real words) in the SB-A.

Vowels in boxes are Turkish vowels and the rest are Arabic ones; underlined /a/, /i/and /u/ are Arabic short vowels. Diamond=Turkish vowels [o] and [œ]; circles = [i] and [ɯ], squares = [a] and [e], triangles = [u] and [y]

As in the PAT experiment, all three listener groups mapped $[a:]_A$>$/e/_T$ instead of $/a/_T$. As can be seen from figure 7-3, $[a:]_A$ and $[a]_A$ are phonetically close to each other and are located centrally between $[e]_T$ and $[a]_T$. Both are acoustically closer in terms of F1 {height} and F2 {backness} to $[a]_T$ based on the measurements in 8) below. This shows that this perceptual map is not phonetically grounded in the SB-A in contrast to what we found in the discussion on the PAT experiment.

8) Comparison of the formants of Arabic $[a]_A$ and $[a:]_A$ with Turkish $/a/_T$ and $/e/_T$

Arabic [a]	F2 1754	Turkish /e/	F2 2194	Turkish /a/	F2 1445
	F1 769		F1 587		F1 766
Arabic [a:]	F2 1789	Turkish /a/	F2 1445	Turkish /e/	F2 2194
	F1 857		F1 766		F1 587

Comparing the same vowel categories in terms of distinctive features (see chapter two for distinctive feature of both Arabic and Turkish) in 9), we can see that $[a:]_A$ and $/e/_T$ agree in both height and frontness whereas $[a:]_A$ and $/e/_T$ agree only in height. This shows that the perceptual map $[a:]_A > /e/_T$ is phonologically determined in the SB-A.

9) Comparison of feature specifications of Arabic $[a:]_A$ and Turkish $/a/_T$ and $/e/_T$

A. $[a:]_A > /a/_T$

Arabic [a:]	**>**	**Turkish /a/**
[-high]	√	[-high]
[+front]	X	[-front]

B. $[a:]_A > /e/_T$

Arabic [a:]	**>**	**Turkish /e/**
[-high]	√	[-high]
[+front]	√	[+front]

The second perceptual map we compare is $[u]_A > /u/_T$ in 10) and 11) below. As can be seen from 10), $/u/_T$ is closer to $[u]_A$ in F2 {backness } than $/y/_T$ is; however, $/y/_T$ is closer in F1 {height} than $/u/_T$ is to $[u]_A$. In both cases, the perceptual mapping is phonetically grounded.

10) Comparison of the formants of Arabic $[u]_A$ and Turkish $/y/_T$ and $/u/_T$

Arabic [u]	F2 800	Turkish/y/	F2 1860	Turkish /u/	F2 1117
	F1 364		F1 446		F1 464.5

As for comparing the distinctive features of Arabic $[u]_A$ and Turkish $/u/_T$ and $/y/_T$, this is given in 11) below. Phonologically, $[u]_A$ and $/u/_T$ agree in height and backness since both are high and back vowels in both Arabic and Turkish. On the other hand, $[u]_A$ and $/y/_T$ agree only in one feature which is height. This makes $/u/_T$ phonologically closer to $[u]_A$ than $/y/_T$ is. Thus, from figure 7-3, 10) and 11) we can that the perceptual map $[u]_A > /u/_T$ is both phonetically and phonologically motivated.

11) Comparison of feature specifications of $[u]_A$ and Turkish $/u/_T$ and $/y/_T$

A. $[u]_A$>$/u/_T$

Arabic [u]	>	Output Turkish /u/
[+high]	√	[+high]
[-front]	√	[-front]

B. $[u]_A$>$/y/_T$

Arabic [u]	>	Output Turkish /y/
[+high]	√	[+high]
[-front]	X	[+front]

Next, I phonetically and phonologically inspect the uvularized categories which were not assimilated to their predicted categories. These are given below in 12- 23.

12) Comparison of the formants of Arabic $[i]_{d^ʕ}$ and Turkish $/i/_T$, $/ɯ/_T$ and $/e/_T$

Arabic $[i]_{d^ʕ}$	Turkish/i/	Turkish /ɯ/	Turkish/e/
F2 1946	F2 2519.5	F2 1594	F2 2194
F1 546	F1 441	F1 473.5	F1 586

The two uvularized vowels $[i]_{d^ʕ}$ and $[i]_q$ are perceived in the same way by the listeners of the T group but not the TA and TQ groups who map $[i]_{d^ʕ}$ onto its predicted category $/ɯ/_T$. In 12) and 14) below we see that Turkish $/ɯ/_T$ is closer in backness {F2} and height {F1} to Arabic **[i]**$_{d^ʕ}$ and **[i]**$_{\mathbf{q}}$ than $/i/_T$ is despite the fact that most T listeners categorized **[i]**$_{d^ʕ}$ and **[i]**$_{\mathbf{q}}$ as $/i/_T$ instead of $/ɯ/_T$. Thus, we can safely maintain that the perceptual maps **[i]**$_{d^ʕ}$ and **[i]**$_{\mathbf{q}}$ as $/i/_T$ are not phonetically supported. Having said this, as a matter of fact, the Turkish vowel /e/ is phonetically closer in F1 and F2 than both $/ɯ/_T$ and $/i/_T$; however, the listeners did not choose it as their response. This strongly suggests that the listeners did not depend on the phonetic details of the vowel categories when categorizing the Arabic **[i]**$_{d^ʕ}$ and **[i]**$_{\mathbf{q}}$. This gives weight to the phonological argument as shown in 13) and 15).

13) Comparison of feature specifications of $[i]_{d^ʕ}$ and Turkish $/i/_T$, $/ɯ/_T$ and $/e/_T$

A. $[i]_{d^ʕ}$ >$/i/_T$

Arabic $[i]_{d^ʕ}$	>	Turkish /i/
[+high]	√	[+high]

[+front] √ [+front]

B. $[i]_{d^ʕ}$ >$/ɯ/_T$

Arabic $[i]_{d^ʕ}$	**>**	**Turkish /ɯ/**
[+high]	√	[+high]
[+front]	x	[-front]

C. $[i]_{d^ʕ}$ >$/e/_T$

Arabic $[i]_{d^ʕ}$	**>**	**Turkish /e/**
[+high]	x	[-high]
[+front]	√	[+front]

Comparing the phonological features of the vowel categories $[i]_{d^ʕ}$, $[i]_q$, $/i/_T$, $/ɯ/_T$ and $/e/_T$, we find that the perceptual maps **$[i]_{d^ʕ}$ >$/i/_T$** and **$[i]_q$ >$/i/_T$** are phonologically determined. This is because $[i]_{d^ʕ}$/$[i]_q$ and $/i/_T$ agree in the two features of height and frontness. Conversely, $[i]_{d^ʕ}$/ $[i]_q$ and $/ɯ/_T$ agree only in one phonological feature which is height but not frontness/backness and $[i]_{d^ʕ}$/$[i]_q$ and $/e/_T$ agree only in frontness but not height.

14) Comparison of the formants of Arabic $[i]_q$ and Turkish $/i/_T$, $/ɯ/_T$ and $/e/_T$

Arabic **$[i]_q$**	Turkish/i/	Turkish /ɯ/	Turkish/e/
F2 1966	F2 2519.5	F2 1594	F2 2194
F1 553	F1 441	F1 473.5	F1 586

15) Comparison of feature specifications of $[i]_q$ and Turkish $/i/_T$ and $/ɯ/_T$

A. $[i]_q$ >$/i/_T$

Arabic $[i]_q$	**>**	**Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i]_q$ >$/ɯ/_T$

Arabic $[i]_q$	**>**	**Turkish /ɯ/**
[+high]	√	[high]
[+front]	x	[-front]

C. $[i]_q$ >$/e/_T$

Arabic $[i]_q$	**>**	**Turkish /e/**

[+high]	x	[-high]
[+front]	√	[+front]

The perceptual map **[i]**$_ʕ$ >/e/$_T$ instead of /i/$_T$ was yielded by the listeners of the T and TQ groups only but not the TA group The phonetic comparison of the formants' measurements of the three vowel categories is given in 16) whereas the phonological comparison of the distinctive features is given in 17).

16) Comparison of the formants of Arabic [i]$_ʕ$ and Turkish /e/$_T$ and /i/$_T$

Arabic **[i]**$_ʕ$	Turkish /e/	Turkish /i/
F2 2822	F2 2194	F2 2519.5
F1 434	F1 586	F1 441

From figure 7-3, we can see that **[i]**$_ʕ$ is phonetically in closer proximity to /i/$_T$ than /e/$_T$. This observation is supported by the measurements in 16) where /i/$_T$ is closer in both F1 and F2 to **[i]**$_ʕ$ than /e/$_T$ is. Therefore, the mapping **[i]**$_ʕ$ >/e/$_T$ is not phonetically supported. Interestingly, /e/$_T$ is phonologically not the closest candidate either as shown in 17). This is because /e/$_T$ and [i]$_ʕ$ agree only in one feature which is frontness but not height. Conversely, /i/$_T$ and [i]$_ʕ$ agree in both height and frontness. Thus, we can say that **[i]**$_ʕ$ >/e/$_T$ neither is phonetically nor phonologically (phonology of Arabic) motivated. One possible source of this mapping is that of vowel harmony of Turkish. That is, the listeners might have been influenced by the phonology of their native language and applied the front-back distinction in the mapping of **[i]**$_ʕ$ because they could hear the lowering effect of F2 of the guttural /ʕ/ in the Arabic cognate and translated it into a backer vowel /e/$_T$. Thus, here we can say that the source of the mapping is Turkish phonology.

17) Comparison of feature specifications of Arabic [i]$_ʕ$ and Turkish /e/$_T$ and /i/$_T$

A. [i]$_ʕ$ >/e/$_T$

Arabic [i]$_ʕ$	**>**	**Turkish /e/**
[+high]	X	[+high]
[+front]	√	[+front]

B. [i]$_ʕ$ >/i/$_T$

Arabic [i]$_ʕ$	>	Turkish /i/
[+high]	√	[+high]
[+front]	√	[+front]

The listeners in the three groups mismatched [i:]$_ʕ$ to /ɯ/$_T$ instead of the predicted /i/$_T$ whereas only the listeners in the T and TQ groups mapped the [i:]$_q$ onto /ɯ/$_T$ instead of /i/$_T$. Phonetically, /i/$_T$ is closer to both [i:]$_ʕ$ and [i:]$_q$ than /ɯ/$_T$ is in F1 and F2 as shown in 18) and 20), meaning that the perceptual maps [i:]$_ʕ$ >/ɯ/$_T$ and [i:]$_q$>/ɯ/$_T$ are not phonetically supported. Moreover, comparing the phonological features of the vowel categories reveals that /i/$_T$ is more similar to both [i:]$_ʕ$ and [i:]$_q$ in the two dimensions of height and frontness as illustrated in 22) and 24). Here too, the the only remaining explanation is that of vowel harmony of Turkish. That is, the listeners might have interpreted the F2 lowering effect of both /q/ and /ʕ/ as a backing effect similar to their Turkish front-back distinction.

18) Comparison of the formants of Arabic [i:]$_q$ and Turkish /ɯ/$_T$ and /i/$_T$

Arabic **[i:]**$_q$	Turkish /ɯ/	Turkish/i/
F2 2800	F2 1594	F2 2519.5
F1 350	F1 473.5	F1 441

19) Comparison of feature specifications of [i:]$_q$ and Turkish /i/$_T$ and /ɯ/$_T$

A. [i:]$_q$ >/i/$_T$

Arabic [i:]$_q$	**>**	**Turkish /ɯ/**
[+high]	√	[high]
[+front]	x	[-front]

B. [i:]$_q$ >/ɯ/$_T$

Arabic [i:]$_q$	**>**	**Turkish /i/**
[+high]	√	[high]
[+front]	√	[-front]

20) Comparison of the formants of Arabic [i:]$_ʕ$ and Turkish /ɯ/$_T$ and /i/$_T$

Arabic **[i:]**$_ʕ$	Turkish /ɯ/	Turkish /i/
F2 2687	F2 1594	F2 2519.5
F1 420	F1 473.5	F1 441

21) Comparison of feature specifications of Arabic [i:]$_ʕ$ and Turkish /ɯ/$_T$ and /i/$_T$

A. [i:]$_{ʕ}$ >/ɯ/$_{T}$

Arabic [i:]$_{ʕ}$	**>**	**Turkish /ɯ/**
[+high]	√	[+high]
[+front]	X	[+front]

B. [i:]$_{ʕ}$ >/i/$_{T}$

Arabic [i]$_{ʕ}$	**>**	**Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

The last perceptual map to discuss is [u]$_{ʕ}$>/o/$_{T}$ instead of /u/$_{T}$ yielded by the monolingual listeners only. Based on the formants values in 22) the vowel /o/$_{T}$ is phonetically more similar [u]$_{ʕ}$ in F2 whereas /u/$_{T}$ is closer to [u]$_{ʕ}$ in F1. This means that the mapping of [u]$_{ʕ}$ to either /o/$_{T}$ or /u/$_{T}$ is phonetically grounded. As for the phonological similarity, we can see from 23) that [u]$_{ʕ}$ and /u/$_{T}$ are similar in both features of height and backness whereas [u]$_{ʕ}$ and /o/$_{T}$ are similar only in the feature of backness but not height. Therefore, the perceptual map [u]$_{ʕ}$>/o/$_{T}$ is only phonetically driven but not phonologically. Noteworthy to mention is that only the monolingual group perceived [u]$_{ʕ}$ as /o/$_{T}$ the latter of which is not an Arabic vowel. In the SB-A experiment, the listeners were specifically asked to write the vowel they heard in Turkish spelling. This may suggest that the TA and TQ groups, who assimilated [u]$_{ʕ}$ as its predicted category /u/$_{T}$, depended on their knowledge of Arabic phonology in addition to their perception in Turkish which gives weight to the role of Arabic phonology.

21) Comparison of the formants of Arabic [u]$_{ʕ}$ and Turkish /o/$_{T}$ and /u/$_{T}$

Arabic [u]$_{ʕ}$ F2 974	Turkish /o/ F2 983	Turkish /u/ F2 1117
F1 481	F1 586	F1 464.5

22) Comparison of feature specifications of Arabic [u]$_{ʕ}$ and Turkish /o/$_{T}$ and /u/$_{T}$

A. [u]$_{ʕ}$ >/o/$_{T}$

Arabic [u]$_{ʕ}$	**>**	**Turkish /o/**
[+high]	X	[-high]
[-front]	√	[-front]

B. $[u]_{ʕ}$ >$/u/_{T}$

Arabic $[u]_{ʕ}$	**>**	**Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

7.5.1 Summary

All in all, in the SB-A experiment the three listener groups manifested different perceptual maps on the contrary to the PAT where all groups had the same mappings. The two groups which knew Arabic reflected a closer perception to that of the corpus with the bilingual group having even a closer perception of 53%, TQ 47% and T 45%. These results were reflected in the logistic regression findings in the interactions of these two groups with the uvularized and plain contexts and the vowel quality i.

Hence, these results together mean that perception alone cannot be responsible for all the mappings in the SB-A experiment and, in turn, in the corpus data. It was also demonstrated that the phonology of Arabic (knowledge of Arabic) and the phonology of Turkish both played a role in the mappings. In the next section, we explore the potential role of orthography in addition to phonology and phonetics.

7.6 Results of chapter six: Perceptual mappings in the Simulated Borrowing experiment Audio, Audio-written and Written tasks- role of orthography

In chapter six, the research question pursued was whether knowledge of Arabic orthography would have an effect on the degrees of match of the source words to their predicted categories in the corpus data. The main finding was that the two groups with knowledge of Arabic manifested higher degrees of match to the corpus patterns at 61.9% in the audio-written condition than the audio one (48.44%). This result indicates that orthography too has an effect in addition to perception on the categorization, which, in turn, gives weight to a hybrid model of phonetics, phonology, orthography and other factors.

Furthermore, the role of the bilinguals as the borrowers is highlighted since the TA bilinguals exhibited the highest degrees of match in all three tasks of the SB experiment, i.e., audio, audio-written and written tasks. The TQ Quranic speakers of Turkish also manifested high degrees of match which were higher than the monolingual T group yet lower than the bilingual TA group. This result was shown in the logistic regression run on the Simulated Borrowing in the Audio-written compared to the Audio condition where the variable **stimulus presentation** was found significant at the Audio-written condition. This is reflected in the fact that the degrees of match in the audio-written level were higher in audio-written condition compared to that of the Audio only condition.

In order to discuss the results of the SB experiment in the audio-written condition, I compare the perceptual maps of each group to their predicted categories in the corpus data. These are provided in tables 7-5 to 7-7 which collectively mirror the finding that the number of matched categories in the SB-AW is higher compared to the SB-A.

Arabic vowel category	**SB-AW perceptual maps of T group**	**Predicted maps based on corpus loanwords**	**As predicted or not**	**Average goodness of fit for each listener group in SB-AW**
[i]	/ɯ/$_T$	/i/$_T$	X	72.22%
[i]$_{d^ʕ}$	/ɯ/$_T$	/ɯ/$_T$	√	88.88%
[i]$_q$	/ɯ/$_T$	/ɯ/$_T$	√	66.66%
[i]$_ʕ$	/e/$_T$	/i/$_T$	X	50%
[i:]	/i/$_T$	/i/$_T$	√	88.88%
[i:]$_{d^ʕ}$	/i/$_T$	/i/$_T$	√	72.22%
[i:]$_q$	/i/$_T$	/i/$_T$	√	44.44%
[i:]$_ʕ$	/i/$_T$	/i/$_T$	√	77.77%
[a]	/e/$_T$	/e/$_T$	√	77.77%
[a]$_{d^ʕ}$	/a/$_T$	/a/$_T$	√	88.88%
[a]$_q$	/a/$_T$	/a/$_T$	√	100%
[a]$_ʕ$	/a/$_T$	/a/$_T$	√	94.44%
[a:]	/a/$_T$	/a/$_T$	√	83.33%
[a:]$_{d^ʕ}$	/a/$_T$	/a/$_T$	√	100%
[a:]$_q$	/a/$_T$	/a/$_T$	√	100%
[a:]$_ʕ$	/a/$_T$	/a/$_T$	√	66.66%
[u]	/u/$_T$	/y/$_T$	X	72.22%
[u]$_{d^ʕ}$	/u/$_T$	/u/$_T$	√	100%
[u]$_q$	/u/$_T$	/u/$_T$	√	77.77%
[u]$_ʕ$	/o/$_T$	/u/$_T$	X	50%

[u:]	/u/$_T$	/u/$_T$	√	83.33%
[u:]$_{\text{dˤ}}$	/u/$_T$	/u/$_T$	√	100%
[u:]$_q$	/u/$_T$	/u/$_T$	√	94.44%
[u:]$_{\text{ʕ}}$	/u/$_T$	/u/$_T$	√	94.44%

Table 7-5: SB-AW vowel patterns and their predicted categories for the T group based on corpus patterns in Turkish

The monolingual group mismatched **four categories** out of twenty four in the SB-AW, all of which were mapped gradiently (less than 80%). These are [i]>/ɯ/$_T$ (predicted /i/$_T$), [i]$_{\text{ʕ}}$>/e/$_T$ (predicted as /i/$_T$), [u]>/y/$_T$ in *sull* 'tuberculosis' and [u]$_{\text{ʕ}}$>/o/$_T$. Three of these were also mapped in the SB-A condition, namely [i]$_{\text{ʕ}}$>/e/$_T$, [u]$_{\text{ʕ}}$>/o/$_T$ and [u]>/y/$_T$, the last being also mapped in the PAT experiment. However, the new mismatched map was that of [i]>/ɯ/$_T$.

Arabic vowel category	**SB-AW perceptual maps of TA group**	**Predicted maps based on corpus loanwords**	**As predicted or not**	**Average goodness of fit for each listener group in SB-AW**
[i]	/ɯ/$_T$	/i/$_T$	X	54.54%
[i]$_{\text{dˤ}}$	/ɯ/$_T$	/ɯ/$_T$	√	100%
[i]$_q$	/ɯ/$_T$	/ɯ/$_T$	√	72.72%
[i]$_{\text{ʕ}}$	/i/$_T$	/i/$_T$	√	72.72%
[i:]	/i/$_T$	/i/$_T$	√	100%
[i:]$_{\text{dˤ}}$	/i/$_T$	/i/$_T$	√	72.72%
[i:]$_q$	/i/$_T$	/i/$_T$	√	72.72%
[i:]$_{\text{ʕ}}$	/i/$_T$	/i/$_T$	√	81.81%
[a]	/e/$_T$	/e/$_T$	√	90.90%
[a]$_{\text{dˤ}}$	/a/$_T$	/a/$_T$	√	90.90%
[a]$_q$	/a/$_T$	/a/$_T$	√	63.63%
[a]$_{\text{ʕ}}$	/a/$_T$ or /e/$_T$	/a/$_T$	√ and X	50%
[a:]	/e/$_T$	/a/$_T$	X	54.54%
[a:]$_{\text{dˤ}}$	/a/$_T$	/a/$_T$	√	100%
[a:]$_q$	/a/$_T$	/a/$_T$	√	100%
[a:]$_{\text{ʕ}}$	/a/$_T$	/a/$_T$	√	72.72%
[u]	/u/$_T$	/y/$_T$	X	72.72%
[u]$_{\text{dˤ}}$	/u/$_T$	/u/$_T$	√	100%
[u]$_q$	/u/$_T$	/u/$_T$	√	81.81%
[u]$_{\text{ʕ}}$	/u/$_T$	/u/$_T$	√	90.90%
[u:]	/u/$_T$	/u/$_T$	√	100%
[u:]$_{\text{dˤ}}$	/u/$_T$	/u/$_T$	√	90.90%
[u:]$_q$	/u/$_T$	/u/$_T$	√	100%
[u:]$_{\text{ʕ}}$	/u/$_T$	/u/$_T$	√	100%

Table 7-6: SB-AW vowel patterns and their predicted categories for the TA group based on corpus patterns in Turkish

The bilingual group reflected **four** mismatched categories out of twenty four compared to their predicted categories in the corpus data. All these were mismatched with a goodness of fit of less than 80%. These include [i]>/ɯ/$_T$ (predicted as /i/$_T$), [a:]>/e/$_T$ (predicted as /a/$_T$), [u]>/u/$_T$ (predicted as /y/$_T$) and [a]$_ʕ$>/e/$_T$ (predicted as /a/$_T$) by half of the TA group participants. As mentioned, both the TA and T groups mismatched [i] to /ɯ/$_T$ but not the TQ group. In the perceptual map [a:]>/e/$_T$, the TA group was the only one that mismatched [a:]>/e/ in the SB-AW condition in contrast to the T and TQ groups that matched it in the AW condition but mismatched it in the A condition. The mismatched map of [u]>/u/$_T$ was incurred by the three groups in both AW and A condition. In the last mismatched pattern, [a]$_ʕ$ was assimilated by half the TA participants as /e/$_T$ whereas the other half triggered the correct category /a/$_T$.

Arabic vowel category	**SB-AW perceptual maps of TQ group**	**Predicted maps based on corpus loanwords**	**As predicted or not**	**Average goodness of fit for each listener group in SB-AW**
[i]	/i/$_T$	/i/$_T$	√	63.63%
[i]$_{dˤ}$	/ɯ/$_T$	/ɯ/$_T$	√	95.45%
[i]$_q$	/ɯ/$_T$	/ɯ/$_T$	√	81.81%
[i]$_ʕ$	/i/$_T$	/i/$_T$	√	45.45%
[i:]	/i/$_T$	/i/$_T$	√	95.45%
[i:]$_{dˤ}$	/i/$_T$	/i/$_T$	√	59.09%
[i:]$_q$	/ɯ/$_T$	/i/$_T$	X	59.09%
[i:]$_ʕ$	/ɯ/$_T$ and /i/	/i/$_T$	X and √	45.45% and 45.45%
[a]	/e/$_T$	/e/$_T$	√	100%
[a]$_{dˤ}$	/a/$_T$	/a/$_T$	√	100%
[a]$_q$	/a/$_T$	/a/$_T$	√	99.90%
[a]$_ʕ$	/a/$_T$	/a/$_T$	√	86.36%
[a:]	/a/$_T$	/a/$_T$	√	59.09%
[a:]$_{dˤ}$	/a/$_T$	/a/$_T$	√	100%
[a:]$_q$	/a/$_T$	/a/$_T$	√	100%
[a:]$_ʕ$	/a/$_T$	/a/$_T$	√	90.90%
[u]	/u/$_T$	/y/$_T$	X	72.72%
[u]$_{dˤ}$	/u/$_T$	/u/$_T$	√	95.45%
[u]$_q$	/u/$_T$	/u/$_T$	√	90.90%
[u]$_ʕ$	/u/$_T$	/u/$_T$	√	86.36%
[u:]	/u/$_T$	/u/$_T$	√	90.90%

$[u:]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	95.45%
$[u:]_q$	$/u/_T$	$/u/_T$	√	100%
$[u:]_ʕ$	$/u/_T$	$/u/_T$	√	90.90%

Table 7-7: SB-AW vowel patterns and their predicted categories for the TQ group based on corpus patterns in Turkish

The TQ group mapped **three** categories not as predicted. These are $[i:]_q$>$/ɯ/_T$ (predicted as $/i/_T$), $[i:]_ʕ$>$/i/_T$ or $/ɯ/_T$ (predicted as $/i/_T$) and [u]>$/u/_T$ (predicted as $/y/_T$). The TQ group mismatched $[i:]_q$ as $/ɯ/_T$ in both the SB-A and SB-AW, and [u] as $/u/_T$ in the PAT, SB-A and SB-AW which suggests that the these perceptual maps are results of either perception and/or phonology but not orthography. Unlike the two other groups, the TQ participants were divided in their responses when assimilating $[i:]_ʕ$ since half responded with $/i/_T$ (i:4, ii:1, i-i:1, ai:2, e-i:1, i-e:1; 10/22) and the other half with $/ɯ/_T$ (ı:10/22) with a goodness of fit of 45.45% for each of the two responses.

In what follows, phonetic proximity and distinctive features are resorted to to determine whether the source of the mismatched perceptual maps by the three groups is phonetic, phonological or orthographic or related to other factors.

