Umlaut in Jerusalem Domari

Moe Kitamura
The University of Tokyo, Japan

Abstract   This study conducts a phonetic and phonological analysis of the umlaut phenomenon within the gender system of Jerusalem Domari. Using descriptive research and acoustic analysis of recordings, the study establishes several key findings: A contrast is observed in the integration of pre-Arabic loanwords (Persian, Kurdish, Turkish) with Indo-Aryan native words, which follow the umlaut rules, whereas the loanwords from Arabic, the most recent contact language, do not. A clear phonemic distinction is identified between the two open vowels, [a(:)] and [ɑ(:)], in pre-Arabic words, while these vowels exhibit allophonic values in the Arabic loanwords.

Keywords   Domari. Umlaut. Descriptive linguistics. Language contact. Phoneme identification.

Summary   1 Introduction. – 2 Introduction to Domari. – 3 Umlaut in Other Languages. – 4 Umlaut Rules in Domari. – 5 Acoustic Analysis. – 6 Discussion: A New Perspective on the Vowel System of Jerusalem Domari. – 7 Conclusion.

1 Introduction

This study investigates the phenomenon of umlaut observed in the nominal or adjectival gender system in Jerusalem Domari. Conducting descriptive research and acoustic analyses of fieldwork data from one Domari speaker recorded in Jerusalem during 2019-20, as well as acoustic analyses of sample audio data from two additional speakers in Matras (2012), it seeks to establish the following arguments:

1. Jerusalem Domari exhibits umlaut on nouns or adjectives with open vowels in their final syllables. Consonant-ending nouns and adjectives, which have lost the final vowels that would have triggered umlaut, also exhibit umlaut.
2. Pre-Arabic loanwords, including Persian, Kurdish, and Turkish, integrate with Indo-Aryan native words and follow the umlaut rules. In contrast, loanwords from Arabic, the latest contact language, do not follow these rules.
3. The two open vowels, front [(ɑː)] and back [(aː)], are recognized as distinct phonemes in pre-Arabic words of Jerusalem Domari. It should be noted, however, that the distinction between these open vowels has existed independently of the umlaut phenomenon.
4. In Arabic loanwords, the open vowels [ɑ(ː)] and [a(ː)] show allophonic values with complementary distribution. This indicates the existence of Parallel System Borrowing in phonology in the bilingual setting of Domari.

Regarding the third argument on distinct phonemes in Jerusalem Domari, Matras (2012, 51) highlights the challenges in identifying phonemes, particularly around the sounds [a], [ʌ], [ɔ], and [o]. This study aims to address these challenges and clarify phoneme identification for these specific sounds.

All Domari examples in this paper, unless otherwise noted, are based on the fieldwork research which I conducted in Jerusalem in 2019-20 with a male Jerusalem Domari speaker, born and raised in Jerusalem.

2 Introduction to Domari

2.1 Domari Language

Domari is an Indo-Aryan language spoken by the Dom people, an ethnic diaspora group in the Middle East. Its relationship with Romani is often mentioned due to shared ethnic and linguistic characteristics. It is indicated in previous research (Matras 2012; Herin 2012; 2014) that there are dialectal variations, specifically Northern

Domari in Syria and Lebanon, and Southern Domari in Palestine and Jordan, although there is still a need for more fieldwork data to establish Domari dialectology. The focus of this study is on Jerusalem Domari, a variant of the Southern dialect. Domari is severely endangered, with most speakers being bilingual in Arabic. Most Domari speakers also have proficient knowledge of Arabic. Due to the intense and prolonged bilingual situation, Domari has been significantly influenced by Arabic. In particular, Jerusalem Domari is in a severely endangered situation, with almost all Dom people speaking Arabic as their first language and having little knowledge of Domari. Currently, only two fluent Domari speakers are found in Jerusalem among the elderly generation.

Through the analysis of phonological innovations, Turner (1926) demonstrated that it belongs to the Central group of the Indo-Aryan branch. It is estimated that the Dom people resided in the Midland of the Indian subcontinent from the Old Indo-Aryan to the Middle Indo-Aryan periods, subsequently relocating to the Northwest from the Middle Indo-Aryan to the New Indo-Aryan periods.