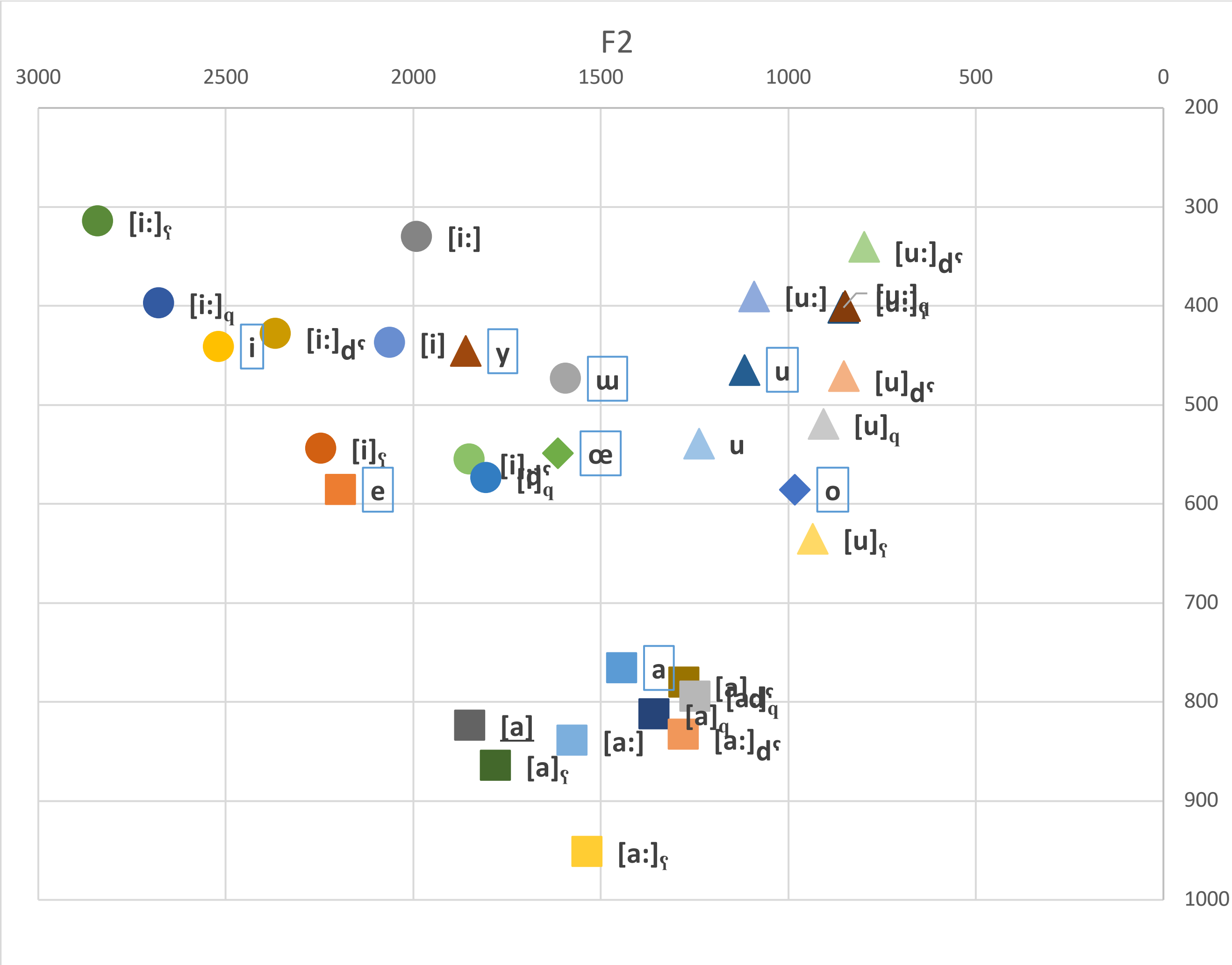


Figure 7-4:Mean frequency values of 2 Turkish speakers and 1 Arabic speaker's vowel formants plot (real words) in the SB-audio-written task

#Vowels in boxes are Turkish vowels and the rest are Arabic ones; underlined /a/, /i/and /u/ are Arabic vowels

First, the four mismatched maps by the T group [(i]>/ɯ/$_{T}$ (predicted /i/$_{T}$), [i]$_{\text{ʕ}}$>/e/$_{T}$ (predicted as /i/$_{T}$), [u]>/u/$_{T}$ (predicted as /y/$_{T}$) and [u]$_{\text{ʕ}}$>/o/$_{T}$ (predicted as /u/$_{T}$)) are examined. As shown in figure 7-4 and 24) below, Arabic [i] is closer to [i]$_{T}$ than [ɯ]$_{T}$ in both F1 [height] and F2 [backness]. Moreover, [i]$_{A}$ and [i]$_{T}$ are phonologically similar in the two features of height and frontness whereas [i]$_{A}$ and [ɯ]$_{T}$ are similar in only the feature high as shown in 25). Hence, the perceptual map [i]>/ɯ/$_{T}$ neither is phonetically nor phonologically supported. Furthermore, the monolingual group does not have access to Arabic orthography, so the audio-written condition does not play a role in the mapping. The only source remaining is native Turkish. That is, the Arabic word *zirr* 'button' sounds similar to the Arabic loanword in Turkish *sır* 'secret' (similar by analogy).

24) Comparison of the formants of Arabic [i] and Turkish /i/$_T$ and /ɯ/$_T$

Arabic [i] F2 2063	Turkish/i/ F2 2519.59	Turkish /ɯ/ F2 1594
F1 437	F1 441	F1 473.5

25) Comparison of feature specifications of [i]$_A$ and Turkish /i/$_T$ and /ɯ/$_T$

A. [i]$_A$>/i/$_T$

Arabic [i]	**>**	**Output Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. [i]$_A$>/ɯ/$_T$

Arabic [i]	**>**	**Output Turkish /ɯ/**
[+high]	√	[+high]
[+front]	X	[-front]

The second perceptual map [i]$_ʕ$>/e/ was yielded both in the SB-A and SB-AW by the T group. This indicates that orthography plays no role in the assimilation. Phonetically, [i]$_ʕ$ is closer to [e]$_T$ in both F1 and F2 (figure 7-4 and 26)). Phonologically, [i]$_ʕ$ is more similar to [i]$_T$ than [e]$_T$ in the two features of height and frontness (27). Therefore, [i]$_ʕ$>/e/ is phonetically grounded but not phonologically.

26) Comparison of the formants of Arabic [i]$_ʕ$ and Turkish /e/$_T$ and /i/$_T$

Arabic **[i]**$_ʕ$ F2 2247	Turkish /e/ F2 2194	Turkish /i/ F2 2519.5
F1 544	F1 586	F1 441

27) Comparison of feature specifications of Arabic [i]$_ʕ$ and Turkish /e/$_T$ and /i/$_T$

A. [i]$_ʕ$ >/e/$_T$

Arabic [i]$_ʕ$	**>**	**Turkish /e/**
[+high]	X	[+high]
[+front]	√	[+front]

B. [i]$_ʕ$ >/i/$_T$

Arabic [i]$_ʕ$	**>**	**Turkish /i/**
[+high]	√	[+high]

[+front]	√	[+front]

The third perceptual map $[u]_T>/u/_T$ was rendered in the SB-A, SB-AW and the PAT by all three groups. This by itself suggests no role of orthography. The Arabic $[u]_A$ and Turkish $[u]_T$ are phonetically more similar in F1 and F2 than $[u]_A$ and $[y]_T$ are (see 28) and figure (7-4). Furthermore, Arabic $[u]_A$ and Turkish $[u]_T$ are phonologically more similar in the distinctive features of height and backness than $[u]_A$ and $[y]_T$ are (29). This means that $[u]_T>/u/_T$ is both phonetically and phonologically sustained.

28) Comparison of the formants of Arabic $[u]_A$ and Turkish $/y/_T$ and $/u/_T$

Arabic [u]	F2 1238	Turkish/y/	F2 1860	Turkish /u/	F2 1117
	F1 539		F1 446		F1 464.5

29) Comparison of feature specifications of $[u]_A$ and Turkish $/u/_T$ and $/y/_T$

A. $[u]_A>/u/_T$

Arabic [u]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. $[u]_A>/y/_T$

Arabic [u]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

The last mismatched perceptual map $[u]_{ʕ}>/o/_T$ was also made by the T group only in both SB-A and SB-AW. As shown in 30) and figure 7-4, $[u]_{ʕ}$ is acoustically closer to $/o/_T$ than $/u/_T$ in F1 and F2. In 31), however, $[u]_{ʕ}$ is shown as being more similar phonologically to $/u/_T$ (B) in height and backness than $/o/_T$ (A) is. Thus, $[u]_{ʕ}>/o/_T$ is phonetically supported but not phonologically. In short, the perceptual maps of the T group in the SB-AW reflect the effect of phonetics ($[u]_{ʕ}>/o/_T$ and $[i]_{ʕ}>/e/$), both phonetics and phonology ($[u]_T>/u/_T$) and Turkish morphology ($[i]>/ɯ/_T$).

30) Comparison of the formants of Arabic [u]$_{ʕ}$ and Turkish /o/$_T$ and /u/$_T$

Arabic [u]$_{ʕ}$ F2 934	Turkish /o/ F2 983	Turkish /u/ F2 1117
F1 635	F1 586	F1 464.5

31) Comparison of feature specifications of Arabic [u]$_{ʕ}$ and Turkish /o/$_T$ and /u/$_T$

A. [u]$_{ʕ}$ >/o/$_T$

Arabic [u]$_{ʕ}$	**>**	**Turkish /o/**
[+high]	X	[-high]
[-front]	√	[-front]

B. [u]$_{ʕ}$ >/u/$_T$

Arabic [u]$_{ʕ}$	**>**	**Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

Similar to the T group, the TA participants categorized [i]$_A$ as /ɯ/$_T$ instead of the predicted /i/$_T$. Here too, we can conclude that [i]$_A$>/ɯ/$_T$ neither is phonetically nor phonologically corroborated since the measurements in 32) (phonetics) and the distinctive featural analysis in 33) sustain mapping onto /i/$_T$ rather than /ɯ/$_T$. Hence, the only remaining explanation would be native Turkish, that the stimulus word *zır* ‘button ’sounded similar to the Arabic loanword *sır* ‘secret’ in Turkish, i.e., giving weight to Turkish morphology. No role of orthography is evident in this context since both the monolingual and bilingual listeners produced the ‘wrong’ vowel category.

32) Comparison of the formants of Arabic [i] and Turkish /i/$_T$ and /ɯ/$_T$

Arabic [i] F2 2063	Turkish/i/ F2 2519.59	Turkish /ɯ/ F2 1594
F1 437	F1 441	F1 473.5

33) Comparison of feature specifications of [i]$_A$ and Turkish /i/$_T$ and /ɯ/$_T$

A. [i]$_A$>/i/$_T$

Arabic [i]	**>**	**Output Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i]_A > /ɯ/_T$

Arabic [i]	>	Output Turkish /ɯ/
[+high]	√	[+high]
[+front]	X	[-front]

The second mismatched perceptual map is that of $[a:]_A$ as $/e/_T$ instead of $/a/_T$. Noteworthy to mention is that the bilingual group was the only one that mismatched it in the SB-AW. According to the acoustic measurements in 34), $[a:]_A$ is closer to $/a/_T$ than $/e/_T$ is in Fi [height] and F2 [backness]. This suggests that the T and TQ groups were influenced by the phonetic proximity of the vowel categories but not the TA group. Phonologically, comparing $[a:]_A$ to $/e/_T$ and $/a/_T$ in 35), it becomes clear that $[a:]_A$ and $/e/_T$ are more similar in height and frontness than $[a:]_A$ and $/a/_T$. This shows that the bilingual group's categorization was driven by the phonology of Arabic. Regarding the effect of orthography, no role is detected for the perceptual map $[a:]_A > /e/_T$ since $[a:]_A$ represented by **alif** would be predicted to be categorized as $/a/_T$ especially that $[a:]_A$ and $/a/_T$ have the same vowel quality.

34) Comparison of the formants of Arabic [a:] and Turkish $/e/_T$ and $/a/_T$

Arabic [a:]	F2 1576	Turkish/e/	F2 2194	Turkish /a/	F2 1445
	F1 839		F1 586		F1 766

35) Comparison of feature specifications of $[a:]_A$ and Turkish $/e/_T$ and $/a/_T$

A. $[a:]_A > /e/_T$

Arabic [a:]	>	Output Turkish /e/
[-high]	√	[-high]
[+front]	√	[+front]

B. $[a:]_A > /a/_T$

Arabic [a:]	>	Output Turkish /a/
[-high]	√	[-high]
[+front]	X	[-front]

The third mismatched perceptual map by the TA group is $[u]_A > /u/_T$ instead of $[u]_A > /y/_T$, which all three groups brought forth in SB-A, SB-AW and the PAT. This indicates no role of orthography in determining the vowel quality of the resulting Turkish vowel. As shown for the T group, $[u]_A$>$/u/_T$ is both phonetically and phonologically supported (36) and 37). To sum up thus far, the TA group's perceptual maps mirror the effects of phonology ($[a:]_A$>$/e/_T$), both phonetics and phonology c and Turkish morphology ([i]>$/ɯ/_T$).

36) Comparison of the formants of Arabic $[u]_A$ and Turkish $/y/_T$ and $/u/_T$

Arabic [u]	F2 1238	Turkish/y/	F2 1860	Turkish /u/	F2 1117
	F1 539		F1 446		F1 464.5

37) Comparison of feature specifications of $[u]_A$ and Turkish $/u/_T$ and $/y/_T$

A. $[u]_A$>$/u/_T$

Arabic [u]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. $[u]_A$>$/y/_T$

Arabic [u]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

The last mismatched perceptual map by half the TA group participants was $[a]_ʕ$>$/e/_T$ instead of the predicted $/a/_T$. As shown in 38) and 39) $[a]_ʕ$ and $/a/_T$ are more similar phonetically and phonologically respectively than $[a]_ʕ$ and $/e/_T$. That is, $[a]_ʕ$ and $/a/_T$ are phonetically closer to each other in both F1 and F2 than $[a]_ʕ$ and $/e/_T$ are. Morover, $[a]_ʕ$ and $/a/_T$ share the two features of height and backness whereas $[a]_ʕ$ and $/e/_T$ share only the feature [-high].

38) Comparison of the formants of Arabic $[a]_ʕ$ and Turkish $/e/_T$ and $/a/_T$

Arabic $[a]_ʕ$	F2 1789	Turkish/e/	F2 2194	Turkish /a/	F2 1445
	F1 864		F1 586		F1 766

39) Comparison of feature specifications of $[a]_ʕ$ and Turkish $/e/_T$ and $/a/_T$

A. $[a]_ʕ$>$/e/_T$

Arabic [a]$_ʕ$	>	Output Turkish /e/
[-high]	√	[-high]
[-front]	X	[+front]

B. [a]$_ʕ$>/a/$_T$

Arabic [a]$_ʕ$	>	Output Turkish /a/
[-high]	√	[-high]
[-front]	√	[-front]

The TQ respondents mismatched three perceptual maps, namely [i:]$_q$>/ɯ/$_T$, [i:]$_ʕ$>/i/$_T$ or /ɯ/$_T$ and [u]$_A$>/u/$_T$. In the [i:]$_q$>/ɯ/$_T$, [i:]$_q$ and /i/$_T$ are closer in F1 and F2 in 40) and according to 41), [i:]$_q$ and /i/$_T$ are more similar than [i:]$_q$ and /ɯ/$_T$ in the two features of height and frontness. This means that the perceptual map [i:]$_q$>/ɯ/$_T$ neither is phonetically nor phonologically grounded. Orthography too cannot be the source of this perceptual map, since [i:]$_q$ represented with a yaa **(** would be expected to be mapped onto a vowel of the same quality, i.e. /i/$_T$. One explanation for this map is that of Turkish phonology, that the participants reflected the backing effect and lowering effect of /q/ as similar to the back-front distinction of their native vowel harmony.

40) Comparison of the formants of Arabic [i:]$_q$ and Turkish /ɯ/$_T$ and /i/$_T$

Arabic **[i:]**$_q$	Turkish /ɯ/	Turkish /i/
F2 2680	F2 1594	F2 2519.5
F1 397	F1 473.5	F1 441

41) Comparison of feature specifications of Arabic [i:]$_ʕ$ and Turkish /ɯ/$_T$ and /i/$_T$

A. [i:]$_q$ >/ɯ/$_T$

Arabic [i:]$_q$	>	Turkish /ɯ/
[+high]	√	[+high]
[+front]	X	[-front]

B. [i:]$_q$ >/i/$_T$

Arabic [i:]$_q$	>	Turkish /i/
[+high]	√	[+high]
[+front]	√	[+front]

The second perceptual map was the one where the responses of the TQ listeners were evenly split since half the participants mapped [i:]$_{ʕ}$ onto /i/$_{T}$ and the other half onto /ɯ/$_{T}$. Phonetically, [i:]$_{ʕ}$ is closer in F1 and F2 to /i/$_{T}$ than /ɯ/$_{T}$ as shown both in figure 7-4 and 42). Moreover, [i:]$_{ʕ}$ is closer to /i/$_{T}$ rather than /ɯ/$_{T}$ in both height and frontness (43). Thus, the perceptual map [i:]$_{ʕ}$>/i/$_{T}$ is both phonetically and phonologically grounded while [i:]$_{ʕ}$>/ɯ/$_{T}$ is not. One explanation for this latter map might be that the TQ participants interpreted the residual effect from the /ʕ/ in the source word as similar to a backing effect employing their Turkish vowel harmony of front-back difference.

42) Comparison of the formants of Arabic [i:]$_{ʕ}$ and Turkish /ɯ/$_{T}$ and /i/$_{T}$

Arabic **[i:]**$_{ʕ}$ F2 2841	Turkish /ɯ/ F2 1594	Turkish /i/ F2 2519.5
F1 314	F1 473.5	F1 441

43) Comparison of feature specifications of Arabic [i:]$_{ʕ}$ and Turkish /ɯ/$_{T}$ and /i/$_{T}$

A. [i:]$_{ʕ}$ >/ɯ/$_{T}$

Arabic [i:]$_{ʕ}$	**>**	**Turkish /ɯ/**
[+high]	√	[+high]
[+front]	X	[-front]

B. [i:]$_{ʕ}$ >/i/$_{T}$

Arabic [i:]$_{ʕ}$	**>**	**Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

The last mismatched perceptual map of [u]$_{A}$>/u/$_{T}$ was found across the three group in the SB-A, SB-AW ad the PAT, meaning that orthography plays no role in this map. As was explained for the T and TA groups, the source of this map is both phonetic, due to phonetic proximity between [u]$_{A}$ and /u/$_{T}$ (figure 7-4 and 44)) and phonological similarity between them (45)). In summary, the perceptual maps of the TQ group can be said to be triggered by phonetics and Arabic phonology ([u]$_{A}$>/u/$_{T}$ and [i:]$_{ʕ}$>/i/$_{T}$), and Turkish phonology ([i:]$_{q}$>/ɯ/$_{T}$ and [i:]$_{ʕ}$>/ɯ/$_{T}$). All in all, the perceptual maps of the three groups of T, TA and TQ in the SB-AW sustain a hybrid model of phonetics, Arabic phonology, Turkish phonology, Turkish morphology and orthography.

44) Comparison of the formants of Arabic [u]$_A$ and Turkish /y/$_T$ and /u/$_T$

Arabic [u] F2 1238 F1 539	Turkish/y/ F2 1860 F1 446	Turkish /u/ F2 1117 F1 464.5

45) Comparison of feature specifications of [u]$_A$ and Turkish /u/$_T$ and /y/$_T$

A. [u]$_A$>/u/$_T$

Arabic [u]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. [u]$_A$>/y/$_T$

Arabic [u]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

In the last part of this section, a comparison of the TA's and TQ's perceptual maps in tables 7-8 and 7-9 to the categories of the corpus data is given.

Arabic vowel category	**SB-W perceptual maps of TA group**	**Predicted maps based on corpus loanwords**	**As predicted or not**	**Average goodness of fit for each listener group in SB-W**
[i]	/i/$_T$	/i/$_T$	√	100%
[i]$_{d^ʕ}$	/i/$_T$	/ɯ/$_T$	X	81.81%
[i]$_q$	/i/$_T$	/ɯ/$_T$	X	90.90%
[i]$_ʕ$	/i/$_T$	/i/$_T$	√	81.81%
[i:]	/i/$_T$	/i/$_T$	√	100%
[i:]$_{d^ʕ}$	/i/$_T$	/i/$_T$	√	81.81%
[i:]$_q$	/i/$_T$	/i/$_T$	√	90.90%
[i:]$_ʕ$	/i/$_T$	/i/$_T$	√	72.72%
[a]	/e/$_T$	/e/$_T$	√	63.63%
[a]$_{d^ʕ}$	/a/$_T$	/a/$_T$	√	45.45%
[a]$_q$	/a/$_T$	/a/$_T$	√	72.72%
[a]$_ʕ$	/a/$_T$	/a/$_T$	√	63.63%
[a:]	/e/$_T$	/a/$_T$	X	63.63%
[a:]$_{d^ʕ}$	/a/$_T$	/a/$_T$	√	63.63%
[a:]$_q$	/a/$_T$	/a/$_T$	√	72.72%

[a:]$_{\text{ʕ}}$	/a/$_T$	/a/$_T$	√	72.72%
[u]	/u/$_T$	/y/$_T$	X	81.81%
[u]$_{d^{\text{ʕ}}}$	/u/$_T$	/u/$_T$	√	72.72%
[u]$_q$	/u/$_T$	/u/$_T$	√	90.90%
[u]$_{\text{ʕ}}$	/u/$_T$	/u/$_T$	√	90.90%
[u:]	/u/$_T$	/u/$_T$	√	63.63%
[u:]$_{d^{\text{ʕ}}}$	/u/$_T$	/u/$_T$	√	90.90%
[u:]$_q$	/u/$_T$	/u/$_T$	√	90.90%
[u:]$_{\text{ʕ}}$	/u/$_T$	/u/$_T$	√	81.81%

Table 7-8: SB-W vowel patterns and their predicted categories for the TA group based on corpus patterns in Turkish

The TA and TQ groups mismatched similar perceptual categories in the written condition of the Simulated Borrowing albeit with varying degrees of match as shown in tables 7-8 and 7-9. These are [i]$_{d^{\text{ʕ}}}$>/i/$_T$ (predicted as /ɯ/$_T$), [i]$_q$>/i/$_T$ (predicted as /ɯ/$_T$), [a:]>/e/$_T$ (predicted as /a/$_T$) and [u]>/u/$_T$ (predicted as /y/$_T$). Having said this, the TQ group was divided even in assimilating [u:] since half the participants mapped it onto /u/$_T$ (45.45%) while others onto /y/$_T$ (45.45%) as shown in table 7-9. Of the mismatched maps, the two groups assimilated [i]$_{d^{\text{ʕ}}}$>/i/$_T$ and [i]$_q$>/i/$_T$ (predicted as /ɯ/$_T$) categorically. However, they both perceived [a:]>/e/$_T$ gradiently (63.63% by both groups) and the TQ group assimilated [u]>/u/$_T$ gradiently (59.09%). In chapter six, we saw how the mismatched maps were expressed in the logistic regression performed. The significant variables included **context** and **vowel.quality** and their interaction **context**:**vowel.quality** which reflects the mismatched patterns.

Arabic vowel category	**SB-W perceptual maps of TQ group**	**Predicted maps based on corpus loanwords**	**As predicted or not**	**Average goodness of fit for each listener group in SB-W**
[i]	/i/$_T$	/i/$_T$	√	90.90%
[i]$_{d^{\text{ʕ}}}$	/i/$_T$	/ɯ/$_T$	X	90.90%
[i]$_q$	/i/$_T$	/ɯ/$_T$	X	95.45%
[i]$_{\text{ʕ}}$	/i/$_T$	/i/$_T$	√	72.72%
[i:]	/i/$_T$	/i/$_T$	√	100%
[i:]$_{d^{\text{ʕ}}}$	/i/$_T$	/i/$_T$	√	86.36%
[i:]$_q$	/i/$_T$	/i/$_T$	√	63.63%
[i:]$_{\text{ʕ}}$	/i/$_T$	/i/$_T$	√	72.72%
[a]	/e/$_T$	/e/$_T$	√	81.81%
[a]$_{d^{\text{ʕ}}}$	/a/$_T$	/a/$_T$	√	68.18%
[a]$_q$	/a/$_T$	/a/$_T$	√	90.90%

$[a]_{ʕ}$	$/a/_T$	$/a/_T$	√	77.27%
[a:]	$/e/_T$	$/a/_T$	X	63.63%
$[a:]_{d^ʕ}$	$/a/_T$	$/a/_T$	√	77.27%
$[a:]_q$	$/a/_T$	$/a/_T$	√	95.45%
$[a:]_{ʕ}$	$/a/_T$	$/a/_T$	√	90.90%
[u]	$/u/_T$	$/y/_T$	X	59.09%
$[u]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	63.63%
$[u]_q$	$/u/_T$	$/u/_T$	√	77.27%
$[u]_{ʕ}$	$/u/_T$	$/u/_T$	√	77.27%
[u:]	$/u/_T$ and $/y/_T$	$/u/_T$	√ and X	45.45% and 45.45%
$[u:]_{d^ʕ}$	$/u/_T$	$/u/_T$	√	72.72%
$[u:]_q$	$/u/_T$	$/u/_T$	√	68.18%
$[u:]_{ʕ}$	$/u/_T$	$/u/_T$	√	77.27%

Table 7-9: vowel patterns and their predicted categories for the TQ group based on corpus patterns in Turkish

Regarding the source of these perceptual maps, orthography is the first one. However, we can safely rule out the role of phonetic/acoustic cues since the stimuli were only written. Phonology still can play a role in the mapping as the participants in both groups possess knowledge of Arabic. Regarding the first mismatched perceptual map $[i]_{d^ʕ}$>$/i/_T$, as shown in 46) $[i]_{d^ʕ}$ agrees with $/i/_T$ in height and frontness but agrees with $/ɯ/_T$ only in height. This means that $[i]_{d^ʕ}$>$/i/_T$ is phonologically and orthographically supported.

46) Comparison of feature specifications of $[i]_{d^ʕ}$ and Turkish $/i/_T$ and $/ɯ/_T$

A. $[i]_{d^ʕ}$ >$/i/_T$

Arabic $[i]_{d^ʕ}$	**>**	**Output Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i]_{d^ʕ}$ >$/ɯ/_T$

Arabic $[i]_{d^ʕ}$	**>**	**Output Turkish /ɯ/**
[+high]	√	[+high]
[+front]	X	[-front]

Similar to $[i]_{d^ˤ}$>$/i/_T$, the perceptual map $[i]_q$>$/i/_T$ is also phonologically grounded since $[i]_q$ and $/i/_T$ are more similar in height and frontness than $[i]_q$ and $/ɯ/_T$ which are similar only in height as shown in 47). Furthermore, $[i]_q$>$/i/_T$ is also orthographically influenced.

47) Comparison of feature specifications of $[i]_q$ and Turkish $/i/_T$ and $/ɯ/_T$

A. $[i]_q$ >$/i/_T$

Arabic $[i]_q$	**>**	**Output Turkish /i/**
[+high]	√	[+high]
[+front]	√	[+front]

B. $[i]_q$ >$/ɯ/_T$

Arabic $[i]_q$	**>**	**Output Turkish /ɯ/**
[+high]	√	[+high]
[+front]	X	[-front]

In 48), 49) and 50), the three perceptual maps [a:]>$/e/_T$, [u]>$/u/_T$ and [u:]>$/u/_T$ are orthographically, phonologically supported or both. Phonologically, these three maps agree in height and frontness/backness whereas [a:]>$/a/_T$, [u]>$/y/_T$ and [u:]>$/y/_T$ agree only in the feature height. In addition, Arabic does not have $/y/_T$ in its phonemic inventory. Orthographically, [a:], [u] and [u:] are not surrounded by gutturals/emphatics which might trigger backness of vowels (under the rules of Turkish vowel harmony); hence are interpreted as front vowels.

48) Comparison of feature specifications of $[a:]_A$ and Turkish $/e/_T$ and $/a/_T$

A. $[a:]_A$>$/e/_T$

Arabic [a:]	**>**	**Output Turkish /e/**
[-high]	√	[-high]
[+front]	√	[+front]

B. $[a:]_A$>$/a/_T$

Arabic [a:]	**>**	**Output Turkish /a/**
[-high]	√	[-high]
[+front]	X	[-front]

49) **Comparison of feature specifications of [u]$_A$ and Turkish /u/$_T$ and /y/$_T$**

A. [u]$_A$>/u/$_T$

Arabic [u]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. [u]$_A$>/y/$_T$

Arabic [u]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

50) Comparison of feature specifications of [u:]$_A$ and Turkish /u/$_T$ and /y/$_T$

A. [u:]$_A$>/u/$_T$

Arabic [u:]	**>**	**Output Turkish /u/**
[+high]	√	[+high]
[-front]	√	[-front]

B. [u:]$_A$>/y/$_T$

Arabic [u:]	**>**	**Output Turkish /y/**
[+high]	√	[+high]
[-front]	X	[+front]

Thus, in the written condition of the Simulated Borrowing experiment, it was demonstrated that knowledge of Arabic phonology and orthography both equally play a role in determining the vowel quality of the resulting Turkish vowel. The fact that both groups with knowledge of Arabic rendered similar perceptual maps but with the TA group yielding even more matched maps and higher degrees of match to the corpus patterns can be taken such that the original borrowers of the Arabic loanwords in Turkish too must have used the Arabic words in everyday life both in the spoken and written forms. In addition, this also implies that the Arabic loanwords infiltrated into Turkish both as spoken and written words by proficient bilinguals.

7.6.1 Summary

To recapitulate, most of the mappings (70%) are due to perception in the PAT. However, there is another 30% which has originated from somewhere (perception and phonology (and probably orthography and/or other effects). If we simulate borrowing with real words (non-forced choice), the percentage which can be accounted for by pure perception drops to 50%. Hence, it appears that another 50% must have come from somewhere. In principle, both the TA and/or TQ could have done significantly better in the SB-audio, but they do not really as we do not see any main effect of listener group in the logistic regression analysis. The TA group matches *slightly* better though, descriptively, so there is some potential effect of pure knowledge of Arabic phonology, but not much.

If we add in orthography, then the match improves, for those that are fluent in Arabic (TA). This suggests a) that the corpus mappings are the result of borrowing by people who knew both languages; and b) that we need a hybrid model of phonetics, phonology of both source and native language, orthography and possibly other factors such as morphology to account for even just the 62% or so. We can model the corpus data based on the SB-AW and W data. This means a big role for perception, which gives us the 'residual guttural' effect, but also a role for knowledge of the source language and mostly a 'normative' knowledge, (i.e. involving citation/written forms). A change in the quality of vowels in the loanwords is certainly plausible. As we know from chapter two, the 'purification machine' was at work since 1932, substituting Arabic and Persian words with Turkish ones and even coining new words altogether at times which sounded Turkish-like.

7.7 Role of Bilingualism

As was evident from the results of the Simulated Borrowing tasks, the TA bilingual group was found to resemble the active borrowers of the loanwords, since in the experiment their mappings most closely matched those of the Ottomans as represented by the corpus patterns. This, in turn, suggests that the speakers of Osmanlica themselves were also bilingual and that they must have introduced many spoken and/or written Arabic words into Turkish/ Osmanlica.

Three points can be raised here regarding the context of bilingualism in the Arabic loanwords in Turkish during the time of the Ottomans, the definition of bilinguals and how that relates to the topic of this thesis. First, according to Versteegh (2001, p. 502), he notes that "in the case of Turkish and Urdu there was no direct context of bilingualism" since the borrowing happened through Persian as an intermediate channel. He points out that despite the abundant Arabic and Persian loanwords in Turkish, the lack of diachronic analysis to date means we cannot deduce when these words entered the language. Versteegh later states that there must have been "bilingual intellectuals" (p. 502) who were responsible for the later words. We know that some Turcologists, spearheaded by Tietze, actually did work on the etymology of Arabic and Persian loanwords in Turkish. Tietze's work, for instance, is documented in two large volumes which he, unfortunately, did not finish due to his passing away. In any case, as was mentioned in chapter two of this thesis, Tietze (1992) sketches the features of the two historical stages of Arabic loanword adaptation in Turkish where he maintains that during the first stage Arabic words were borrowed via Persian. However, they were borrowed directly during the second stage when Classical Arabic was used to correct the older words with Persian pronunciation (p. 350). During the latter stage, more words were borrowed directly from Arabic through trade and religious schooling 'madrasas' which resembled Classical Arabic pronunciation more than the ones in the first stage, which matches Versteegh's description of "bilingual intellectuals."

This leads us to the definition of bilinguals, and discussion of how proficient in Arabic the Osmanlica speakers were. According to Paradis and LaCharité (2001a, p. 4), bilinguals are speakers "who have access to the codes of both the source and borrowing languages." Moreover, Paradis and LaCharité (ibid) do not use the term bilinguals to refer to *balanced bilinguals*, i.e., bilinguals who are equally proficient in two languages. In this sense, the TA bilinguals in the SB experiment are probably all early/balanced bilinguals since they either learned Arabic at an early age and/or one or both of their parents speaks Arabic.

Therefore, since the bilinguals in the SB experiment were the closest to the mappings of the Ottomans, this can by extension be interpreted such that the Osmanlica speakers were likely to have been proficient bilinguals themselves. However, we see a real effect in the SB experiment only when they have the written form; the TA were not significantly better than TQ or T in the SB-audio task. For example, there was only an interaction with age in that there was a reduced effect of age for the TA listeners than for the T or TQ groups. Thus, although it is the TA listeners who match more, it is their knowledge of written Arabic which they make use of. Moreover, this also sheds light on the SB stimuli material that the words used in the SB experiment are similar to those that were borrowed during the second stage of adaptation by Osmanlica speakers. Thus, the fact that the bilinguals were the group that did better in all three conditions implies that the Ottoman intellectuals possessed proficiency levels in spoken and written Arabic similar to those of day-to-day users. Moreover, it also mirrors very closely Versteegh's description of 'bilingual intellectuals' as the main agents of borrowing which this thesis confirmed with empirical data.

7.8 Conclusions, limitations, contributions and recommendations for future research

A number of conclusions can be drawn from this thesis. First, perception is important and is capable of explaining the most part of loanwords in the form of non-words. However, real non-borrowed words were found to challenge the Perceptual model and the role of listener group became more prominent, where knowledge of the source language was key. Bilingual speakers were found to be the active borrowers in the three tasks of SB; audio-only, audio-written and written conditions. Moreover, it was concluded that the input to the adaptation process was phonetic, phonological and orthographic as orthography assisted in matching the predicted categories. This, in turn, sustains the notion that a hybrid model of both perception, phonology and orthography can account for the Arabic loanwords corpus patterns identified in chapter three.

One of the biggest contributions of this thesis is that it has shown empirically in the three separate tasks of the Simulated Borrowing experiment that a hybrid model of phonetics and phonology is needed to account for the Arabic loanwords corpus facts presented in chapter three. It also showed that orthography plays a pivotal role in determining the mappings and that bilinguals were consistently and systematically the active borrowers in the SB-tasks which ultimately sustains the hybrid model. Another contribution is the original data, primarily in the form of the corpus of Arabic loanwords into Turkish, established from primary sources, in chapter three, but also in the form of the set of carefully selected non-word and real word stimuli created for use in the SB experiment.

Regarding the limitations, one of these is that the thesis, as it stands now, provides only a sketch of how the findings could be analyzed in a formal model. This limitation is due to the complexity of the data and the fact that four datasets were used to manipulate both the properties of the stimuli and the stimulus presentation conditions. Future work could draw on certain findings from the different analyses presented in this thesis, perhaps, and formalize them independently. Moreover, future research could tackle polysyllabic words and the syllable structure of the resulting words since all the stimuli words in the SB condition looked solely at monosyllabic words. In addition, consideration of emphasis spread in vowels, and its directionality, in Arabic loanwords in Turkish seems also to be a promising topic worthy of investigation. Nevertheless, the present work has laid a solid empirical foundation for further investigations of this type, which we hope will shed further light on the complex linguistic situation which resulted in the unique phonological properties of Arabic loanwords in the present day Turkish lexicon.

Appendices

Appendix 4-1: Language Background Questionnaire

This questionnaire aims to gather some information about your language background and your language proficiency level. I would like to find out what languages you are fluent in and how proficient you are in using them.
Date: ________

A. General questions:

1. Gender: ____

2. Age: ____

3. City and province of birth: ____

4. What is your highest level of education? (Please place a ✓)

High school ____ University ____ Master's ____ Doctorate _____ Other ______

5. Have you ever had any of the following? (Check all applicable).

1. Vision problem ☐
2. Hearing impairment ☐
3. Language disability ☐
4. Learning disability ☐
5. Other☐

6. In the previous question, if yes, please explain (including any corrections):
____

B. Language background:

7. What is your 1st language? _____________2nd ____________3rd ____________

8. What is your father's 1st language? ___________2nd __________3rd___________

9. What is your mother's 1st language? ____________2nd _________3rd__________

10. What languages do you use at home?1st_________2nd_________3rd___________

11. How would you rate your reading, understanding, speaking and writing skills in Turkish on a scale from '0' to '10' where '0' is poor and '10' is good?

Proficiency rate in Turkish							
Reading	☐	Understanding	☐	Speaking	☐	Writing	☐

12. If you speak Arabic as your 1st, 2nd or 3rd language or know Arabic through reciting the Holy Qur'an, please proceed to questions 13, 14, 15 and 16 below; otherwise, click on the 'Not applicable' option below.

I am a bilingual or know Arabic from reciting Qur'an ☐

Not applicable ☐

13. How long have you been using Arabic?

____

14. In what context(s) do you mostly use your knowledge of Arabic? (e.g. at home, reciting Qur'an, etc).

____

15. How would you rate your reading, understanding, speaking and writing skills in Arabic on a scale from '0' to '10' where '0' is poor and '10' is good from the scroll down menus?

Proficiency rate in Arabic							
Reading	☐	Understanding	☐	Speaking	☐	Writing	☐

16. Translate the word 'كتاب' to its equivalent in Turkish.

____

Thank you!

Dil Geçmişi Anketi

Bu anket, dil geçmişiniz ve dil yetkinlik düzeyiniz hakkında kimi bilgileri toplamayı amaçlamaktadır.Hangi dilleri akıcı bir biçimde konuştuğunuzu ve bu dilleri kullanmakta ne kadar yetkin olduğunuzu öğrenmek istiyorum.

Tarih: ________

A. Genel Sorular:

1. Cinsiyet: ____

2. Yaş: ____

3. Doğduğu şehir ve eyalet: ____

4. Mezun olduğunuz en yüksek eğitim derecesi hangisidir? (Lütfen bir '✓' koyun)

Lise ____ Üniversite ____ Yüksek Lisans ____ Doktora _____ Diğer ______

5. Aşağıdakilerden herhangi birini yaşadınız mı? (Size uyan tüm seçenekleri işaretleyin).

   1. bir görme sorununuz ☐
   2. işitme zayıflığınız ☐
   3. dil engeliniz ☐
   4. öğrenme güçlüğünüz ☐
   5. Diğer ☐

6. Bir önceki soruya evet dediyseniz, lütfen açıklayınız (tedavileri ile).
____

B. Dil geçmişi:

7. İlk (1'inci) diliniz hangisidir? ___________2'nci ___________3'üncü____________

8. Babanızın ilk (1'inci) dili hangisidir? ________ 2'nci_________3'üncü___________

9. Annenizin ilk (1'inci) dili hangisidir? ___________2'nci________3'üncü_________

10. Evde hangi dilleri kullanıyorsunuz? 1'inci_______2'nci_________3'üncü________

11. '0'dan '10'a kadar olan ve okuma, anlama, konuşma ve yazma becerilerinizi nasıl derecelendirirsiniz? '0'ın zayıf ve '10'un iyi anlamına geldiği aşağıya açılır menüden ulaşılabilen ölçekte,

Türkçedeki yetkinlik derecesi							
Okuma	☐	Anlama	☐	Konuşma	☐	Yazma	☐

12. Birinci, ikinci ya da üçüncü dil olarak Arapça konuşuyorsanız, ya da Kur'an-ı Kerim okuma aracılığıyla Arapçayla aşinaysanız, lütfen aşağıdaki 13, 14, 15 ve 16'ncı sorulara ilerleyiniz; aksi takdirde aşağıda bulunan 'Uygun değil' seçeneğine tıklayınız.

☐ İki dili birden biliyorum ya da Kuran okuduğum için Arapça biliyorum

☐ Uygun değil

13. Arapçayı ne kadar süredir kullanmaktasınız?

____

14. Arapça bilgisinizi özellikle hangi bağlamda kullanırsınız? (Örneğin, günlük yaşamdan veya Kuran-ı Kerim'i okuyabilen yada her ikisinden dolayı veya diğer bağlamlarda)

----

15. '0'dan '10'a kadar olan ve '0'ın zayıf ve '10'un iyi anlamına geldiği aşağıya ölçekte, okuma, anlama, konuşma ve yazma becerilerinizi nasıl derecelendirirsiniz?

Arapçadaki yetkinlik derecesi							
Okuma	☐	Anlama	☐	Konuşma	☐	Yazma	☐

16. Lütfen كتاب kelimesini Arapçadan Türkçeye çevirin.

Teşekkür ederim!

Appendix 4-2: Information sheet and consent form

1 Information Sheet

Hello and thank you for visiting this webpage. My name is Shadiya al-Hashmi and I am a Linguistics PhD student at the University of York. Feel free to contact me at any time with questions or comments relating to the study or for the results using the contact details provided below and at the end of the survey. Please completely read the "Questions You May Be Asking" section below before commencing the survey.

Title of the research: **The sound system of Turkish**

QUESTIONS YOU MAY BE ASKING

What is the research about?
The aim of the study is to find out how non-Turkish words are pronounced and written by Turkish speakers, and whether knowing another language besides Turkish makes a difference. The experiment on this webpage is for people who know Turkish (only) or for people who know Turkish and Arabic (either from daily life or from reading the Quran or both).

Who can participate?

Participants can be people who know Turkish (only) or those who know Turkish and Arabic (either from daily life or from reading the Quran or both).

What does the study involve?

The study involves a number of tasks such as listening to recordings and reading words from screen and then responding by identifying what vowel you hear from a given list (in the 1st part of the study) and writing down the words you hear in Turkish (in the 2nd part of the study).

Do I have to take part?

No, participation is entirely voluntary. You are free to quit at any time before the end of the survey, and your data will be destroyed and will not be used in the research. To quit, just exit your browser or navigate to a different website.

Are there any risks to taking part?

No risks are involved in taking part in this study.

Are there any benefits to participating?

You will be enhancing our knowledge of how non-Turkish words are pronounced and written by Turkish speakers, and whether knowing another language besides Turkish makes a difference. In addition, there will be a prize drawing of a $100 Amazon gift card for those interested to enter the draw.

What will happen to the data I provide?

The data you provide will be used alongside the data of other participants to statistically analyze group results in order to either confirm or refute the research hypotheses. Your data will be stored securely in the University of York, Department of Language and Linguistic Science.

What about confidentiality?

Your identity will be kept strictly confidential. Your IP address will not be recorded. No real names will be used in any presentations or publications or in my dissertation. In the event that you email me for any reason, your email address will be stored securely.

Will I know the results?

You may contact me (the researcher) for the results of the survey via email after I finish analyzing the results of the study.

Contact Details:

Researcher name: Shadiya al-Hashmi
Department of Language and Linguistic Science, University of York
Email: saah500@york.ac.uk

Supervisors name and details
Dr. Sam Hellmuth
Department of Language and Linguistic Science, University of York
Email: sam.hellmuth@york.ac.uk

The study has been reviewed and approved by the Departmental Ethics Committee of the Department of Language and Linguistic Science. If you have any questions regarding this, you can contact the head of the ethics committee, Traci Walker, email: traci.walker@york.ac.uk

By clicking "Next" and beginning the survey, you confirm that you:

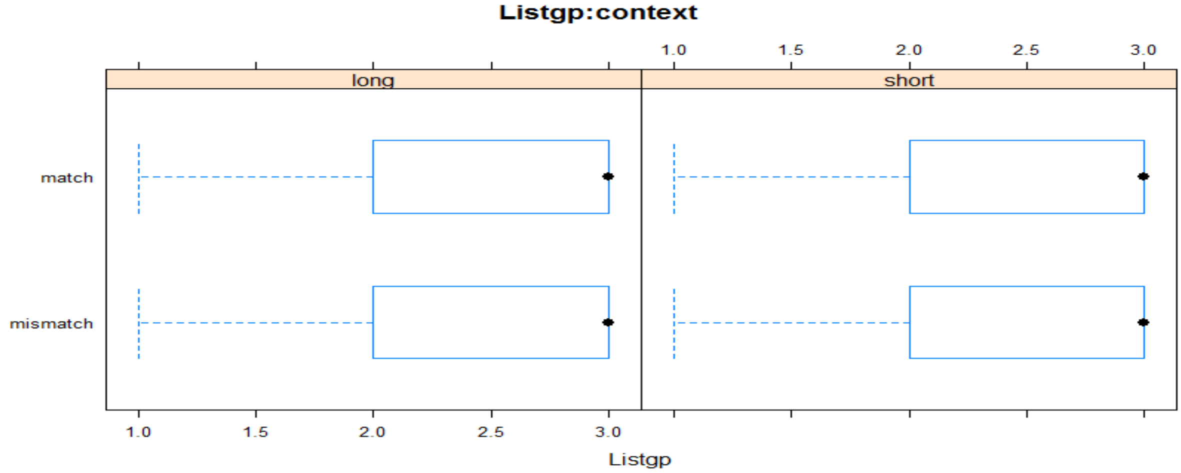
- Have read and understood the above information
- understand that the information you provide will be held in confidence by the researcher, and your name or identifying information about you will not be mentioned in any publication
- Understand that you can withdraw at any time before the end of the survey if you no longer wish to take part in the survey, and that in such a case all your data will be destroyed
- Agree to participate in the study

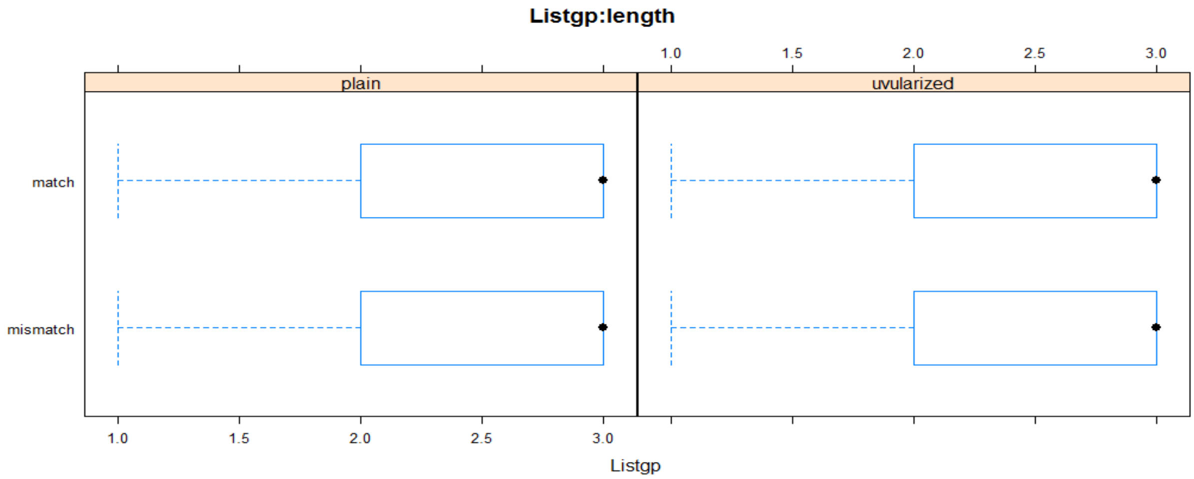
2 Consent form

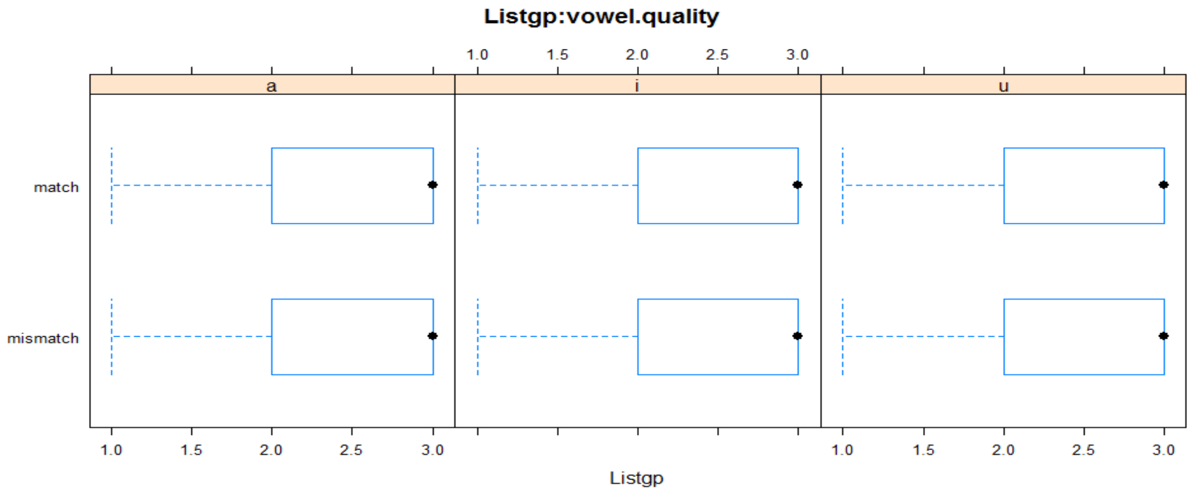
This form is for you to state whether or not you agree to take part in the study. Please read and answer every question. If there is anything you do not understand, or if you want more information, please email the researcher at saah500@york.ac.uk

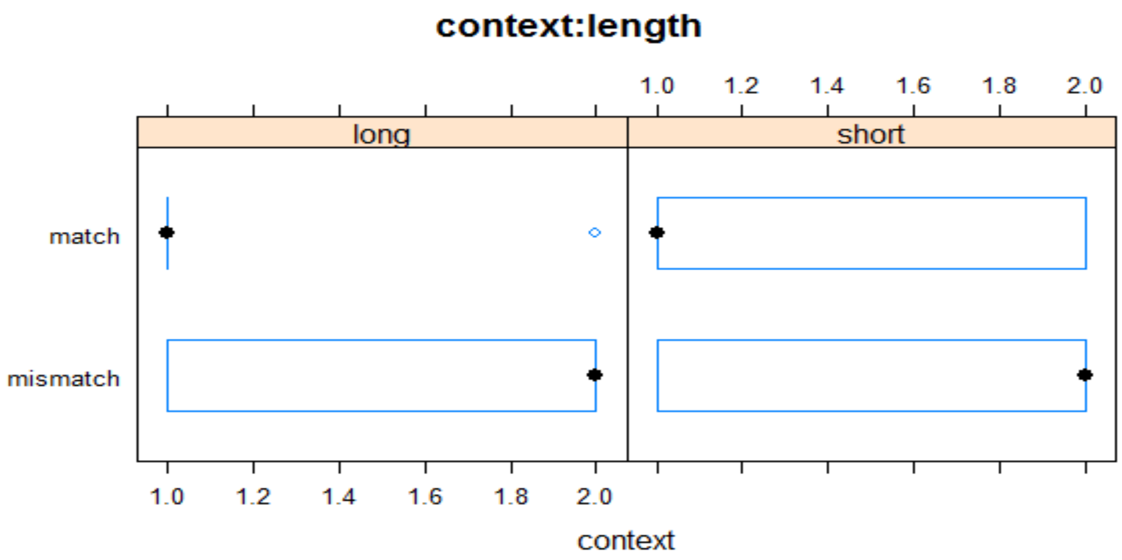
Have you read and understood the information about the study?	Yes ❒ No ❒
Have you had an opportunity to ask questions about the study or email the researcher and have these been answered satisfactorily?	Yes ❒ No ❒
Do you understand that the information you provide will be held in confidence by the researcher, and your name or identifying information about you will not be mentioned in any publication?	Yes ❒ No ❒
Do you understand that you may withdraw from the study at any time before the end of the data collection session without giving any reason, and that in such a case all your data will be destroyed?	Yes ❒ No ❒
Do you agree to take part in the study?	Yes ❒ No ❒