In figure 1, the estimated migration path of the Dom people is illustrated [fig. 1]. It depicts their movement from the midland of the Indian subcontinent towards the Northwest, their traversal of the Iranian-speaking area, and their eventual arrival in Arabic-speaking areas. Consequently, the term 'Pre-Arabic words' is used in this study to refer to the vocabulary in Domari that existed prior to contact with Arabic. This category includes native Indo-Aryan words, along with loanwords from Iranian languages such as Persian and Kurdish, as well as from Turkic languages.
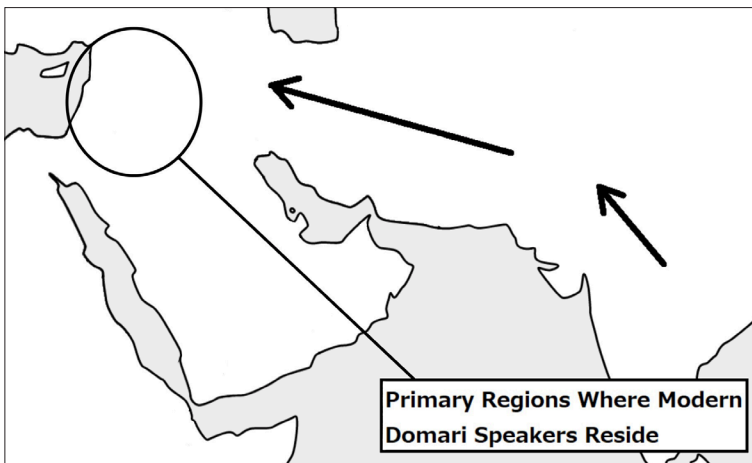

Primary Regions Where Modern Domari Speakers Reside

Figure 1   Estimated Migration Path of Domari

2.2 General Grammatical Background

The basic word order in Domari is VSO, and subject nouns are often omitted when they are clear from the context, a pattern that aligns with Jerusalem Arabic.

(1)   a.   *feː-r-a*   *zaːr-a*   *oːr-as.*
               hit-PRF-3.SG.M   boy-M.SG.NOM   that-OBL.M
               'The boy hit that man.'
        b.   *lah-am-r-i*   *kull*   *diːs.*
               see-1.SG.A-2.SG.P-PRS   every   day
               'I see you every day.'

Domari has predication markers, which construct the subject complement and agree in number and gender with the subject, as illustrated in example (2).

(2)   a.   *ame*   *giʃ*   *doːm-eːni.*
               1.PL.NOM   all   Dom-PRED.PL
               'We are all Doms.'
        b.   *ama*   *till-eːk.*
               1.SG.NOM   old-PRED.M.SG
               'I am old.'
        c.   *ato*   *ʃtɐt-ik.*
               2.SG.NOM   young-PRED.F.SG
               'You are young.'
        d.   *kariːm*   *doːm-i.*
               Kareem   Dom-PRED.SG
               'Kareem is a Dom man.'

2.3 Domari Vowel System

Figure 2 displays the vowel system based on the one presented in Matras's (2012) *A Grammar of Domari*. Additionally, it should be noted that Domari distinguishes between short and long pairs of each vowel phoneme. Figure 2 only displays short vowels; however, each vowel also has a long counterpart, which is recognized as a distinct phoneme **[fig. 2]**.
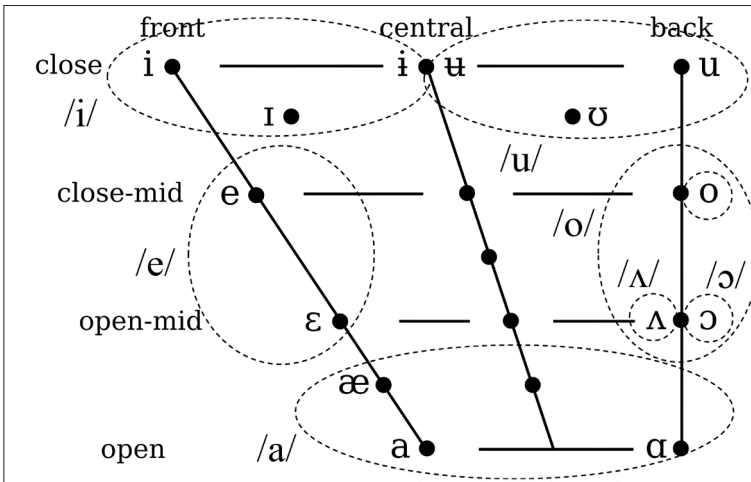
Figure 2 Vowel System based on Matras 2012, 36

In discussing the Jerusalem Domari vowel system, Matras (2012) mentioned the challenges associated with identifying the vowel sounds [a], [ʌ], [ɔ], and [o].