Appendix 4-3: data visualization

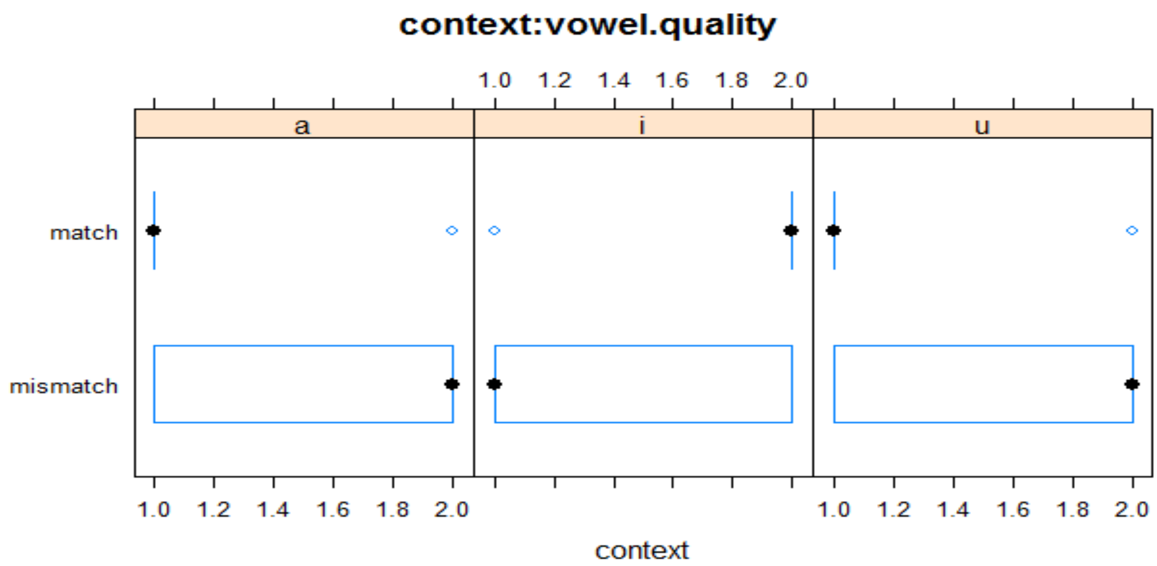




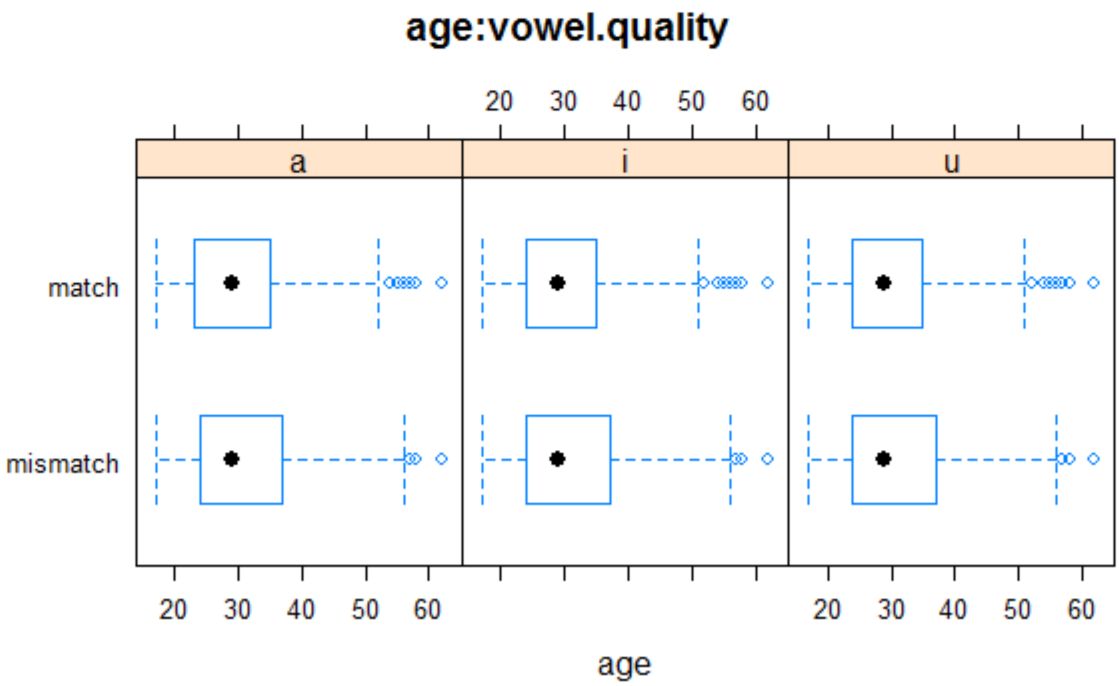


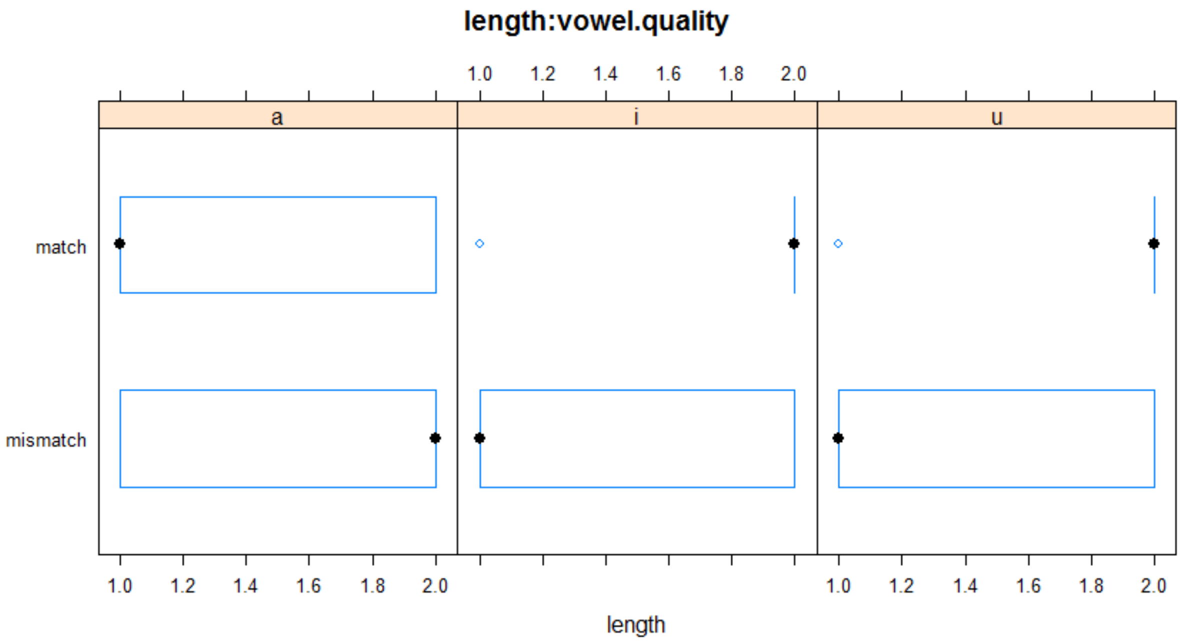


context:length
long
short
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
context



context:vowel.quality
a
i
u
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
context


age:vowel.quality
a
i
u
match
mismatch
20
30
40
50
60
age


length:vowel.quality
a
i
u
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
length

Appendix 4-4: Logistic regression step-wise model selection

modelPATset<-glmer(match~Listgp+context+length+age+ vowel.quality + Listgp:length + Listgp:context + Listgp:vowel.quality +context:length + context:vowel.quality+ age:vowel.quality + (Listgp|stimulus)+(vowel.quality+length+context|listener) , data = PATset, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.284792	1.654989	0.172	0.86337	
ListgpTA	0.342786	0.605126	0.566	0.57107	
ListgpTQ	0.308373	0.435176	0.709	0.47856	
contextuvularized	-4.195321	1.721005	-2.438	0.01478	*
lengthshort	1.834031	1.511649	1.213	0.22503	
age	-0.013337	0.009758	-1.367	0.17173	
vowel.qualityi	-3.664941	1.856179	-1.974	0.04833	*
vowel.qualityu	-2.706976	1.889257	-1.433	0.15191	
ListgpTA:lengthshort	-0.600504	0.551291	-1.089	0.27604	
ListgpTQ:lengthshort	-0.650172	0.402534	-1.615	0.10627	
ListgpTA:contextuvularized	0.205941	0.516040	0.399	0.68983	
ListgpTQ:contextuvularized	0.139500	0.368180	0.379	0.70477	
ListgpTA:vowel.qualityi	0.265055	0.615715	0.430	0.66684	
ListgpTQ:vowel.qualityi	0.208071	0.433322	0.480	0.63110	
ListgpTA:vowel.qualityu	0.462933	0.640003	0.723	0.46948	
ListgpTQ:vowel.qualityu	0.630412	0.459616	1.372	0.17019	
contextuvularized:lengthshort	0.602552	1.245700	0.484	0.62859	
contextuvularized:vowel.qualityi	4.261465	1.476958	2.885	0.00391	**
contextuvularized:vowel.qualityu	2.597769	1.544575	1.682	0.09259	.
age:vowel.qualityi	0.023681	0.011000	2.153	0.03134	*
age:vowel.qualityu	0.007302	0.011736	0.622	0.53380	
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> dropterm(modelPATset, scale = 0, test = "Chisq", k = 2, sorted = FALSE,
trace = TRUE)
trying - Listgp:length
trying - Listgp:context
trying - Listgp:vowel.quality
trying - context:length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + age + vowel.quality + Listgp:length +
Listgp:context + Listgp:vowel.quality + context:length +
context:vowel.quality + age:vowel.quality + (Listgp | stimulus) +
(vowel.quality + length + context | listener)

	Df	AIC	LRT	Pr(Chi)
<none>		5091.8		

Listgp:length	2	5090.4	2.6068	0.2716	
Listgp:context	2	5087.9	0.1453	0.9299	
Listgp:vowel.quality	4	5086.4	2.6386	0.6200	
context:length	1	5090.0	0.2037	0.6518	
context:vowel.quality	2	5094.5	6.6880	0.0353	*
age:vowel.quality	2	5092.1	4.3437	0.1140	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_1<-update(modelPATset, .~.-Listgp:context)
> dropterm(step_1, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - context:length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions

Model:
match ~ Listgp + context + length + age + vowel.quality + (Listgp |
    stimulus) + (vowel.quality + length + context | listener) +
    Listgp:length + Listgp:vowel.quality + context:length + context:vowel.
quality +
    age:vowel.quality
```

	Df	AIC	LRT	Pr(Chi)	
<none>		5087.9			
Listgp:length	2	5086.5	2.5994	0.27262	
Listgp:vowel.quality	4	5082.5	2.6189	0.62348	
context:length	1	5086.2	0.2411	0.62345	
context:vowel.quality	2	5090.5	6.5977	0.03693	*
age:vowel.quality	2	5088.3	4.3521	0.11349	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_2<-update(step_1, .~.-Listgp:vowel.quality)
> dropterm(step_2, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - context:length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + age + vowel.quality + (Listgp |
stimulus) + (vowel.quality + length + context | listener) +
Listgp:length + context:length + context:vowel.quality +
age:vowel.quality

	Df	AIC	LRT	Pr(Chi)
<none>		5082.5		
Listgp:length	2	5080.7	2.1076	0.34862
context:length	1	5080.8	0.2173	0.64108

context:vowel.quality	2	5084.9	6.3862	0.04104 *
age:vowel.quality	2	5082.8	4.2618	0.11873
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

```
> step_3<-update(step_2, .~.-Listgp:length)
> dropterm(step_3, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp
trying - context:length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions

Model:
match ~ Listgp + context + length + age + vowel.quality + (Listgp |
    stimulus) + (vowel.quality + length + context | listener) +
    context:length + context:vowel.quality + age:vowel.quality
```

	Df	AIC	LRT	Pr(Chi)
<none>		5080.7		
Listgp	2	5077.0	0.3543	0.83764
context:length	1	5078.9	0.2405	0.62386
context:vowel.quality	2	5082.9	6.2491	0.04396 *
age:vowel.quality	2	5080.9	4.2067	0.12205
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

```
> step_4<-update(step_3, .~.-Listgp)
> dropterm(step_4, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - context:length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions
```

Model:
match ~ context + length + age + vowel.quality + (Listgp | stimulus) +
(vowel.quality + length + context | listener) + context:length +
context:vowel.quality + age:vowel.quality

	Df	AIC	LRT	Pr(Chi)
<none>		5077.0		
context:length	1	5075.2	0.2221	0.6374
context:vowel.quality	2	5079.5	6.4528	0.0397 *
age:vowel.quality	2	5077.2	4.1840	0.1234
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

```
> step_5<-update(step_4, .~.-context:length)
> dropterm(step_5, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - length
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions
```

Model:
match ~ context + length + age + vowel.quality + (Listgp | stimulus) +

(vowel.quality + length + context | listener) + context:vowel.quality +
age:vowel.quality

	Df	AIC	LRT	Pr(Chi)	
<none>		5075.2			
length	1	5074.3	1.0949	0.29539	
context:vowel.quality	2	5077.5	6.2800	0.04328	*
age:vowel.quality	2	5075.4	4.1811	0.12362	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_6<-update(step_5, .~.-length)
> dropterm(step_6, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - context:vowel.quality
trying - age:vowel.quality
Single term deletions
```

Model:
match ~ context + age + vowel.quality + (Listgp | stimulus) +
(vowel.quality + length + context | listener) + context:vowel.quality +
age:vowel.quality

	Df	AIC	LRT	Pr(Chi)	
<none>		5074.3			
context:vowel.quality	2	5076.2	5.8668	0.05322	.
age:vowel.quality	2	5074.5	4.1998	0.12247	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> summary(step_6)
```

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.001281	0.839931	1.192	0.23322	
contextuvularized	-3.510293	1.119183	-3.136	0.00171	**
age	-0.013855	0.009713	-1.426	0.15375	
vowel.qualityi	-3.047220	1.179543	-2.583	0.00978	**
vowel.qualityu	-1.627860	1.225611	-1.328	0.18411	
contextuvularized:vowel.qualityi	4.194038	1.590575	2.637	0.00837	**
contextuvularized:vowel.qualityu	2.796262	1.646459	1.698	0.08944	.
age:vowel.qualityi	0.023305	0.010964	2.126	0.03354	*
age:vowel.qualityu	0.007432	0.011743	0.633	0.52678	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

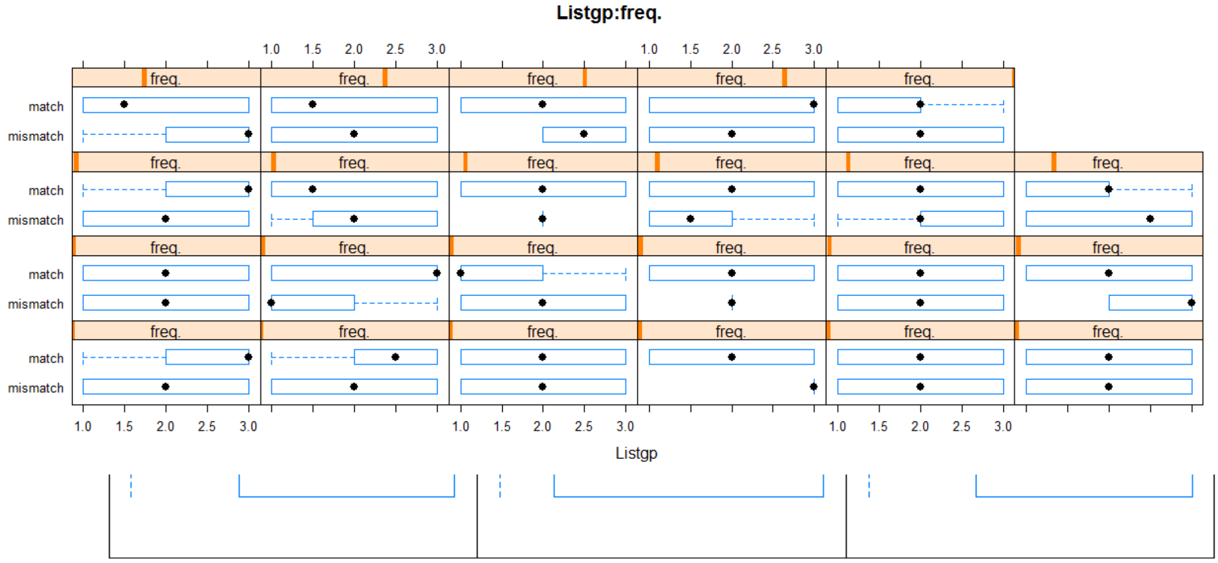
step_7<-glmer(match~context+age+ vowel.quality + length + context:vowel.quality+ length: vowel.quality + age:vowel.quality + (Listgp|stimulus)+(vowel.quality+length+context|listene r) , data = PATset, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAG Q =1)

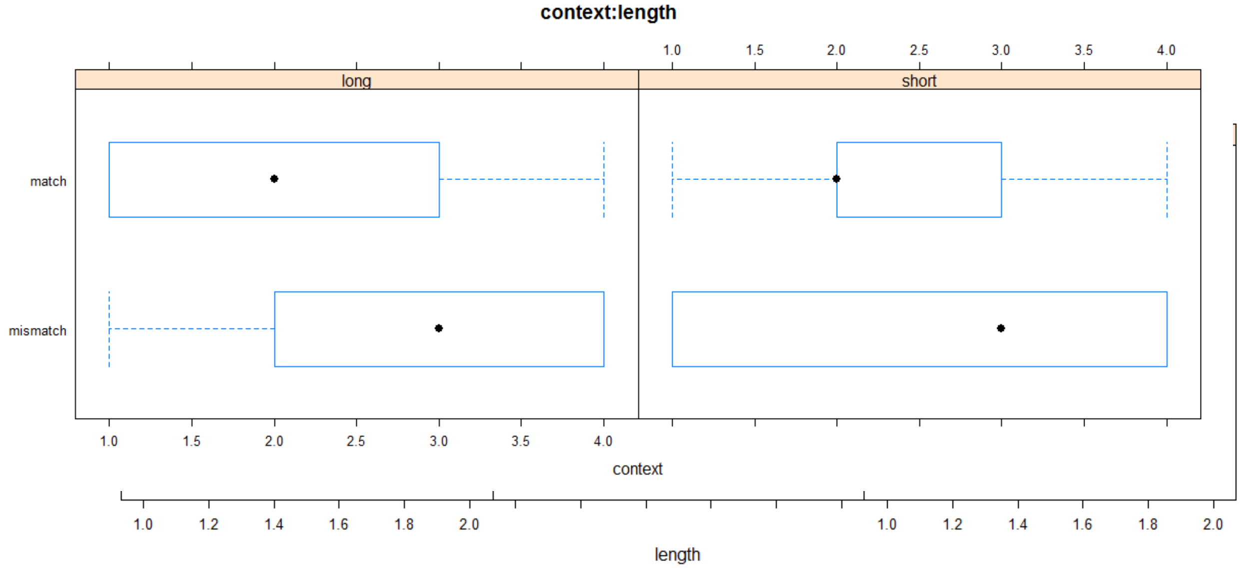
```
> summary(step_7)
```

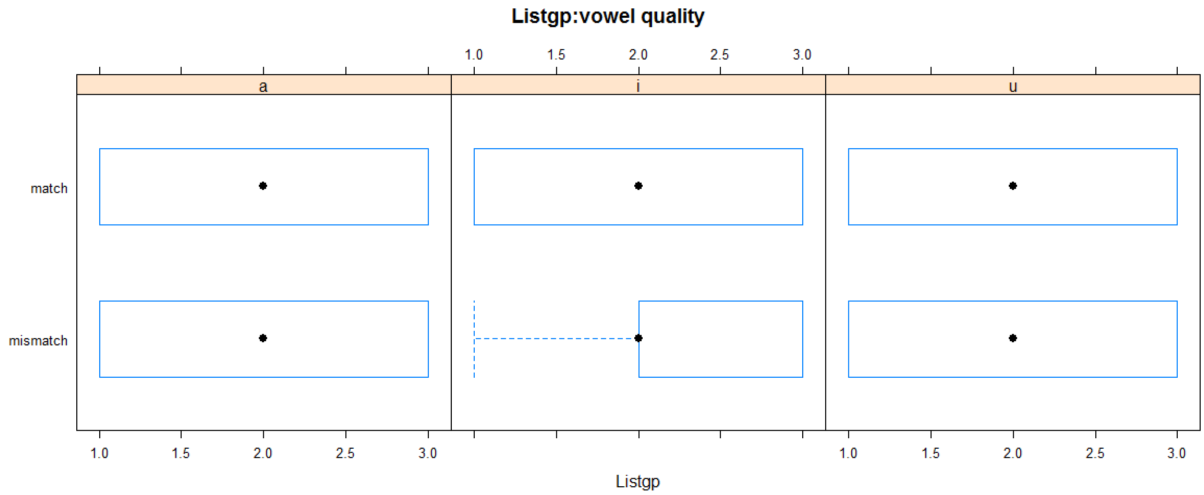
Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.492582	0.865270	1.725	0.084529	.
contextuvularized	-3.648663	0.896210	-4.071	4.68e-05	***
age	-0.013691	0.009707	-1.410	0.158399	

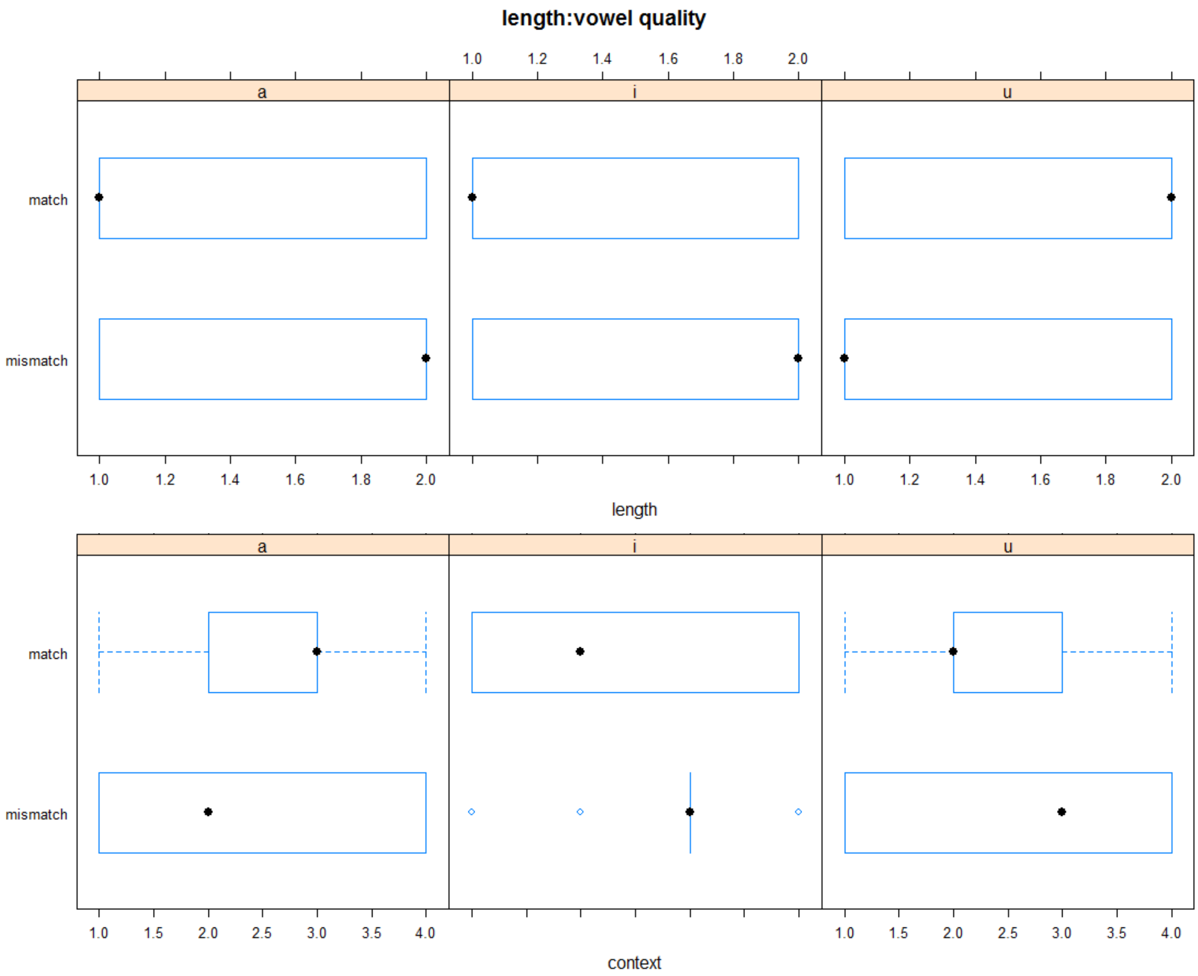
lengthshort	-0.776080	0.962151	-0.807	0.419892	
vowel.qualityi	-5.136480	1.280216	-4.012	6.02e-05	***
vowel.qualityu	-2.892269	1.232041	-2.348	0.018898	*
contextuvularized:vowel.qualityi	4.731263	1.279049	3.699	0.000216	***
contextuvularized:vowel.qualityu	2.269573	1.309578	1.733	0.083085	.
lengthshort:vowel.qualityi	3.041907	1.337849	2.274	0.022982	*
lengthshort:vowel.qualityu	2.475961	1.396428	1.773	0.076217	.
age:vowel.qualityi	0.023302	0.010959	2.126	0.033477	*
age:vowel.qualityu	0.007495	0.011709	0.640	0.522098	
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

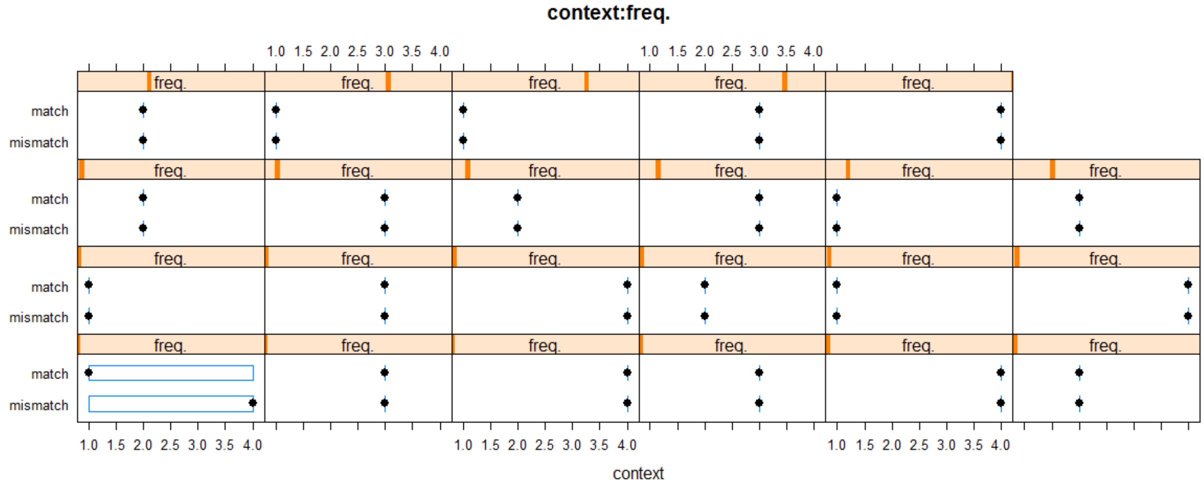
Appendix 5-1: Data exploration plots



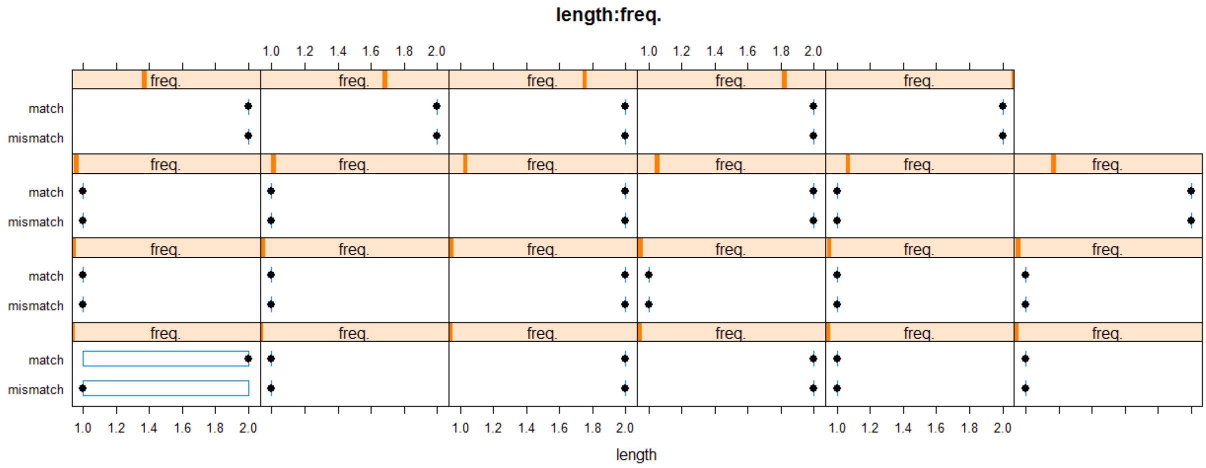




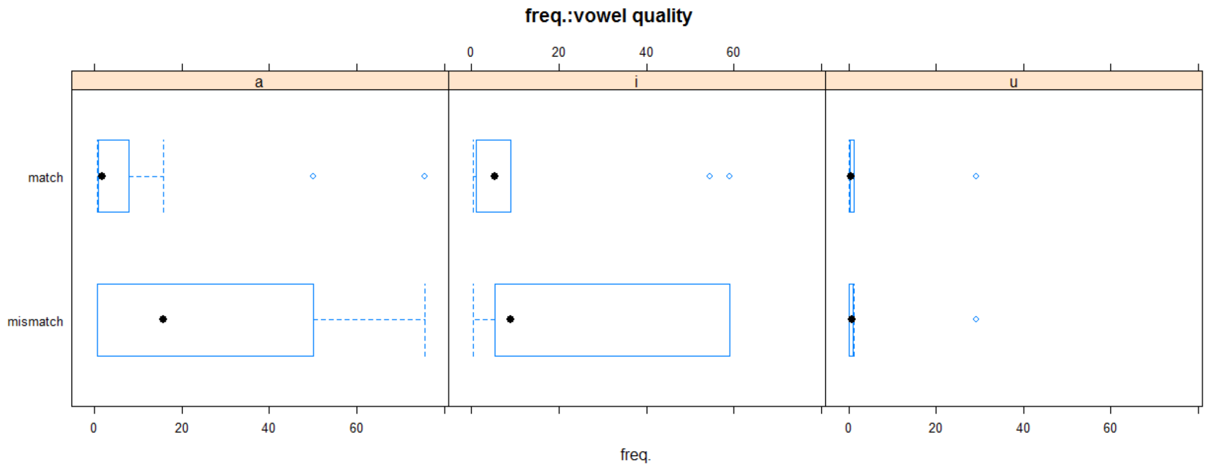

length:vowel quality
a
i
u
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
length
1.5
2.5
3.0
3.5
4.0
context

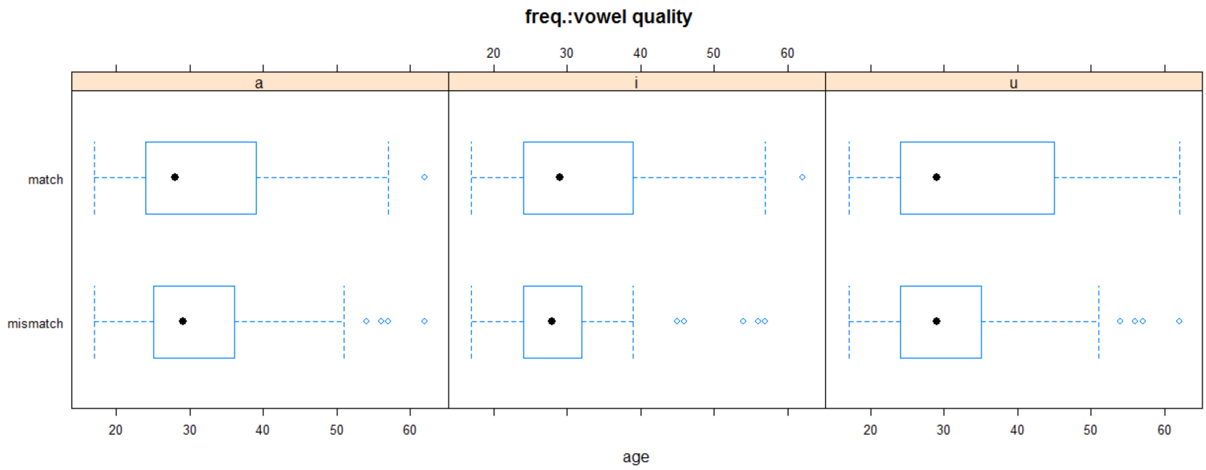


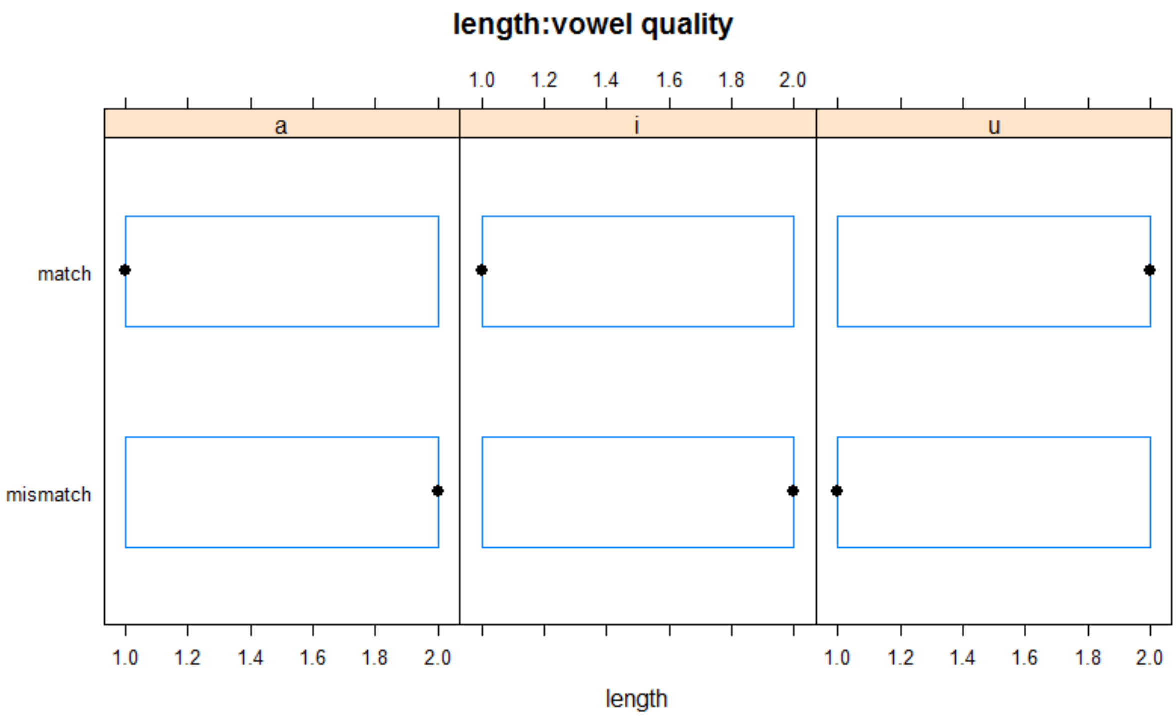
context:freq.
freq.
match
mismatch
1.0 1.5 2.0 2.5 3.0 3.5 4.0
context



length:freq.
freq.
match
mismatch
1.0 1.2 1.4 1.6 1.8 2.0
length


freq.:vowel quality
a
i
u
match
mismatch
0
20
40
60
freq.


freq.:vowel quality
a
i
u
match
mismatch
20
30
40
50
60
age


length:vowel quality
a
i
u
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
length

Appendix 5-2: Summary tables of fixed effects regression models

databasedmsba1<-glmer(match~Listgp + context+length+freq.+vowel.quality+ age+ Listgp:length + Listgp:freq. + Listgp:context + Listgp:vowel.quality + context:length+context:freq.+ context:vowel.quality+length:vowel.quality+ age:vowel.quality +freq.:vowel.quality + (Listgp|stimulus) + (context + length|listener) , data = msba1 , family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.5038843	0.8323136	-0.605	0.544912	
ListgpTA	0.1103299	0.9388776	0.118	0.906454	
ListgpTQ	0.2905367	0.8317390	0.349	0.726855	
contextpharyngeal	-2.1192874	0.8720994	-2.430	0.015095	*
contextplain	1.7333792	0.6972650	2.486	0.012920	*
contextq	-0.3689900	0.6947857	-0.531	0.595360	
lengthshort	-5.3857188	1.6539282	-3.256	0.001129	**
freq.	0.0911038	0.0346033	2.633	0.008468	**
vowel.qualityi	1.9252034	1.1781023	1.634	0.102226	
vowel.qualityu	-0.1753534	0.8712295	-0.201	0.840487	
age	-0.0004011	0.0149775	-0.027	0.978634	
ListgpTA:lengthshort	0.3020473	0.9820664	0.308	0.758415	
ListgpTQ:lengthshort	0.3789083	0.8867678	0.427	0.669167	
ListgpTA:freq.	-0.0040919	0.0164830	-0.248	0.803940	
ListgpTQ:freq.	-0.0212656	0.0163178	-1.303	0.192501	
ListgpTA:contextpharyngeal	-0.4673536	0.7538805	-0.620	0.535303	
ListgpTQ:contextpharyngeal	-0.4028556	0.7236783	-0.557	0.577748	
ListgpTA:contextplain	0.6074585	0.7913297	0.768	0.442699	
ListgpTQ:contextplain	1.3050237	0.7248495	1.800	0.071796	.
ListgpTA:contextq	0.1316402	0.7762492	0.170	0.865337	
ListgpTQ:contextq	-0.5072573	0.7592379	-0.668	0.504061	
ListgpTA:vowel.qualityi	-3.1809115	0.7612005	-4.179	2.93e-05	***
ListgpTQ:vowel.qualityi	-1.1275190	0.7314758	-1.541	0.123212	
ListgpTA:vowel.qualityu	-0.4860686	0.6492470	-0.749	0.454059	
ListgpTQ:vowel.qualityu	-0.6027517	0.6102960	-0.988	0.323330	
contextpharyngeal:lengthshort	-7.0386580	2.4992213	-2.816	0.004857	**
contextplain:lengthshort	5.9620366	1.3217566	4.511	6.46e-06	***
contextq:lengthshort	1.0655139	0.8912732	1.195	0.231893	
contextpharyngeal:freq.	0.7919074	0.2161283	3.664	0.000248	***
contextplain:freq.	-0.1422381	0.0343181	-4.145	3.40e-05	***
contextq:freq.	-0.0488685	0.0230343	-2.122	0.033875	*
contextpharyngeal:vowel.qualityi	4.6796594	1.5171519	3.085	0.002039	**
contextplain:vowel.qualityi	-5.1392891	1.0170320	-5.053	4.34e-07	***
contextq:vowel.qualityi	0.7816474	1.0452407	0.748	0.454571	

contextpharyngeal:vowel.qualityu	0.7541106	0.8065328	0.935	0.349787	
contextplain:vowel.qualityu	-3.4327476	0.8394297	-4.089	4.33e-05	***
contextq:vowel.qualityu	-2.0698468	0.8283821	-2.499	0.012466	*
lengthshort:vowel.qualityi	5.0648242	1.7496259	2.895	0.003794	**
lengthshort:vowel.qualityu	4.8843178	1.3682859	3.570	0.000357	***
vowel.qualityi:age	0.0733599	0.0238888	3.071	0.002134	**
vowel.qualityu:age	0.0244899	0.0174103	1.407	0.159536	
freq.:vowel.qualityi	-0.0701187	0.0431046	-1.627	0.103799	
freq.:vowel.qualityu	-0.5393535	0.1582398	-3.408	0.000653	***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> dropterm(databasedmsba1, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace = TRUE)
trying - Listgp:length
trying - Listgp:freq.
trying - Listgp:context
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + age +
Listgp:length + Listgp:freq. + Listgp:context + Listgp:vowel.quality +
context:length + context:freq. + context:vowel.quality +
length:vowel.quality + age:vowel.quality + freq.:vowel.quality +
(Listgp | stimulus) + (context + length | listener)

	Df	AIC	LRT	Pr(Chi)	
<none>		1106.5			
Listgp:length	2	1102.8	0.220	0.8956622	
Listgp:freq.	2	1104.5	1.967	0.3739601	
Listgp:context	6	1101.9	7.326	0.2917313	
Listgp:vowel.quality	4	1127.1	28.511	9.827e-06	***
context:length	3	1136.3	35.704	8.649e-08	***
context:freq.	3	1129.7	29.184	2.049e-06	***
context:vowel.quality	6	1151.9	57.372	1.536e-10	***
length:vowel.quality	2	1120.2	17.663	0.0001460	***
vowel.quality:age	2	1112.6	10.048	0.0065774	**
freq.:vowel.quality	2	1118.2	15.680	0.0003938	***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_1<-update(databasedmsba1, .~.-Listgp:length)
> dropterm(step_1, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:freq.
trying - Listgp:context
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + age +
(Listgp | stimulus) + (context + length | listener) + Listgp:freq. +
Listgp:context + Listgp:vowel.quality + context:length +
context:freq. + context:vowel.quality + length:vowel.quality +
vowel.quality:age + freq.:vowel.quality

	Df	AIC	LRT	Pr(Chi)	
<none>		1102.8			
Listgp:freq.	2	1100.8	1.973	0.3728867	
Listgp:context	6	1098.0	7.196	0.3031408	
Listgp:vowel.quality	4	1114.4	19.652	0.0005848	***
context:length	3	1132.4	35.612	9.047e-08	***
context:freq.	3	1125.9	29.137	2.096e-06	***
context:vowel.quality	6	1148.1	57.344	1.557e-10	***
length:vowel.quality	2	1116.8	18.049	0.0001204	***
vowel.quality:age	2	1108.8	10.042	0.0065970	**
freq.:vowel.quality	2	1114.5	15.726	0.0003847	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> step_2<-update(step_1, .~.-Listgp:freq.)
> dropterm(step_2, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:context
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + age +
(Listgp | stimulus) + (context + length | listener) + Listgp:context +
Listgp:vowel.quality + context:length + context:freq. + context:vowel.quality +
length:vowel.quality + vowel.quality:age + freq.:vowel.quality

	Df	AIC	LRT	Pr(Chi)
<none>		1100.8		
Listgp:context	6	1096.1	7.311	0.2930340

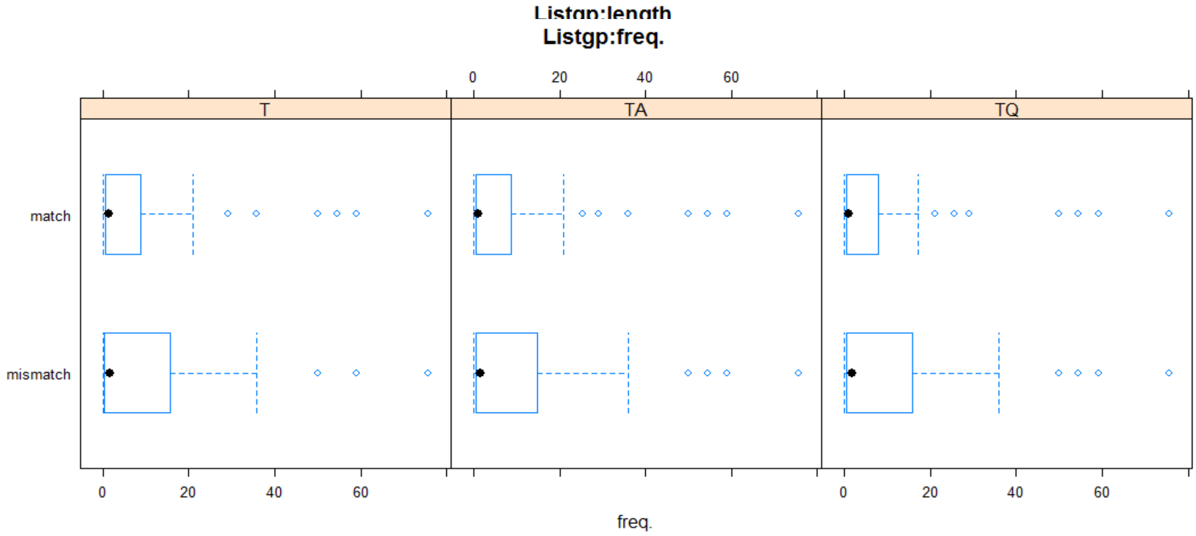
Listgp:vowel.quality	4	1112.2	19.502	0.0006262	***
context:length	3	1130.0	35.297	1.054e-07	***
context:freq.	3	1124.2	29.450	1.801e-06	***
context:vowel.quality	6	1146.9	58.150	1.068e-10	***
length:vowel.quality	2	1114.9	18.189	0.0001123	***
vowel.quality:age	2	1106.9	10.195	0.0061124	**
freq.:vowel.quality	2	1112.8	16.086	0.0003213	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

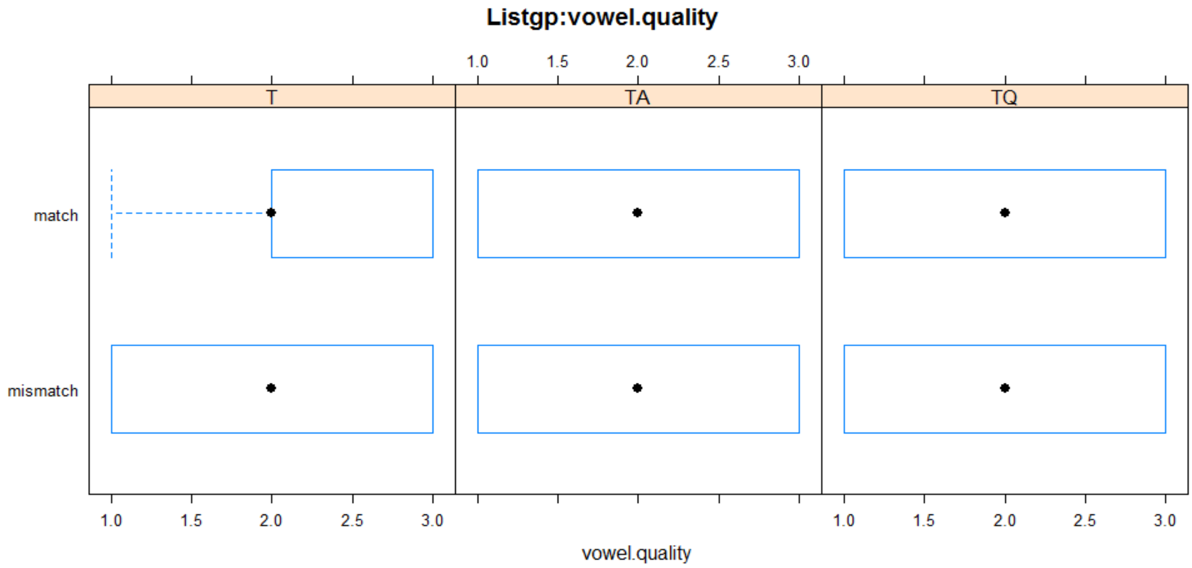
Summary(step_2)

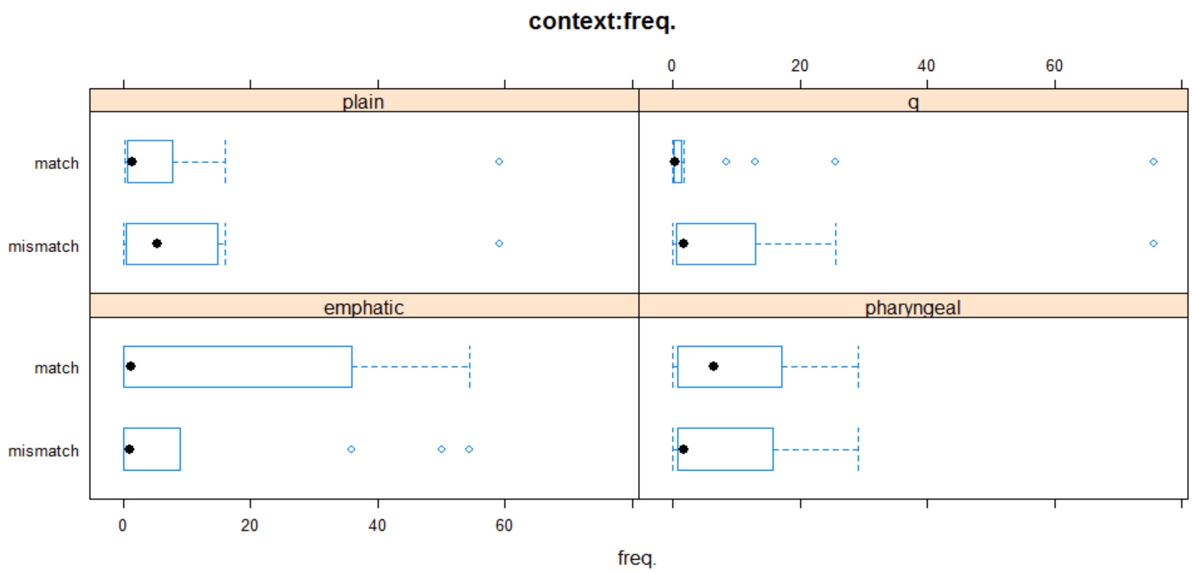
Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.4119516	0.7635410	-0.540	0.589523	
ListgpTA	0.2076415	0.6264908	0.331	0.740315	
ListgpTQ	-0.0374532	0.6251703	-0.060	0.952228	
contextpharyngeal	-2.1892083	0.8867330	-2.469	0.013555	*
contextplain	1.7069877	0.7094280	2.406	0.016122	*
contextq	-0.3974466	0.7111787	-0.559	0.576260	
lengthshort	-5.0940069	1.5704405	-3.244	0.001180	**
freq.	0.0804284	0.0334369	2.405	0.016156	*
vowel.qualityi	2.1100369	1.1950146	1.766	0.077446	.
vowel.qualityu	-0.2792246	0.8721506	-0.320	0.748850	
age	-0.0002146	0.0149235	-0.014	0.988525	
ListgpTA:contextpharyngeal	-0.3810594	0.7561737	-0.504	0.614310	
ListgpTQ:contextpharyngeal	-0.2207218	0.7647029	-0.289	0.772859	
ListgpTA:contextplain	0.6280358	0.7855106	0.800	0.423986	
ListgpTQ:contextplain	1.4300320	0.7489715	1.909	0.056220	.
ListgpTA:contextq	0.1778608	0.7902864	0.225	0.821934	
ListgpTQ:contextq	-0.3956727	0.7931782	-0.499	0.617889	
ListgpTA:vowel.qualityi	-3.3131208	0.7808000	-4.243	2.20e-05	***
ListgpTQ:vowel.qualityi	-1.2639100	0.7592963	-1.665	0.095996	.
ListgpTA:vowel.qualityu	-0.4334212	0.6104758	-0.710	0.477721	
ListgpTQ:vowel.qualityu	-0.3640794	0.6066042	-0.600	0.548378	
contextpharyngeal:lengthshort	-7.2016469	2.4924847	-2.889	0.003860	**
contextplain:lengthshort	5.8285269	1.3116794	4.444	8.85e-06	***
contextq:lengthshort	0.9368842	0.9068071	1.033	0.301525	
contextpharyngeal:freq.	0.7978377	0.2152008	3.707	0.000209	***
contextplain:freq.	-0.1401666	0.0339762	-4.125	3.70e-05	***
contextq:freq.	-0.0455771	0.0229981	-1.982	0.047504	*
contextpharyngeal:vowel.qualityi	4.5839259	1.5110872	3.034	0.002417	**
contextplain:vowel.qualityi	-5.2603374	1.0244086	-5.135	2.82e-07	***
contextq:vowel.qualityi	0.7200049	1.0505651	0.685	0.493123	
contextpharyngeal:vowel.qualityu	0.7402418	0.8110165	0.913	0.361383	
contextplain:vowel.qualityu	-3.4457779	0.8433414	-4.086	4.39e-05	***
contextq:vowel.qualityu	-2.0579810	0.8279735	-2.486	0.012935	*

lengthshort:vowel.qualityi	5.1509513	1.7328924	2.972	0.002954	**
lengthshort:vowel.qualityu	4.8654053	1.3551603	3.590	0.000330	***
vowel.qualityi:age	0.0732599	0.0239829	3.055	0.002253	**
vowel.qualityu:age	0.0243314	0.0174065	1.398	0.162164	
freq.:vowel.qualityi	-0.0725039	0.0426335	-1.701	0.089012	.
freq.:vowel.qualityu	-0.5409178	0.1572894	-3.439	0.000584	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

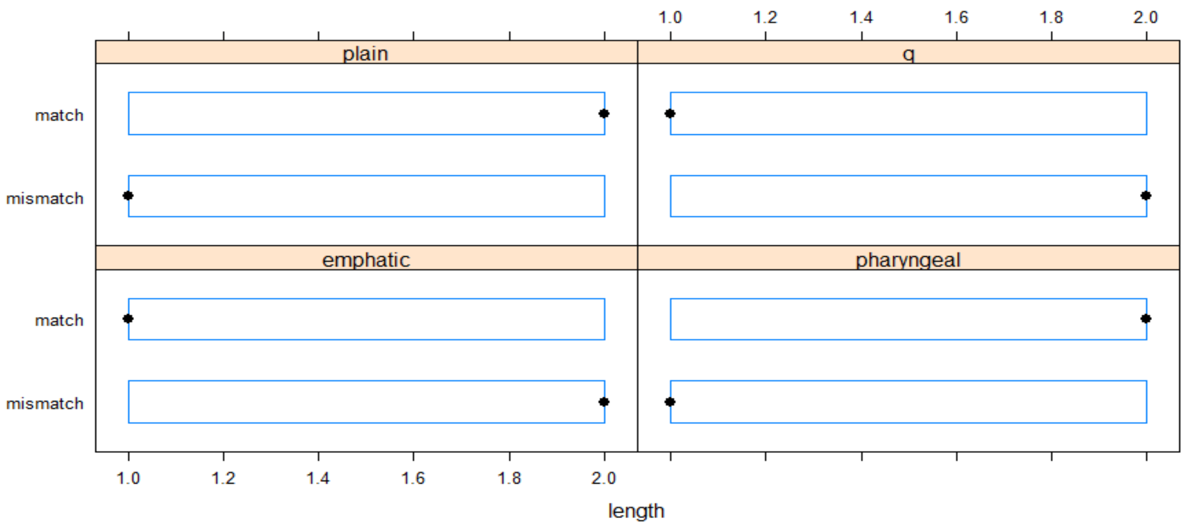
Appendix 6-1: SBAAW dataset data exploration



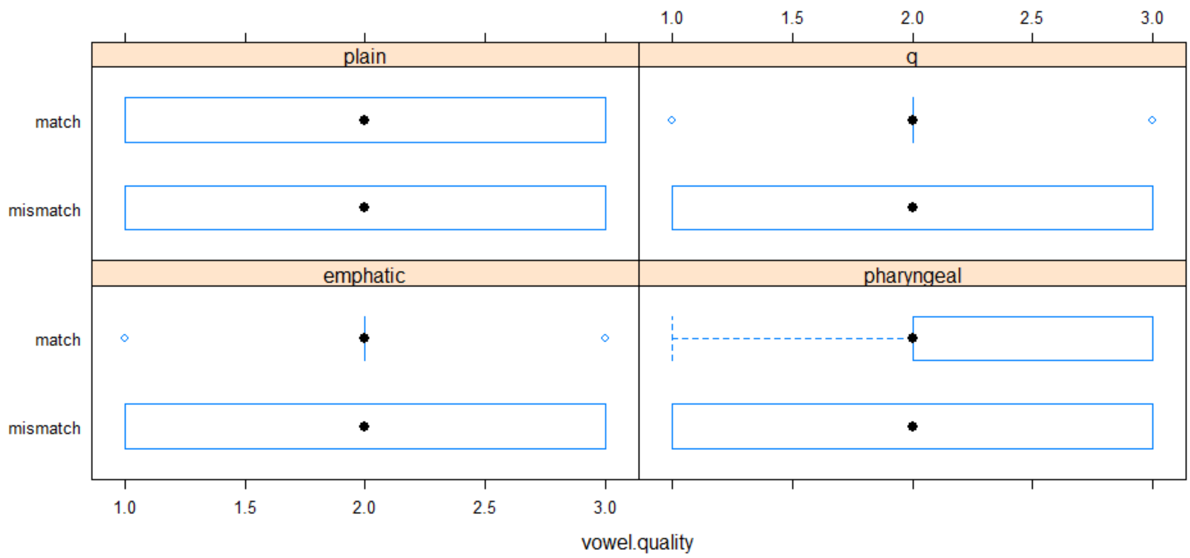




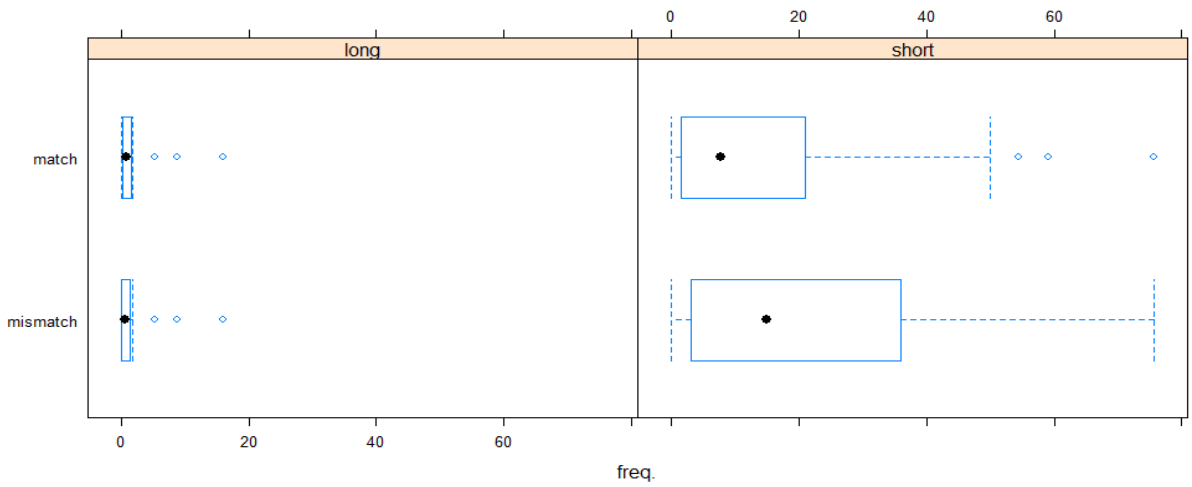
context:length

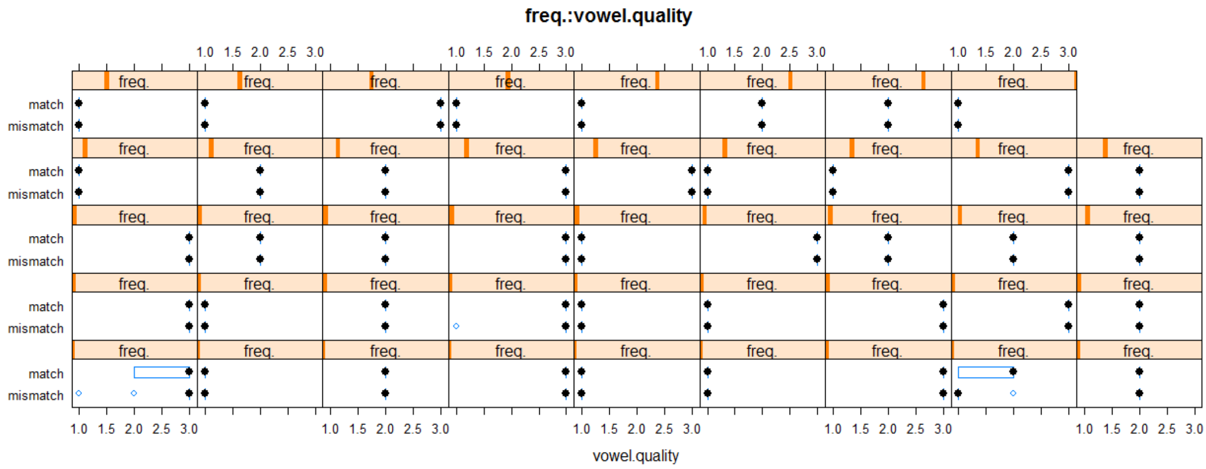


context:vowel.quality

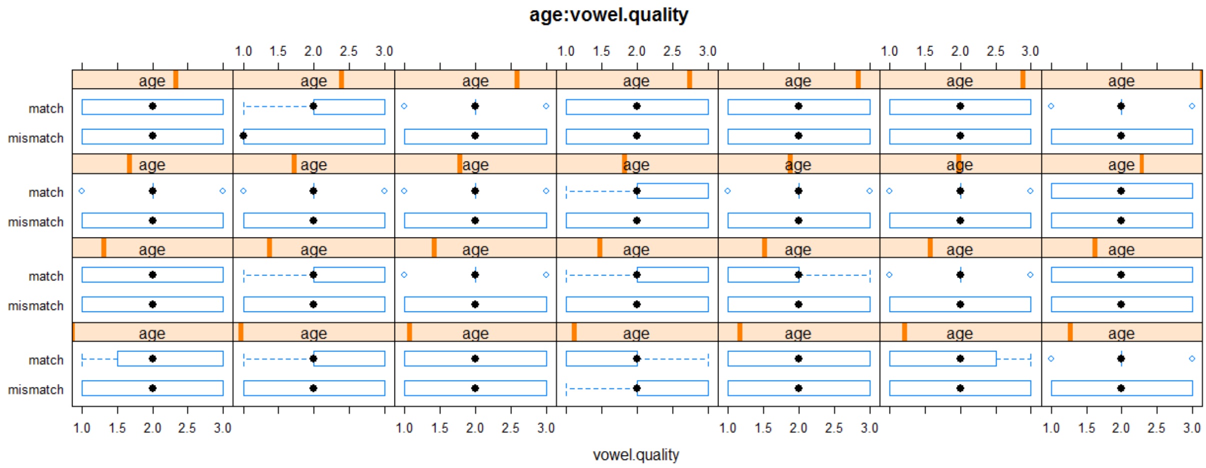


length:freq.





freq.:vowel.quality
freq.
match
mismatch
vowel.quality



age:vowel.quality
age
match
mismatch
vowel.quality

Appendix 6-2: SB Audio and audio-written regression models

stpmodel<-glmer(match~Listgp + context + length + freq. + vowel.quality + stimulus.present ation + age +Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality + Listgp:stimu lus.presentation+ context:length + context:freq. + context:vowel.quality + length:freq. + age :vowel.quality + freq.:vowel.quality + (Listgp|stimulus) + (stimulus.presentation|listener) , d ata = SBAAW , family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ = 1)

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.023286	0.703470	-0.033	0.973593	
ListgpTA	0.387843	0.526329	0.737	0.461193	
ListgpTQ	0.272108	0.505923	0.538	0.590683	
contextpharyngeal	-0.227506	0.696490	-0.327	0.743935	
contextplain	0.459979	0.635811	0.723	0.469402	
contextq	-0.261577	0.677829	-0.386	0.699568	
lengthshort **	-2.680630	0.626953	-4.276	1.91e-05	*
freq.	0.067751	0.063522	1.067	0.286163	
vowel.qualityi	0.540797	0.851753	0.635	0.525479	
vowel.qualityu	-0.254307	0.772653	-0.329	0.742054	
stimulus.presentationaudio+written *	-0.716525	0.264663	-2.707	0.006783	*
age	0.002028	0.013494	0.150	0.880511	
ListgpTA:lengthshort	0.704622	0.392403	1.796	0.072549	.
ListgpTQ:lengthshort	-0.142910	0.385067	-0.371	0.710541	
ListgpTA:contextpharyngeal	-0.436963	0.457631	-0.955	0.339661	
ListgpTQ:contextpharyngeal	0.031859	0.463020	0.069	0.945143	
ListgpTA:contextplain	0.448176	0.468256	0.957	0.338508	
ListgpTQ:contextplain	0.340396	0.451120	0.755	0.450514	
ListgpTA:contextq	0.355460	0.462224	0.769	0.441880	
ListgpTQ:contextq	0.066240	0.467190	0.142	0.887250	
ListgpTA:freq.	-0.017687	0.012823	-1.379	0.167812	
ListgpTQ:freq.	-0.016326	0.012645	-1.291	0.196681	
ListgpTA:vowel.qualityi **	-2.446858	0.410670	-5.958	2.55e-09	*
ListgpTQ:vowel.qualityi	-0.511971	0.417485	-1.226	0.220078	
ListgpTA:vowel.qualityu	-1.020602	0.396408	-2.575	0.010035	*
ListgpTQ:vowel.qualityu	-0.048221	0.392811	-0.123	0.902298	
ListgpTA:stimulus.presentationaudio+written	-0.255395	0.370459	-0.689	0.490571	
ListgpTQ:stimulus.presentationaudio+written	-0.784534	0.358953	-2.186	0.028844	*
contextpharyngeal:lengthshort *	3.611735	1.241036	2.910	0.003611	*
contextplain:lengthshort **	4.598890	0.789268	5.827	5.65e-09	*
contextq:lengthshort	1.057497	0.866090	1.221	0.222085	
contextpharyngeal:freq.	-0.123480	0.067098	-1.840	0.065726	.
contextplain:freq. **	-0.163716	0.028057	-5.835	5.37e-09	*
contextq:freq.	-0.041509	0.026367	-1.574	0.115427	

contextpharyngeal:vowel.qualityi	1.637780	0.954088	1.717	0.086054	.
contextplain:vowel.qualityi	-1.518703	0.851683	-1.783	0.074557	.
contextq:vowel.qualityi	1.227316	0.957468	1.282	0.199900	
contextpharyngeal:vowel.qualityu	-0.309369	0.927714	-0.333	0.738776	
contextplain:vowel.qualityu	-0.655663	0.847096	-0.774	0.438924	
contextq:vowel.qualityu	-0.760498	0.924589	-0.823	0.410777	
lengthshort:freq.	-0.035142	0.064587	-0.544	0.586373	
vowel.qualityi:age	0.040325	0.012794	3.152	0.001623	**
vowel.qualityu:age	0.007222	0.011335	0.637	0.524041	
freq.:vowel.qualityi	0.050492	0.025581	1.974	0.048406	*
freq.:vowel.qualityu	0.154024	0.043596	3.533	0.000411	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> dropterm(stpmodel, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace = TRUE)
trying - Listgp:length
trying - Listgp:context
trying - Listgp:freq.
trying - Listgp:vowel.quality
trying - Listgp:stimulus.presentation
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:freq.
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + stimulus.presentation + age + Listgp + Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality + Listgp:stimulus.presentation + context:length + context:freq. + context:vowel.quality + length:freq. + age:vowel.quality + freq.:vowel.quality + (Listgp | stimulus) + (stimulus.presentation | listener)

	Df	AIC	LRT	Pr(Chi)	
<none>		2456.0			
Listgp:length	2	2457.8	5.8342	0.054091	.
Listgp:context	6	2450.2	6.1784	0.403508	
Listgp:freq.	2	2454.2	2.1880	0.334879	
Listgp:vowel.quality	4	2479.3	31.3117	2.644e-06	***
Listgp:stimulus.presentation	2	2457.2	5.2393	0.072829	.
context:length	3	2477.3	27.2577	5.198e-06	***
context:freq.	3	2480.1	30.0903	1.321e-06	***
context:vowel.quality	6	2462.5	18.4965	0.005104	**
length:freq.	1	2454.3	0.3071	0.579488	
vowel.quality:age	2	2462.8	10.7996	0.004517	**
freq.:vowel.quality	2	2465.6	13.5576	0.001138	**
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> step_1<-update(stpmodel, .~.-length:freq.)
```

```
> dropterm(step_1, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:context
trying - Listgp:freq.
trying - Listgp:vowel.quality
trying - Listgp:stimulus.presentation
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + stimulus.presentation +
age + (Listgp | stimulus) + (stimulus.presentation | listener) +
Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality +
Listgp:stimulus.presentation + context:length + context:freq. +
context:vowel.quality + vowel.quality:age + freq.:vowel.quality

	Df	AIC	LRT	Pr(Chi)	
<none>		2454.3			
Listgp:length	2	2456.1	5.8003	0.0550147	.
Listgp:context	6	2448.5	6.1718	0.4042220	
Listgp:freq.	2	2452.5	2.1769	0.3367358	
Listgp:vowel.quality	4	2477.6	31.2557	2.715e-06	***
Listgp:stimulus.presentation	2	2455.6	5.2686	0.0717703	.
context:length	3	2475.6	27.2787	5.146e-06	***
context:freq.	3	2478.8	30.4460	1.112e-06	***
context:vowel.quality	6	2461.0	18.6613	0.0047756	**
vowel.quality:age	2	2461.1	10.7662	0.0045936	**
freq.:vowel.quality	2	2465.0	14.6813	0.0006486	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> step_2<-update(step_1, .~.-Listgp:context)
> dropterm(step_2, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:freq.
trying - Listgp:vowel.quality
trying - Listgp:stimulus.presentation
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + stimulus.presentation +
age + (Listgp | stimulus) + (stimulus.presentation | listener) +
Listgp:length + Listgp:freq. + Listgp:vowel.quality + Listgp:stimulus.presentation +
context:length + context:freq. + context:vowel.quality +
vowel.quality:age + freq.:vowel.quality

	Df	AIC	LRT	Pr(Chi)
<none>		2448.5		

Listgp:length	2	2448.9	4.4322	0.1090322	
Listgp:freq.	2	2446.4	1.9566	0.3759522	
Listgp:vowel.quality	4	2468.6	28.1160	1.182e-05	***
Listgp:stimulus.presentation	2	2449.7	5.2138	0.0737630	.
context:length	3	2469.5	27.0248	5.817e-06	***
context:freq.	3	2472.7	30.2445	1.226e-06	***
context:vowel.quality	6	2454.8	18.3376	0.0054416	**
vowel.quality:age	2	2455.2	10.7539	0.0046219	**
freq.:vowel.quality	2	2458.7	14.2154	0.0008188	***
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> step_3<-update(step_2, .~.-Listgp:freq.)
> dropterm(step_3, scale = 0, test = "Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - Listgp:stimulus.presentation
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - vowel.quality:age
trying - freq.:vowel.quality
Single term deletions
```