In Matras (2012, 51), it is stated that the contrast between [a] and [ʌ] is distinctive despite the fact that there is only one lexeme including the phoneme [ʌ], as illustrated by this near minimal pair:

[ˈpandʒi]	*pandži*	'he/she'	(Matras 2012, 51)
[ˈpʌndʒes]	*pandžes*	'five'	(Matras 2012, 51)

The phonemic contrast between /ɔ/ and /o/ is described as ambiguous, suggesting the possibility of their status as free variants. /ɔ/ predominantly precedes semi-vowels, and has minimal pairs contrasting to /a/, as demonstrated by the following examples:

[rɔˈwari]	*rɔwari*	'he/she cries'	(Matras 2012, 51)
[raˈwari]	*rawari*	'he/she travels'	(Matras 2012, 51)
[bɔˈjom]	*bɔyom*	'my father'	(Matras 2012, 51)
[baˈjom]	*bayom*	'my wife'	(Matras 2012, 51)
[dɔˈwari]	*dɔwari*	'he/she washes'	(Matras 2012, 51)
[daˈwari]	*dawari*	'he/she dances'	(Matras 2012, 51)

Additionally, it is implied that [ɔ] and [o] do not show complementary distribution, and evidence suggests that native speakers regard these two vowels as distinct phonemes. This study aims to address the challenges associated with identifying vowel phonemes in Jerusalem Domari.

2.4   Domari Nouns and Gender System

Domari shows a two-way gender system: masculine and feminine. Typically, masculine nouns in Domari are marked with the suffix -*a* (-*α* in later transcription), and feminine nouns with the suffix -*i* in the nominative singular. Furthermore, Domari has what are known as consonant-ending nouns. These nouns exhibit gender distinction but lack suffixes in the nominative singular form. Consequently, their gender cannot be identified based on suffixes alone.

Table 1 presents Domari nouns in the nominative singular form, based on Matras 2012. For this table, I have adopted the phonological system outlined by Matras 2012. The nouns in the leftmost column are marked with masculine suffixes -*a*. Those in the middle column are marked with feminine suffixes -*i*, while the nouns in the rightmost column are consonant-ending nouns, lacking gender-specific suffixes [tab. 1].

Table 1   Domari nouns in Nominative Singular based on Matras 2012

Masculine suffix	Feminine suffix	Consonant-ending
qrar-a 'Bedouin man'	*qrar-i* 'Bedouin woman'	
šōn-a 'non-Dom boy'	*šōn-i* 'non-Dom girl'	
zar-a 'Dom boy'	*lāč-i* 'girl'	*ūyar* 'market, Jerusalem' (ꜰ)
ṣnoṭ-a 'dog'	*gor-i* 'horse'	*qar* 'donkey' (ᴍ)

In Domari, adjectives follow the same declension system as nouns. While most Domari adjectives have suffixes, there is also an example of the consonant-ending adjective, *qar-* 'stupid; donkey'. This is a result of pattern replication or a calque, modelled on the Arabic *ħimaːr* 'donkey; stupid'.

According to Masica (1991, 217-23), nominal gender is common across many Indo-Aryan languages. Sanskrit originally had three genders, a system preserved in Pali and Prakrit, although there was some confusion between the masculine and neuter genders. In New Indo-Aryan languages, the most prevalent gender system is a two-gender system, resulting from the merger of the old masculine and neuter genders.

Thus, it is evident that the Domari gender system aligns with the common pattern in modern Indo-Aryan languages. In languages where the original final vowels have been lost, 'unmarked' nouns ending in consonants are often found. The gender assignment of these nouns generally depends on the vowels that have been lost. This appears to be the origin of the consonant-ending nouns in Domari.

Considering these facts, the Domari gender system is inherited from Indo-Aryan languages and appears to be shared in many New Indo-Aryan languages.

3   Umlaut in Other Languages

Umlaut is a type of sound change originally named after the phenom-
enon observed in Germanic languages. Hock (1991, 66) defines um-
laut as "the assimilation of a class of vowels to a set of [+vocalic] seg-
ments in an immediately neighbouring syllable".