Model:
match ~ Listgp + context + length + freq. + vowel.quality + stimulus.presentation +
age + (Listgp | stimulus) + (stimulus.presentation | listener) +
Listgp:length + Listgp:vowel.quality + Listgp:stimulus.presentation +
context:length + context:freq. + context:vowel.quality +
vowel.quality:age + freq.:vowel.quality

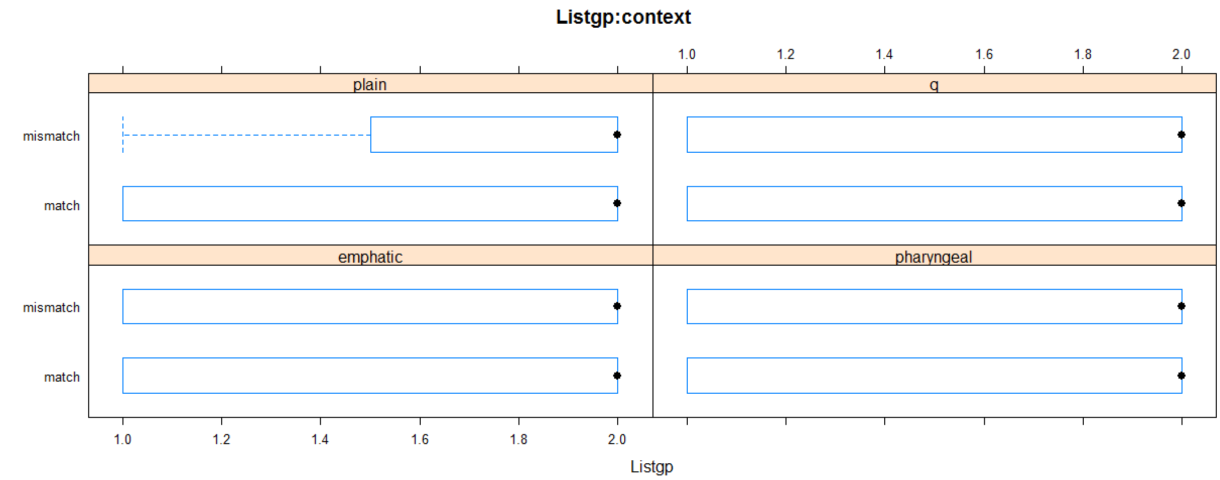
	Df	AIC	LRT	Pr(Chi)	
<none>		2446.4			
Listgp:length	2	2448.4	6.0050	0.049663	*
Listgp:vowel.quality	4	2466.0	27.5447	1.542e-05	***
Listgp:stimulus.presentation	2	2447.0	4.5108	0.104830	
context:length	3	2467.0	26.5765	7.222e-06	***
context:freq.	3	2469.7	29.2727	1.963e-06	***
context:vowel.quality	6	2452.9	18.4161	0.005272	**
vowel.quality:age	2	2453.2	10.7909	0.004537	**
freq.:vowel.quality	2	2456.2	13.7907	0.001013	**
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

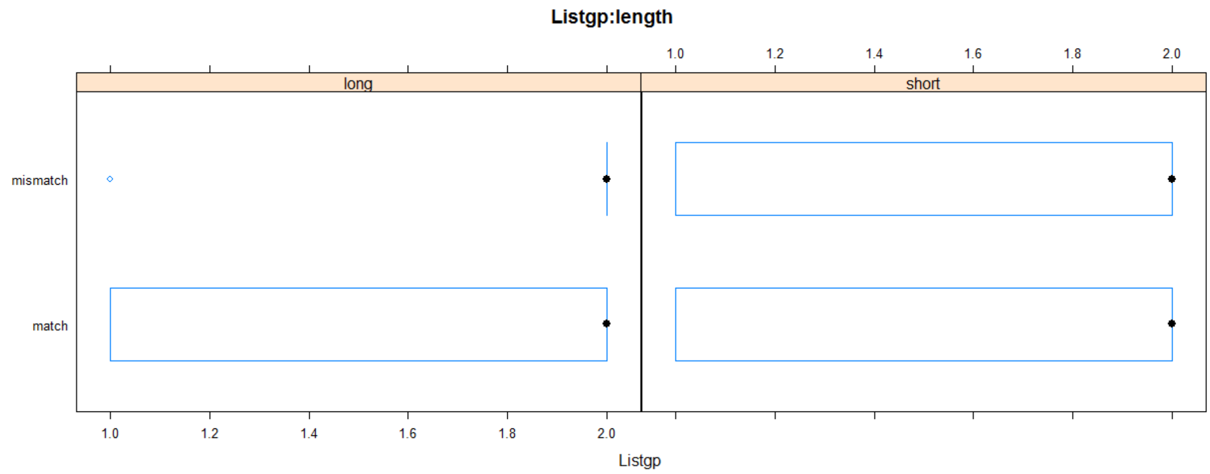
```
> summary(step_4)
```

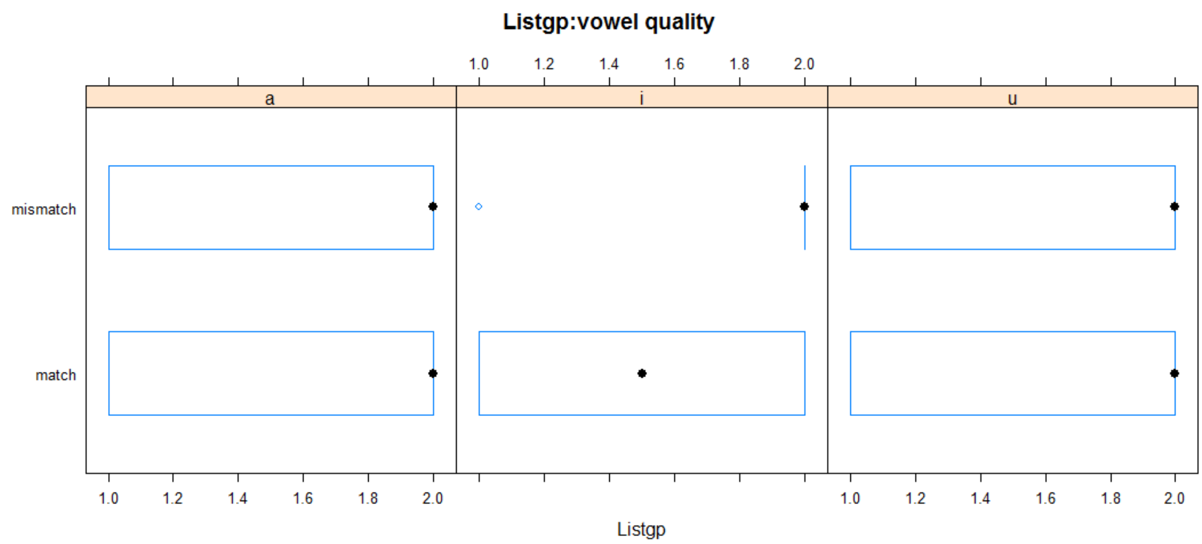
Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.116582	0.680685	-0.171	0.864011
ListgpTA	0.325316	0.456759	0.712	0.476325
ListgpTQ	0.250002	0.411839	0.607	0.543825

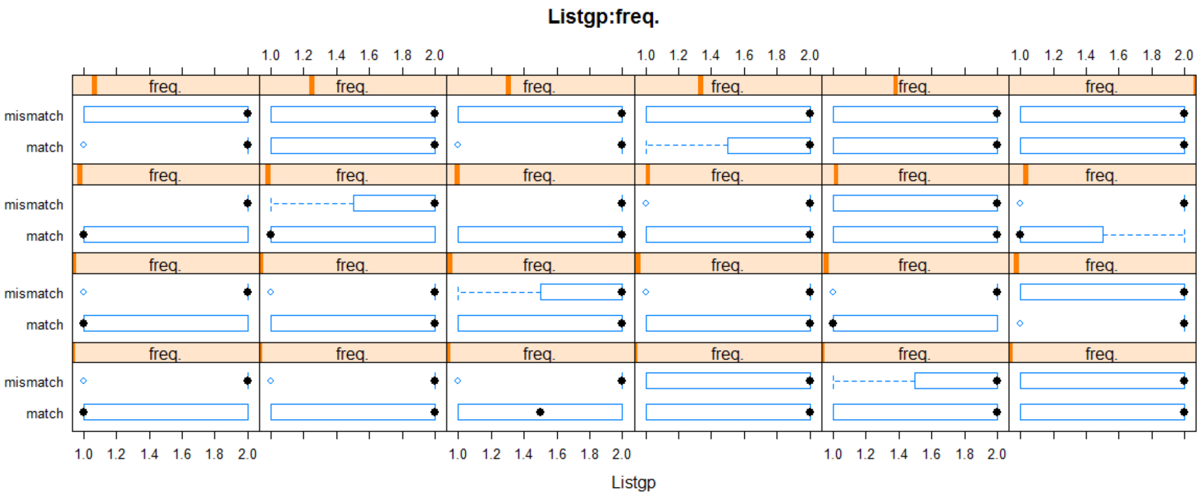
contextpharyngeal	-0.126394	0.641918	-0.197	0.843906	
contextplain	0.791864	0.631386	1.254	0.209781	
contextq	-0.069078	0.645544	-0.107	0.914784	
lengthshort	-3.867144	1.008391	-3.835	0.000126	***
freq.	0.060855	0.028433	2.140	0.032333	*
vowel.qualityi	0.987745	0.803557	1.229	0.218990	
vowel.qualityu	-0.390851	0.747542	-0.523	0.601080	
stimulus.presentationaudio+written	-0.641654	0.255015	-2.516	0.011865	*
age	0.002404	0.013349	0.180	0.857095	
ListgpTA:lengthshort	0.383944	0.345320	1.112	0.266203	
ListgpTQ:lengthshort	-0.351230	0.325415	-1.079	0.280442	
ListgpTA:vowel.qualityi	-2.325697	0.431281	-5.393	6.95e-08	***
ListgpTQ:vowel.qualityi	-0.488148	0.415100	-1.176	0.239604	
ListgpTA:vowel.qualityu	-0.920474	0.411007	-2.240	0.025119	*
ListgpTQ:vowel.qualityu	0.008727	0.385647	0.023	0.981947	
ListgpTA:stimulus.presentationaudio+written	-0.177953	0.383631	-0.464	0.642744	
ListgpTQ:stimulus.presentationaudio+written	-0.669892	0.352482	-1.901	0.057367	.
contextpharyngeal:lengthshort	4.180992	1.164568	3.590	0.000330	***
contextplain:lengthshort	4.738237	0.775841	6.107	1.01e-09	***
contextq:lengthshort	1.419348	0.857238	1.656	0.097778	.
contextpharyngeal:freq.	-0.131919	0.061037	-2.161	0.030672	*
contextplain:freq.	-0.145273	0.025420	-5.715	1.10e-08	***
contextq:freq.	-0.061919	0.027582	-2.245	0.024777	*
contextpharyngeal:vowel.qualityi	1.327294	0.914976	1.451	0.146882	
contextplain:vowel.qualityi	-2.103533	0.866679	-2.427	0.015219	*
contextq:vowel.qualityi	1.022597	0.929646	1.100	0.271338	
contextpharyngeal:vowel.qualityu	-0.553982	0.865729	-0.640	0.522236	
contextplain:vowel.qualityu	-0.907525	0.847620	-1.071	0.284316	
contextq:vowel.qualityu	-0.973517	0.884063	-1.101	0.270816	
lengthshort:vowel.qualityi	0.288984	0.864819	0.334	0.738263	
lengthshort:vowel.qualityu	1.707455	0.849639	2.010	0.044471	*
vowel.qualityi:age	0.040462	0.012811	3.158	0.001586	**
vowel.qualityu:age	0.007409	0.011328	0.654	0.513063	
freq.:vowel.qualityi	0.019460	0.033504	0.581	0.561345	
freq.:vowel.qualityu	0.108146	0.046910	2.305	0.021144	*
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

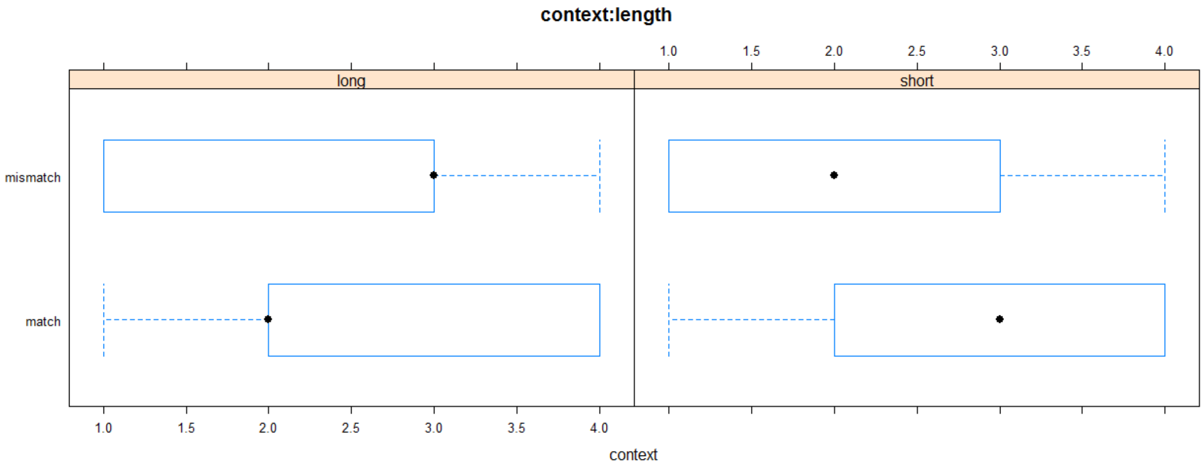
Appendix 6-3: SB-written dataset data exploration

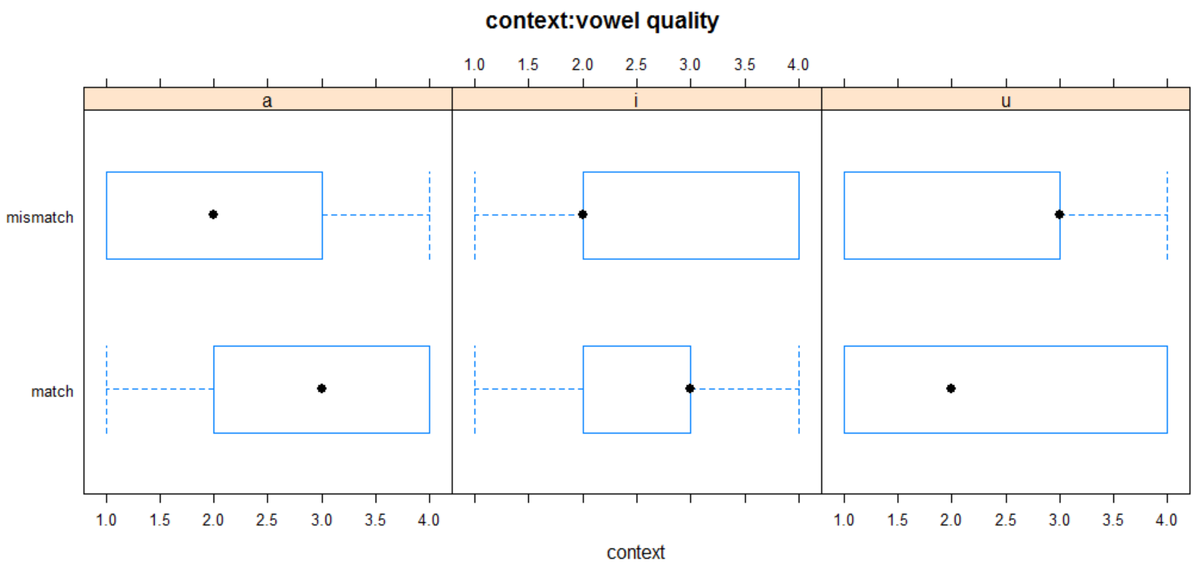


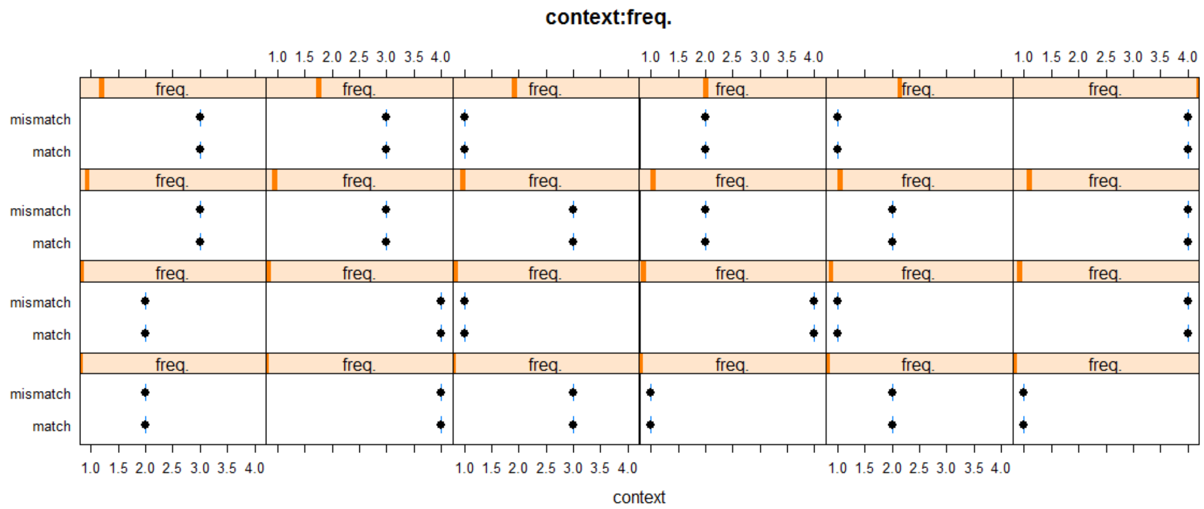


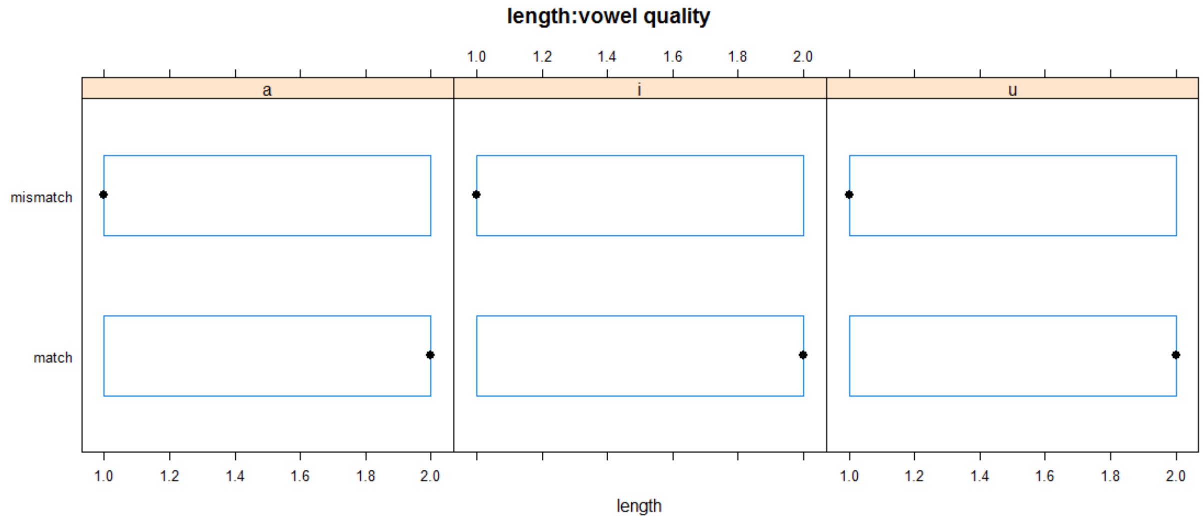


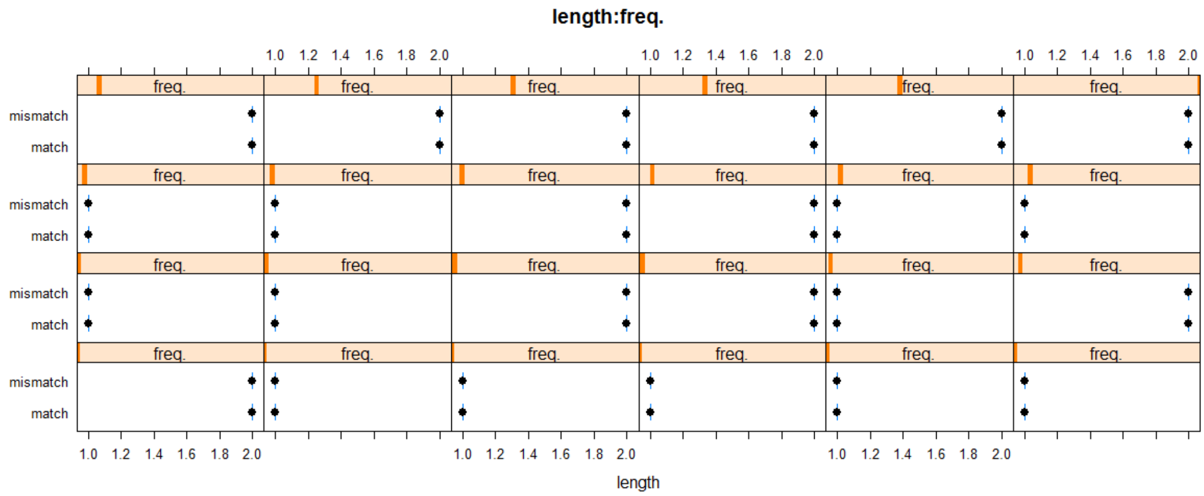

Listgp:freq.
freq.
mismatch
match
1.0 1.2 1.4 1.6 1.8 2.0
Listgp

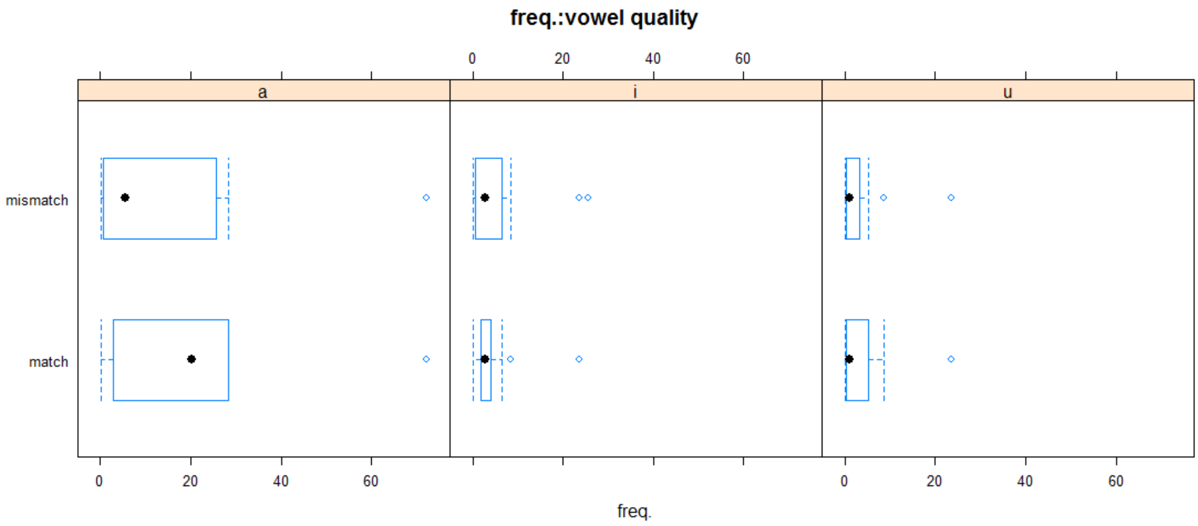

context:length
long
short
mismatch
match
1.0 1.5 2.0 2.5 3.0 3.5 4.0
context

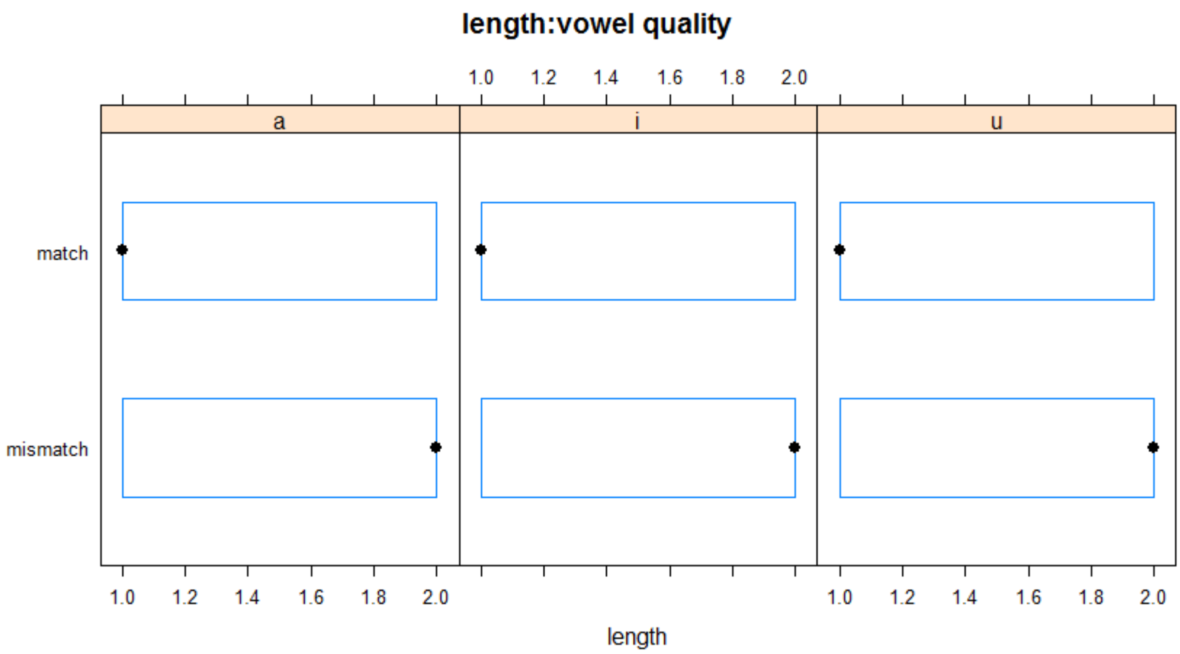

context:vowel quality
a
i
u
mismatch
match
1.0 1.5 2.0 2.5 3.0 3.5 4.0
context


context:freq.
freq.
mismatch
match
context


length:vowel quality
a
i
u
mismatch
match
length


length:freq.
freq.
mismatch
match
length


freq.:vowel quality
a
i
u
mismatch
match
0
20
40
60
freq.


length:vowel quality
a
i
u
match
mismatch
1.0
1.2
1.4
1.6
1.8
2.0
length

Appendix 6-4: SB-written condition regression models

wdatadriven<-glmer(match~Listgp+context+length+vowel.quality+age+freq.+ Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality +context:length+context:freq.+ context:vowel.quality+length:freq.+ age:vowel.quality +freq.:vowel.quality + (Listgp|stimulus) + (length+context|listener) , data = SBwritten, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.045025	1.190330	0.878	0.3800	
ListgpTQ	0.668211	1.006180	0.664	0.5066	
contextpharyngeal	-1.388187	1.363369	-1.018	0.3086	
contextplain	-0.267354	0.924636	-0.289	0.7725	
contextq	-1.037242	0.785232	-1.321	0.1865	
lengthshort	0.720075	0.889892	0.809	0.4184	
vowel.qualityi	-2.545126	1.071763	-2.375	0.0176	*
vowel.qualityu	-1.665911	1.042678	-1.598	0.1101	
age	-0.020167	0.025923	-0.778	0.4366	
freq.	0.021804	0.157871	0.138	0.8902	
ListgpTQ:lengthshort	-1.667604	0.741496	-2.249	0.0245	*
ListgpTQ:contextpharyngeal	0.325962	0.831605	0.392	0.6951	
ListgpTQ:contextplain	0.422577	0.785801	0.538	0.5907	
ListgpTQ:contextq	-0.003625	0.669515	-0.005	0.9957	
ListgpTQ:freq.	-0.004153	0.019714	-0.211	0.8331	
ListgpTQ:vowel.qualityi	2.583188	0.637461	4.052	5.07e-05	***
ListgpTQ:vowel.qualityu	0.886332	0.620258	1.429	0.1530	
contextpharyngeal:lengthshort	-0.611310	0.882524	-0.693	0.4885	
contextplain:lengthshort	-0.488142	1.021300	-0.478	0.6327	
contextq:lengthshort	-1.191944	0.966373	-1.233	0.2174	
contextpharyngeal:freq.	0.085364	0.066690	1.280	0.2005	
contextplain:freq.	0.001137	0.070666	0.016	0.9872	
contextq:freq.	0.041131	0.035804	1.149	0.2506	
contextpharyngeal:vowel.qualityi	1.624351	1.543446	1.052	0.2926	
contextplain:vowel.qualityi	-1.980249	0.954106	-2.076	0.0379	*
contextq:vowel.qualityi	1.336321	1.115198	1.198	0.2308	
contextpharyngeal:vowel.qualityu	0.089599	1.330630	0.067	0.9463	
contextplain:vowel.qualityu	1.149734	0.945783	1.216	0.2241	
contextq:vowel.qualityu	0.490593	0.900277	0.545	0.5858	
lengthshort:freq.	-0.056380	0.153173	-0.368	0.7128	
vowel.qualityi:age	0.043442	0.022611	1.921	0.0547	.
vowel.qualityu:age	0.017837	0.022599	0.789	0.4300	
vowel.qualityi:freq.	0.088991	0.042157	2.111	0.0348	*
vowel.qualityu:freq.	0.070684	0.051105	1.383	0.1666	
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

```
> dropterm(wdatadriven, scale = 0, test ="Chisq", k = 2, sorted = FALSE, t
race = TRUE)
trying - Listgp:length
trying - Listgp:context
```

```
trying - Listgp:freq.
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:freq.
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
Listgp:length + Listgp:context + Listgp:freq. + Listgp:vowel.quality +
context:length + context:freq. + context:vowel.quality +
length:freq. + age:vowel.quality + freq.:vowel.quality +
(Listgp | stimulus) + (length + context | listener)

	Df	AIC	LRT	Pr(Chi)	
<none>		925.08			
Listgp:length	1	928.08	4.9979	0.0253778	*
Listgp:context	3	919.50	0.4155	0.9370246	
Listgp:freq.	1	923.10	0.0144	0.9043198	
Listgp:vowel.quality	2	934.22	13.1386	0.0014028	**
context:length	3	920.76	1.6796	0.6414759	
context:freq.	3	921.99	2.9044	0.4066005	
context:vowel.quality	6	936.45	23.3639	0.0006833	***
length:freq.	1	923.26	0.1737	0.6768652	
vowel.quality:age	2	925.15	4.0692	0.1307332	
vowel.quality:freq.	2	924.99	3.9066	0.1418042	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_1<-update(wdatadriven, .~.-Listgp:context)
> dropterm(step_1, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:freq.
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:freq.
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
(Listgp | stimulus) + (length + context | listener) + Listgp:length +
Listgp:freq. + Listgp:vowel.quality + context:length + context:freq. +
context:vowel.quality + length:freq. + vowel.quality:age +
vowel.quality:freq.

	Df	AIC	LRT	Pr(Chi)	
<none>		919.50			
Listgp:length	1	922.88	5.3854	0.0203064	*
Listgp:freq.	1	917.57	0.0726	0.7875340	
Listgp:vowel.quality	2	928.69	13.1901	0.0013671	**
context:length	3	915.14	1.6474	0.6486871	

context:freq.	3	916.35	2.8560	0.4143635	
context:vowel.quality	6	930.81	23.3172	0.0006969	***
length:freq.	1	917.66	0.1639	0.6855720	
vowel.quality:age	2	919.51	4.0135	0.1344259	
vowel.quality:freq.	2	919.50	4.0071	0.1348547	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> dropterm(step_2, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - length:freq.
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
(Listgp | stimulus) + (length + context | listener) + Listgp:length +
Listgp:vowel.quality + context:length + context:freq. + context:vowel.quality +
length:freq. + vowel.quality:age + vowel.quality:freq.

	Df	AIC	LRT	Pr(Chi)	
<none>		917.57			
Listgp:length	1	922.20	6.6294	0.0100310	*
Listgp:vowel.quality	2	929.19	15.6234	0.0004050	***
context:length	3	913.25	1.6830	0.6407182	
context:freq.	3	914.51	2.9448	0.4002140	
context:vowel.quality	6	929.28	23.7092	0.0005906	***
length:freq.	1	915.74	0.1682	0.6817135	
vowel.quality:age	2	917.58	4.0106	0.1346173	
vowel.quality:freq.	2	917.66	4.0862	0.1296250	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_3<-update(step_2, .~.-length:freq.)
> dropterm(step_3, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - context:length
trying - context:freq.
trying - context:vowel.quality
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
(Listgp | stimulus) + (length + context | listener) + Listgp:length +
Listgp:vowel.quality + context:length + context:freq. + context:vowel.quality +
vowel.quality:age + vowel.quality:freq.

	Df	AIC	LRT	Pr(Chi)
<none>		915.74		

Listgp:length	1	920.37	6.6369	0.0099886	**
Listgp:vowel.quality	2	929.61	17.8675	0.0001319	***
context:length	3	911.68	1.9415	0.5846411	
context:freq.	3	912.51	2.7748	0.4276701	
context:vowel.quality	6	928.62	24.8812	0.0003592	***
vowel.quality:age	2	915.72	3.9849	0.1363612	
vowel.quality:freq.	2	917.57	5.8346	0.0540808	.
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_4<-update(step_3, .~.-context:length)
> dropterm(step_4, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - context:freq.
trying - context:vowel.quality
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
(Listgp | stimulus) + (length + context | listener) + Listgp:length +
Listgp:vowel.quality + context:freq. + context:vowel.quality +
vowel.quality:age + vowel.quality:freq.

	Df	AIC	LRT	Pr(Chi)	
<none>		911.68			
Listgp:length	1	916.25	6.5667	0.0103902	*
Listgp:vowel.quality	2	929.47	21.7901	1.855e-05	***
context:freq.	3	907.89	2.2118	0.5296243	
context:vowel.quality	6	927.31	27.6324	0.0001102	***
vowel.quality:age	2	911.56	3.8825	0.1435273	
vowel.quality:freq.	2	916.61	8.9358	0.0114717	*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> step_5<-update(step_4, .~.-context:freq.)
> dropterm(step_5, scale = 0, test ="Chisq", k = 2, sorted = FALSE, trace
= TRUE)
trying - Listgp:length
trying - Listgp:vowel.quality
trying - context:vowel.quality
trying - vowel.quality:age
trying - vowel.quality:freq.
Single term deletions
```

Model:
match ~ Listgp + context + length + vowel.quality + age + freq. +
(Listgp | stimulus) + (length + context | listener) + Listgp:length +
Listgp:vowel.quality + context:vowel.quality + vowel.quality:age +
vowel.quality:freq.