This term indicates not only the diachronic sound change itself
but also frequently denotes the synchronic phenomenon resulting
from this historical linguistic development. Table 2 provides exam-
ples of synchronic umlaut phenomena in various languages [tab. 2].[1]

Table 2   Umlaut in other languages

Languages	Examples	
Modern Standard German	*Huhn* /huːn/ 'hen.SG'	*Hühn-er* /hyːnɐ/ 'hen-PL'
	Vogel /foːgl̩/ 'bird.SG'	*Vögel* /føːgl̩/ 'bird.PL'
	Hund /hʊnt/ 'dog.M'	*Hünd-in* /hʏndɪn/ 'dog-F'
Kashimiri	*šur* 'child.M'	*šuir* 'child.F'
	koṭ 'boy'	*kəṭ* 'girl'
Palula	*ǰaanu* 'masculine person'	*ǰeeni* 'feminine person'
	moomu 'mother's father'	*meemi* 'mother's mother'
	praaču 'guest'	*preeči* 'female guest'
Nsong (Bantu B80)	*-lam-* 'cook'	*-lɛmɛn* 'cook for'
Mpur (Bantu B80)	*-lám-* 'cook'	*-lɛ́ɛ́m* 'cook for'

He also described some additional observations on the conditions
of umlaut.

> Note however that crosslinguistically, umlaut most frequently is
> conditioned by final syllables. The reason seems to be that word-fi-
> nal position is a highly conducive environment for the loss of seg-
> ments and syllables, including vowels, the most common condition-
> ing environments for vowel assimilations. (Hock 1991, 68)

Consequently, umlaut vowels may remain in certain morphological
contexts, even in the absence of the original morpheme that provided
the phonological context necessary for the assimilation. For instance,
in the German examples in Table 2, *Hühner* 'hen-PL' maintains the
suffix *-er* which was originally *-ir* in Old High German and triggers
the umlaut of the vowel in the word stem, while *Vögel* 'birds.PL' has
lost the suffix that triggered the umlaut but still preserves the vowel

---

1   Modern Standard German: Wiese 1996; Kashmiri: Koul 2003; Palula: Liljegren 2019;
Nsong and Mpur: Bostoen, Koni Muluwa 2014.

quality conditioned by the umlaut. Regarding the High German Umlaut, various scholars have studied the process whereby the sounds /u(ː)/ or /o(ː)/ change to /y(ː)/ (<ü>) or /ø:/ (<ö>), conditioned by /i(ː)/ or /j/ in subsequent syllables. This change led to these vowels obtaining distinct phonemic statuses, evolving from Old High German to Middle High German (Janda 2003).

In Indo-Aryan languages, umlaut phenomena in the gender system, similar to the umlaut in Domari that will be reported in this study, have been observed in the so-called Dardic languages of the Hindu-Kush area, located in the northwest region of the Indian subcontinent.

In Kashmiri, five types of vowel alternation are observed: 1. Lowering of /ə/, /ə̄/, and /ū/ of monosyllabic stems to /a/, /ā/, and /ō/ respectively before a plural suffix *-i* or *-ɨ*. 2. Raising of /a/ and /ā/ in CVC stems to /ə/ and /ə̄/ respectively before a suffix with *-i*. 3. Centralization of /u/, /ū/, /o/, and /ō/ to /ɨ/, /ɨ̄/, /ə/, and /ə̄/ respectively before a suffix with *-i* or *-y*. 4. Centralization of the second vowel /u/ of disyllabic words with the structure CVCVC to the central vowel /a/ before a plural suffix *-ø* (Koul 2003, 904). The second and third alternations are assimilation processes, which match the definition of umlaut. These umlaut phenomena are widely observed in Kashmiri, including in the gender system, similar to the Domari case discussed in section 4. In the gender formation process from masculine to feminine in Kashmiri, the following two umlaut processes regarding vowels are observed:

a. /u, ū, o, ō/ in masculine nouns with the structure CVC are diphthongized or replaced by the central vowels at the same height: masculine *šur* 'child', *gūr* 'milkman', *gob* 'heavy', *koṭ* 'boy': feminine *šuɨr* 'child', *gūər* 'milkwoman', *goəb* 'heavy', *kəṭ* 'girl'.

b. Penultimate /u/ of masculine nouns with the structure CVCVC is replaced by /ɨ/: *kōtur* 'pigeon', *kɔkur* 'cock': feminine *kōətɨr*, *kɔkɨr* 'hen' (Koul 2003, 905-6).