	Df	AIC	LRT	Pr(Chi)	
<none>		907.89			
Listgp:length	1	912.35	6.4587	0.0110410	*

Listgp:vowel.quality	2	925.40	21.5058	2.138e-05	***
context:vowel.quality	6	921.48	25.5875	0.0002657	***
vowel.quality:age	2	907.55	3.6570	0.1606549	
vowel.quality:freq.	2	913.25	9.3629	0.0092654	**
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

```
> summary(step_5)
```

Fixed effects:					
	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.718943	1.086998	0.661	0.50835	
ListgpTQ	1.005533	0.711515	1.413	0.15759	
contextpharyngeal	-0.008717	0.581685	-0.015	0.98804	
contextplain	-0.075453	0.552763	-0.137	0.89143	
contextq	-1.019399	0.595433	-1.712	0.08689	.
lengthshort	-0.061705	0.519231	-0.119	0.90540	
vowel.qualityi	-2.479175	1.046130	-2.370	0.01780	*
vowel.qualityu	-1.211021	0.941807	-1.286	0.19850	
age	-0.018910	0.026426	-0.716	0.47426	
freq.	-0.000855	0.011188	-0.076	0.93908	
ListgpTQ:lengthshort	-1.712043	0.661080	-2.590	0.00960	**
ListgpTQ:vowel.qualityi	2.683390	0.590444	4.545	5.5e-06	***
ListgpTQ:vowel.qualityu	0.838722	0.553127	1.516	0.12944	
contextpharyngeal:vowel.qualityi	0.536017	0.849033	0.631	0.52783	
contextplain:vowel.qualityi	-1.754800	0.806424	-2.176	0.02955	*
contextq:vowel.qualityi	1.502375	0.911221	1.649	0.09920	.
contextpharyngeal:vowel.qualityu	-1.179042	0.727380	-1.621	0.10503	
contextplain:vowel.qualityu	0.728307	0.714689	1.019	0.30818	
contextq:vowel.qualityu	-0.125062	0.799315	-0.156	0.87567	
vowel.qualityi:age	0.041161	0.022697	1.814	0.06975	.
vowel.qualityu:age	0.016903	0.022646	0.746	0.45543	
vowel.qualityi:freq.	0.098830	0.036753	2.689	0.00717	**
vowel.qualityu:freq.	0.073686	0.046615	1.581	0.11394	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

step_6<-glmer(match~Listgp+context+length+vowel.quality+age+freq.+ Listgp:length+ Listgp:vowel.quality + context:vowel.quality+ **length:vowel.quality** + age:vowel.quality +freq.:vowel.quality + (Listgp|stimulus) + (length+context|listener) , data = SBwritten, family = "binomial", control=glmerControl(optCtrl=list(maxfun=2e5)), nAGQ =1)

Fixed effects	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.833896	1.113904	0.749	0.4541	
ListgpTQ	1.002307	0.726003	1.381	0.1674	
contextpharyngeal	0.003590	0.594173	0.006	0.9952	
contextplain	-0.076145	0.563917	-0.135	0.8926	
contextq	-1.228073	0.684598	-1.794	0.0728	.
lengthshort	-0.466140	0.815756	-0.571	0.5677	
vowel.qualityi	-2.655220	1.070751	-2.480	0.0131	*
vowel.qualityu	-1.278548	0.985371	-1.298	0.1944	
age	-0.019823	0.026858	-0.738	0.4605	
freq.	0.007846	0.017837	0.440	0.6600	

ListgpTQ:lengthshort	-1.726584	0.674180	-2.561	0.0104	*
ListgpTQ:vowel.qualityi	2.718546	0.609496	4.460	8.18e-06	***
ListgpTQ:vowel.qualityu	0.802186	0.572841	1.400	0.1614	
contextpharyngeal:vowel.qualityi	0.444133	0.890653	0.499	0.6180	
contextplain:vowel.qualityi	-1.785754	0.833343	-2.143	0.0321	*
contextq:vowel.qualityi	1.705846	0.971802	1.755	0.0792	.
contextpharyngeal:vowel.qualityu	-1.199898	0.745276	-1.610	0.1074	
contextplain:vowel.qualityu	0.692812	0.731887	0.947	0.3438	
contextq:vowel.qualityu	0.069507	0.868078	0.080	0.9362	
lengthshort:vowel.qualityi	0.581692	0.839912	0.693	0.4886	
lengthshort:vowel.qualityu	0.336401	0.821025	0.410	0.6820	
vowel.qualityi:age	0.042666	0.022842	1.868	0.0618	.
vowel.qualityu:age	0.017526	0.022812	0.768	0.4423	
vowel.qualityi:freq.	0.085704	0.042396	2.021	0.0432	*
vowel.qualityu:freq.	0.073253	0.059680	1.227	0.2197	
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1					

Abbreviations

Transliteration	IPA symbol	Arabic letter
thth	/ð/	ذ
Dh	/dˤ/	ض
Q	/q/	ق
?	/ʕ/ or /ʔ/	ع/ ء
Th	/θ/	ث
H	/ħ/	ح
T	/t/	ت
T	/tˤ/	ط
H	/h/	ه
S	/s/	س
S	/sˤ/	ص
D	/d/	د
M	/m/	م
N	/n/	ن
L	/l/	ل
B	/b/	ب
F	/f/	ف
J	/dʒ/	ج
Y	/j/	ي
W	/w/	و
Z	/z/	ز
R	/r/ or /ɾ/	ر
K	/k/	ك
Kh	/x/	خ
SH	/ʃ/	ش
TH	/ðˤ/	ظ
GH	/ɣ/	غ

Turkish alphabet

Capital letters	Small letters	IPA symbol
A	A	/a/
B	B	/b/
C	C	/dʒ/
Ç	Ç	/tʃ/
D	D	/d/
E	E	/e/
F	F	/f/
G	G	/g/
Ğ	Ğ	No equivalent; a lengthened preceding vowel
H	H	/h/
I	I	/ɯ/
İ	I	/i/
J	J	/ʒ/
K	K	/k/
L	L	/l/
M	M	/m/
N	N	/n/
O	O	/o/
Ö	Ö	œ/
P	P	/p/
R	R	/r/
S	S	/s/
Ş	Ş	/ʃ/
T	T	/t/
U	U	/u/
Ü	Ü	/y/
V	V	/v/
Y	Y	/j/
Z	Z	/z/

References

arabdict. (2008). Retrieved from http://www.arabdict.com/en/عربي-عربي

Baayen, R. (2008). Analyzing Linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.

Aksan, D. (1993). Die neueren Entwicklungen im modernen Türkei-Türkischen. In J. P. Laut & K. Ro_hrborn (Eds.), Sprach- und Kulturkontakte der türkischen Volker (pp. 5-11). Wiesbaden: O. Harrassowitz.

Al-Ani, S. (1970). Arabic phonology: An acoustical and physiological investigation (Vol. 61): Walter de Gruyter.

Al-Ani, S., & El-Dalee, M. (1983). Tafkhim in Arabic: the acoustic and [physiological] parameters. Paper presented at the Proceedings of the Xth International Meeting of Phonetic Sciences, ed. MPR Van den Broecke and A. Cohen.

Albashir, A. (2008). Production and Perception of Libyan Arabic vowels. (Ph.D. dissertation), University of Newcastle Upon Tyne, University of Newcastle Upon Tyne.

Allatif, O. (2008). Contrôle des corrélats temporels et spectraux de la quantité vocalique: de l'arabe syrien de l'Euphrate au français de Savoie. Université Stendhal-Grenoble III.

Almbark, R. (2008). A Sociophonetic Study of Emphasis in Syrian Arabic. University of York.

Almbark, R. (2012). The perception and production of SSBE vowels by Syrian Arabic learners: The foreign language model. (Unpublished doctoral dissertation), University of York, York, UK.

Al-Wer, E. (1998). Raising of /a/ and related vocalic movements in the emerging dialect of Amman. Paper presented at the Third International Conference on Arabic Dialectology, Malta.

Ar-Rajhi, A. Al-Lahajaat al-'arabiyyah fil-Qiraa'aat al-Qur'aaniyyah. Egypt: Dar AlMa'aarif.

Baayen, R. (2008). Analyzing Linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.

Balpinar, Z. (2011). Turkish phonology, morphology and syntax: Anadolu Universitesi.

Barkat, M. (2006a). Vowel backing. In K. Versteegh (Ed.), Encyclopedia of Arabic Language and Linguistics (Vol. 3, pp. 669-674).

Barkat, M. (2006b). Vowel raising. In K. Versteegh (Ed.), Encyclopedia of Arabic Language and Linguistics (Vol. 3, pp. 678-682): Brill.

Barkat, M., Hombert, J., & Taine-Cheikh. (1997). Détermination d'indices acoustiques pour l'identification des parles arabes:Etats d'(a)mes vocaliques. Paper presented at the Actes des Journées d'Etudes Linguistiques La voyelle dans tous ses états, Nantes.

Barkat-Defradas, Al-Tamimi, J., & Benkirane, T. (2003). Phonetic variation in production and perception of speech: A comparative study of two Arabic dialects. Paper presented at the The fifteen international Congress of Phonetic Sciences, Barcelona.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of memory and language, 68(3), 255-278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1).

Benkirane, T. (1981). Durée, prosodie et syllabation en arabe marocain. Travaux de l'Institut de Phonétique d'Aix Aix-en-Provence, 8, 45-84.

Bergsträsser, G. (1918). Zur Phonetik des Türkischen nach gebildeter Konstantinopler Aussprachep. Aynı yazarın, Review of the Grammatik der osmanich-türkischen Sprache by Gotthold Weil, Berlin. ZDMG, 72, 233-262.

Best, C. (1994). Development of Language-specific Influences on Speech Perceptionand Production in Pre-verbal Infancy.

Best, C. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), Speech perception and linguistic experience: Issues in cross-language research.Speech perception and linguistic experience: Issues in cross-language research (pp. 171-204). Baltimore: York Press.

Best, C., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. Journal of phonetics, 20(3), 305-330.

Boersma, P. (2009). Cue constraints and their interactions in phonological perception and production. Phonology in perception, 15, 55-110.

Boersma, P., & Hamann, S. (2009). Loanword adaptation as first-language phonological perception. Loanword phonology, 11-58.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. Linguistic Inquiry, 32(45–86).

Boersma, P., & Weeink, D. (2009). Praat: doing phonetics by computer (Version 5.1.14) [Computer program]. Retrieved from http://www.praat.org/

Calabrese, A., & Wetzels, W. (2009). Loan phonology (Vol. 307): John Benjamins Publishing.

Clements, G. N., & Sezer, E. (1982). Vowel and consonant disharmony in Turkish. The structure of phonological representations, 2, 213-255.

Cowell, M. (1964). A reference grammar of Syrian Arabic. Washington, DC: Georgetown University Press.

Crystal, D. (2008). Dictionary of linguistics and phonetics: Wiley-Blackwell.

Davis, S. (1995). Emphasis spread in Arabic and grounded phonology. Linguistic Inquiry, 465-498.

Detey, S., & Nespoulous, J.-L. (2008). Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters. Lingua, 118(1), 66-81.

Dohlus, K. (2013). *Phonetics or phonology: Asymmetries in loanword adaptations; French and German mid front rounded vowels in Japanese*. Universitätsbibliothek Johann Christian Senckenberg.

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.

Eren, H. (1999). Türk dilinin etimoloji sözlüğü. Ankara: Bizim Büro basımevi.

Faber, A. (1997). Genetic subgrouping of the Semitic languages. In R. Hetzron (Ed.). The Semitic Languages (pp. 3-15). London: Routledge.

Ferguson, C. A. (1959). Diglossia. word, 15(2), 325-340.

Fischer, W. (1969). Probleme der silbenstruktur im Arabischen. Paper presented at the Proceedings of the International Conference on Semitic Studies.

Flege, J. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. Journal of Phonetics, 15, 47–65.

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), Speech perception and linguistic experience: Theoretical and methodological issues (pp. 233–277). Baltimore: York Press.

Fox, J. (2003). Effects displays in R for generalised linear models. Journal of statistical software, 8(15), 1-17.

Gairdner, W. H. T. (1925). The phonetics of Arabic. London: Oxford University Press.

Ghazali, S. (1983). La coarticulation de l'empathse en arabe. Arabica, 28, 251-277.

Gilichinskaya, Y., & Strange, W. (2010). Perceptual assimilation of American English vowels by inexperienced Russian listeners. The Journal of the Acoustical Society of America, 128, 80-85.

Göksel, A., & Kerslake, C. (2005). Turkish: A comprehensive grammar: Psychology Press.

Greenberg, J. (1950). The patterning of root morphemes in Semitic. Word, 6(2), 162-181.

Haugen, E. (1950). The analysis of linguistic borrowing. Language, 26(2), 210-231.

Heffernan, K. (2007). The role of phonemic contrast in the formation of Sino-Japanese. Journal of East Asian Linguistics, 16, 61–86.

Hellmuth, S. (2013). Phonology. In J. Owens (Ed.), The Oxford Handbook of Arabic Linguistics. The Oxford Handbook of Arabic Linguistics. Oxford: Oxford University Press.

Herzallah, R. (1990). Aspects of Palestinian Arabic phonology: A non-linear approach. Phonetics Laboratory, Department of Modern Languages and Linguistics, Cornell University.

Hetzron, R. (1972). Ethiopian Semitic: studies in classification: Manchester University Press.

Hsieh, F., Kenstowicz, M., & Mou, X. (Eds.). (2009). Mandarin adaptations of coda nasals in English loanwords (Vol. 307): John Benjamins Publishing company.

International Phonetic, A. (1999). Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet: Cambridge University Press.

Jacobs, H., & Gussenhoven, C. (2000). Loan phonology: perception, salience, the lexicon and OT. Optimality Theory: Phonology, syntax, and acquisition, 193-209.

Jacobs, H., & Gussenhoven, C. (2000). Loan phonology: perception, salience, the lexicon and OT. Optimality Theory: Phonology, syntax, and acquisition, 193-209.

Johanson, L., & Bulut, C. (2006). Turkic-Iranian Contact Areas: Historical and Linguistic Aspects (Vol. 62): Otto Harrassowitz Verlag.

Johanson, L., & Bulut, C. (2006). Turkic-Iranian Contact Areas: Historical and Linguistic Aspects (Vol. 62): Otto Harrassowitz Verlag.

Kaneko, E. (2006). Vowel selection in Japanese loanwords from English. Proceedings LSO Working Papers in Linguistics, 49-62.

Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. Phonology, 20(2), 219-274.

Kang, Y. (2011). Loanword phonology. In M. van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (Ed.), Companion to Phonology: Wiley-Blackwell.

Kaye, A. (Ed.) (1997). Phonologies of Asia and Africa. Eisenbraun: Winona Lake, Ind.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, M. (2005). The phonetics and phonology of Korean loanword adaptation. Paper presented at the Proceedings of the first European conference on Korean linguistics.

Kenstowicz, M. (2007). Salience and similarity in loanword adaptation: a case study from Fijian. Language Sciences, 29(2), 316-340.

Kenstowicz, M., & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. Lingua, 116(7), 921-949.

Khattab, G., Al-Tamimi, F., & Heselwood, B. (2006). Acoustic and Auditory differences in the/t/-/T/Opposition in Male and Female Speakers of Jordanian Arabic. Paper presented at the Boudelaa Sami (Ed.), Perspectives on Arabic Linguistics XVI: Papers from the sixteenth annual symposium on Arabic linguistics.

Kiliç, M., & Giriç, I. (2003). Türkiye Türkçesi'ndeki ünlülerin sesbilgisel özellikleri. In A. S. Özsoy, Taylan, E.E., Aksu-Koç, A., Akar, D. Nakipoğlu, M (Ed.), Studies in Turkish Linguistics. (pp. 3-18). Istanbul: Boğaziçi University Press.

Kiliç, M., & Öğüt, F. (2004). A high unrounded vowel in Turkish: is it a central or back vowel? Speech communication, 43(1), 143-154.

Kim, H. (2009). Korean adaptation of English affricates and fricatives in a feature driven model of loanword adaptation. In C. a. Wetzels (Ed.), Loan Phonology. Current Issues in Linguistic Theory (Vol. 307, pp. 155–180). Amestrdam: John Benjamins publishing Company.

Kirchner, R. (1993). Turkish vowel harmony and disharmony: An Optimality Theoretic account. Paper presented at the Rutgers Optimality Workshop I.

Kornfilt, J. (1997). Turkish. London and New York: Routledge.

Krämer, M. (1998). Correspendence Approach to Vowel Harmony and Disharmony: Sonderforschungsbereich 282.

Kuhl, P. (2000). A new view of language acquisition. Proceedings of the National Academy of Sciences, 97(22), 11850-11857.

LaCharité, D., & Paradis, C. (2000). Phonological evidence for the bilingualism of borrowers. Paper presented at the Proceedings of the 2000 Annual Conference of the Canadian Linguistic Association, Ottawa, Canada.

LaCharité, D., & Paradis, C. (2005). Category preservation and proximity versus phonetic approximation in loanword adaptation. Linguistic Inquiry, 36(2), 223-258.

Levin, A. (1998). Arabic linguistic thought and dialectology (Vol. 1): JSAI.

Lewis, G. (1999). The Turkish Language Reform: A Catastrophic Success: A Catastrophic Success: OUP Oxford.

Lewis, G. (2000). Turkish Grammar. Oxford Oxford University Press.

Martin, K. (2014). Regression models: How do you know you need a polynomial? The analysis factor. Retrieved from http://www.theanalysisfactor.com/regression-modelshow-do-you-know-you-need-a-polynomial/

Martin, K. G. (2014). Regression models: How do you know you need a polynomial? The analysis factor. Retrieved from http://www.theanalysisfactor.com/regression-modelshow-do-you-know-you-need-a-polynomial/

McCarthy, J. (1991). Semitic Gutturals and Distinctive Feature Theory. Paper presented at the UMOP

McCarthy, J. (1994). The phonetics and phonology of Semitic pharyngeals. Paper presented at the Papers in Laboratory Phonology III: Phonological structure and phonetic form.

McCarthy, J. (1997). Process-specific constraints in Optimality Theory. Linguistic Inquiry, 28, 231-251.

McCarthy, J. (2011). Perceptually grounded faithfulness in Harmonic Serialism. Linguistic Inquiry, 42(1), 171-183.

Murray, R. (2015). The Early History of Historical Phonology. In P. H. a. J. Salmons (Ed.), The Oxford Handbook of Historical Phonology. Oxford: Oxford University Press.

Norlin, K. (1987). A phonetic study of emphasis and vowels in Egyptian Arabic. Lund University.

Odden, D. (2011). The Representation of Vowel Length. In van Oostendorp, Marc, Colin J. Ewen, Elizabeth Hume and Keren Rice (Eds.), The Blackwell Companion to Phonology. Blackwell Reference.

Owens, J. (2005). Pre-diaspora Arabic: Dialects, statistics and historical reconstruction. Diachronica, 22(2), 271-308.

Owens, J. (2013). The Oxford handbook of Arabic linguistics: Oxford University Press.

Paradis, C. (1995). Derivational constraints in phonology: Evidence from loanwords and implications. Paper presented at the Proceedings of the 31st annual Meeting of the Chicago Linguistic Society.

Paradis, C., & LaCharité, D. (1997). Preservation and minimality in loanword adaptation. Journal of Linguistics, 33(02), 379-430.

Paradis, C., & LaCharité, D. (2001). Guttural deletion in loanwords. Phonology, 18(2), 255-300.

Paradis, C., & LaCharité, D. (2008). Apparent phonetic approximation: English loanwords in old Quebec French. Journal of linguistics-cambridge-, 44(1), 87.

Paradis, C., & Prunet, J. (2000). Nasal vowels as two segments: Evidence from borrowings Language, 324-357

Paradis, C., & Tremblay, A. (2009). Nondistinctive features in loanword adaptation. In C. a. Wetzels (Ed.), Loan Phonology. Current Issues in linguistics Theory (Vol. 307, pp. 211-224). Amsterdam: John Benjamins

Parkinson, D. (2009). arabiCorpus. Retrieved on February 2015 from http://arabicorpus.byu.edu/index.php

Peperkamp, S. (2004). A psycholinguistic theory of loanword adaptations. Paper presented at the Proceedings of the Annual Meeting of the Berkeley Linguistics Society.

Peperkamp, S., & Dupoux, E. (2003). Reinterpreting loanword adaptations: the role of perception. Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences.

Peperkamp, S., Vendelin, I., & Nakamura, K. (2008). On the perceptual origin of loanword adaptations: Experimental evidence from Japanese. Phonology, 25(01), 129-164.

Perry, J. (1984). at and-a: Arabic Loanwords with the Feminine Ending in Turkish. Turkish Studies Association Bulletin, Perry, J. R. (1984). at and-a: Arabic Loanwords with the Feminine Ending in Turkish'. Turkish Studies Association Bulletin,8(2), 16-25.

Prince, A., & Smolensky, P. (1993). Optimality Theory.

Prince, A., & Smolensky, P. (2004). Optimality Theory: Constraint interaction in generative grammar.

Ratcliffe, R. (Ed.) (2013). Morphology. Oxford: Oxford University Press.

Rose, S. (1996). Variable laryngeals and vowel lowering. Phonology, 13, 73-117.

Ruhlen, M. (1994). The origin of language: tracing the evolution of the mother tongue: Wiley New York.

Sarkar, D. (2008). Lattice: multivariate data visualization with R: Springer Science & Business Media.

Saydam, Y. (2008). Language use in the Ottoman Empire and its problems, 1299-1923.

Schaade, A. (1927). Der Vokalismus der arabischen Fremdwörter im osmanischen Türkisch Festschrift Meinhof. Sprachwissenschaftliche und andere Studien. (pp. 449-460). Hamburg: Friederichsen.

Sevan, N. (Ed.) (2007) "Sözlerin Soyağacı – Çağdaş Türkçe'nin Etimolojik Sözlüğü". Istanbul.

Shahin, K. (1997). Acoustics of pharyngealization vs. uvularization harmony. Amsterdam studies in the theory and historyof linguistic science series, 4, 215-238.

Shinohara, S. (2004). Emergence of universal grammar in foreign word adaptations. In R. Kager, J. Pater, & W. Zonneveld (Eds.), Constraints in Phonological Acquisition (pp. 292-320). Cambridge ; New York Cambridge University Press.

Silverman, D. (1992). Multiple scansions in loanword phonology: evidence from Cantonese. Phonology, 9(2), 298-328.

Smith, J. L. (2006). Loan phonology is not all perception: Evidence from Japanese loan doublets. Japanese/Korean Linguistics, 14, 63-74.

Sözlük, T., & Maddesi, N. (2005). Türk Dil Kurumu Yayınları. Türk Tarih Kurumu.

Steele, F. (2008). Module 5: Introduction to multilevel modeling concepts. University of Bristol: Centre for Multilevel Modelling.

Stein, H. (2006). Palatal-velar vocalism of Arabic-Persian loanwords in 16th-century Ottoman Turkish. Turkic-Iranian Contact Areas: Historical and Linguistic Aspects, 62, 143.

Steriade, D. (2001). Directional asymmetries in place assimilation. The role of speech perception in phonology, 219-250.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S., Nishi, K., & Jenkins, J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. Journal of Phonetics, 26, 311-344.

Team. (2015). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from http://www.R-project.org/

Tietze, A. (1958 and 1999). Direkte arabische Entlehnungen im anatolischen Türkisch. In M. J. D. et.al. (Ed.), (pp. 255–333). Ankara: Türk Dil Kurumu: Türk Dil Kurumu.

Tietze, A. (1992). "Überlegungen über die lautliche Form der arabischen und persischen Lehnwörter im anatolischen Türkisch. Wiener Zeitschrift für die Kunde des Morgenlandes 82, 349–358.

Tietze, A. (2002a). Tarihi ve Etimolojik Türkiye Türkcesi Lugati Cilt 2 F-J: Verlag der Österreichischen Akademie der Wissenschaften.

Tietze, A. (2002b). Tarihi ve Etimolojik Türkiye Türkçesi Lugatı. CI AE, Simurg Yayınları, Ġstanbul-Wien.

Torreblanca, M. (1994). On Hispano-Arabic historical phonology: Latin and Romance evidence. Amestrdam studies in the theory and history of linguistic science, 4, 37-37.

Turkey political map. (2016). Retrieved from http://www.bizbilla.com/country-maps/maps/turkey-political-map.jpg

Turkish Dictionary (2015). Retrieved from http://www.turkishdictionary.net/

Underhill, R. (1986). Turkish. Studies in Turkish linguistics, 7-21.

Van Der Hulst, H., & Van De Weijer, J. (1991). Topics in Turkish phonology. Turkish linguistics today, 11-59.

Venables, W. & Ripley, B. (2003). Modern Applied Statistics with S-Plus. (4th edition ed.). New York: Springer.

Vendelin, I., & Peperkamp, S. (2006). The influence of orthography on loanword adaptations. Lingua, 116(7), 996-1007.

Versteegh, K. (2001). Linguistic contacts between Arabic and other languages. Arabica, 48(4), 470-508.

Watson, J. (1999). The directionality of emphasis spread in Arabic. Linguistic Inquiry, 30(2), 289-300.

Watson, J. (2002). The phonology and morphology of Arabic: Oxford University Press on Demand.

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications.

Woidich, M. (1994). Cairo Arabic and the Egyptian dialects.

Yip, M. (1993). Cantonese loanword phonology and Optimality Theory. Journal of East Asian Linguistics, 2(3), 261-291.

Yip, M. (2002). Necessary but not sufficient: perceptual influences in loanword phonology. The Journal of the Phonetic society of Japan, 6(1), 4-21.

Yip, M. (2006). The symbiosis between perception and grammar in loanword phonology. Lingua, 116(7), 950-975.

Younes, M. (1991). Emphasis spread in three Arabic dialects. Ms., Cornell University, Ithaca, NY.

Zaborski, A. (1994). Archaic Semitic in the light of Hamito-Semitic. Zeitschrift für Althebraistik, 7(2), 234-244.

Zaborski, A. (1997). Qaatala and Qattala in Semitic and Hamitosemitic. Rocznik Orientalistyczny, 50, 257-262.

Zawaydeh, B. (1997). An acoustic analysis of uvularization spread in Ammani-Jordanian Arabic. Studies in the Linguistic Sciences, 27(1), 185-200.

Zawaydeh, B. (1999). The phonetics and phonology of gutturals in Arabic. Indiana University.

Zawaydeh, B., & de Jong, K. (2003). Uvularization spread in Arabic. Speech Prosody and Timing: Dynamic Aspects of Speech, 4, 93.

Zuur, A., Ieno, E., Walker, N. J., Saveliev, A., & Smith, G. (2009). Mixed effects modelling for nested data. In Mixed effects models and extensions in ecology with R. New York: Springer.