The condition which causes the first case of umlaut is similar to the case of umlaut in the Domari gender system, as described in section 4.

Similar umlaut phenomena in the gender system have been reported in some other Dardic languages. For instance, as described by Liljegren (2019), in Palula, one of the Dardic languages in the Hindu-Kush area, 'masculine person' is expressed as *ǰaanu*, while 'feminine person' is *ǰeeni* (Liljegren 2019, 302).

As noted in section 2.1, Domari belongs to the Central group of Indo-Aryan languages, so the similarity regarding umlaut between Domari and Dardic languages does not seem to be due to the preservation of an inherited genetic feature. Although language contact between Domari and Dardic languages was suggested in Turner's

(1926) analysis of phonological changes, there is no evidence so far to suggest that the similarity of umlaut in the gender systems of Domari and Dardic languages, such as Kashmiri, is due to contact-induced change.

In addition, Bostoen and Koni Muluwa (2014) reported umlaut phenomena in some Bantu B80 languages. They identify four types of umlaut:

> (a) the raising of the low central vowel a to mid front vowel (y)ʒ or æ; (b) the fronting of the open-mid back vowel ɔ to (w)ɛ or œ; (c) the fronting of the open-mid back vowel o/ʊ to wi/we or ø; (d) the fronting of the closed back vowel u to wi or ü (=[y])'. (Bostoen and Koni Muluwa 2014, 228)

Among these languages, in Yans, Mpur, Ding, and Ntsambaan, the umlaut leads to a phonemic split. In contrast, in Nsong, Mpiin, Mbuun, and Ngong, it does not lead to such a split because, in these languages, the outcome vowel of the umlaut was already present in the inherited Proto-Bantu vowels.

In both Germanic and Hindu-Kush Indo-Aryan languages, but not in some of the Bantu languages such as Nsong, Mpiin, Mbuun, and Ngong, umlaut phonemes emerged as new phonemes, distinct from certain existing phonemes. This occurred through the process of umlaut producing allophones of existing phonemes, and these allophones being phonemicized due to the loss of the conditioning environment. However, in Domari, the phonemic distinction between the open vowels /a/ and /ɑ/, which primarily interacts with umlaut rules in modern Jerusalem Domari, seems to have existed prior to the occurrence of umlaut. This is evidenced by the fact that the front vowel /a/ frequently occurs in verbal stems unrelated to the umlaut rule. Similarly, the back vowel /ɑ/ is also present in a few verbal stems.

| ex. | *ga-r-a* 'go-PRF-3.SG.M' | *bag-id-om* 'break-PRF-1.SG' |
| | *raw-r-i* 'cry-PRF-3.SG.F' | *waz-r-a* 'flee-PRF-3.SG.M |

Investigating the diachronic processes of these vowels could be a topic for future research.

4   Umlaut Rules in Domari

The analysis in this section is based on the fieldwork research which I conducted in 2019-20 with a 67-year-old male Jerusalem Domari speaker. Five rules have been identified that govern umlaut in the speaker.

1.   Fronting of Open Vowel:

The back open vowels /ɑ(ː)/ in masculine nouns become front vowels /a(ː)/ in feminine nouns.

> ex.   *qraːr-a* 'Bedouin-ᴍ'   *qraːr-i* 'Bedouin-ꜰ'
>   *aːn-a* 'egg-ᴍ'   *kaːb-i* 'door-ꜰ'

2.   Centralization of Close-Mid Vowel:

The back close-mid vowels /o(ː)/ in masculine nouns become central vowels /ɵ(ː)/ in feminine nouns.

> ex.   *ʃoːn-a* 'non-Dom.boy-ᴍ'   *ʃɵːn-i* 'non-Dom.girl-ꜰ'
>   *ʃtot-a* 'small-ᴍ'   *ʃtɵt-i* 'small-ꜰ'

3.   Application to a Consonant-Ending Feminine Noun:

Although most consonant-ending nouns are masculine (ex. *wɑt* 'stone', *sɑːl* 'rice'), only one documented example of a consonant-ending feminine noun *wijar* 'market, Jerusalem' also follows these rules. It exhibits the front vowel in the last syllable, even though it lacks the suffix that would have triggered umlaut.

4.   Application to Declined Feminine Nouns with the Palatal Glide *-j-*:

5.   The umlaut is also observed in declined feminine nouns accompanying the feminine suffix with the palatal glide *-j-* instead of *-i*.

> ex.   *ʃɵːn-j-a* 'girl-ꜰ-ᴏʙʟ.ꜰ'   *kaːb-j-a* 'door-ꜰ-ᴏʙʟ.ꜰ'

6.   Non-Application to Masculine Nouns:

These rules are not applicable to masculine nouns, even when they share the same phonological environments as feminine nouns.

> ex.   *wɑt-i* 'stone-ᴘʀᴇᴅ'   *doːm-i* 'Dom-ᴘʀᴇᴅ'

Table 3 displays Domari nouns and adjectives that have a shared word stem and variable gender, while Table 4 presents Domari nouns with invariable gender **[tabs 3-4]**.

Table 3   Domari Nouns/Adjectives with Variable Gender

Masculine	Feminine	Consonant-ending
qraːr-a 'Bedouin man'	*qraːr-i* 'Bedouin woman'	
ʃoːn-a 'non-Dom boy'	*ʃɵːn-i* 'non-Dom girl'	
	qar-i 'stupid (F)'	*qar* 'stupid (M)'
ʃtot-a 'small (M)'	*ʃtɵt-i* 'small (F)'	

Table 4   Domari Nouns with Invariable Gender

Masculine	Feminine	Consonant-ending
zaːr-a 'Dom boy'	*laːʃ-i* 'Dom girl'	*wat* 'stone' (M)
aːn-a 'egg'	*kaːb-i* 'door'	*saːl* 'rice' (M)
aːt-a 'tahini'	*ʃmaːl-i* 'chicken'	*ʃaːl* 'well' (M)
man-a 'bread'	*baːn-i* 'water'	*wijar* 'market, Jerusalem' (F)
snoːt-a 'dog'	*kɵr-i* 'house'	

Concerning rules 1-3 (Fronting of Open Vowel; Centralization of Close-Mid Vowel; Application to a Consonant-Ending Feminine Noun), it can be observed that the nouns with masculine suffixes in the left-most column in Table 3 and 4 contain back vowels [ɑ(ː)] or [o(ː)] in the word stem. In contrast, the nouns with feminine suffixes in the middle column have front or central vowels [a(ː)] or [ɵ(ː)] in their word stems. Regarding the consonant-ending nouns in the right-most column, masculine nouns typically contain back vowels, while only one feminine noun exhibits the front vowel [a].

Examples (3) and (4) demonstrate Rule 4: Application to Declined Feminine Noun with the Palatal Glide *-j-*. In example (3), the feminine noun *ʃɵːn-i*, which means 'non-Dom girl', shows the front vowel [ɵː] in the oblique case *ʃɵːn-j-a*. Similarly, in example (4), the feminine noun *kaːb-i*, meaning 'door', exhibits the front vowel [aː] in the oblique case *kaːb-j-a*.

(3)   *f-ar-i*   *ʃɵːn-j-a.*
       hit-3.SG-PRES   girl-F-OBL.F
       'He/She hits the girl.' < *ʃɵːn-i* 'non-Dom girl'

(4)   *qoːl-am-i*   *kaːb-j-a.*
       open-1.SG-PRES   door-F-OBL.F
       'I open the door'   < *kaːb-i* 'door'

This rule is typically observed in kinship terms, as Domari kinship terms generally appear with possessive pronouns. Table 5 displays Domari kinship terms accompanied by the first-person singular pronominal suffix *-oːm*, meaning 'my'. Notably, the front vowel [aː] is consistently retained in the feminine kinship terms **[tab. 5]**.

Table 5 Umlaut in Kinship Terms

Masculine	Feminine
baːj-oːm 'my father'	*daː-j-oːm* 'my mother'
baːd-oːm 'my grandfather'	*daːd-j-oːm* 'my grandmother'
maːm-oːm 'my paternal uncle'	*maːm-j-oːm* 'my paternal aunt'
xaːl-oːm 'my maternal uncle'	*xaːl-j-oːm* 'my maternal aunt'
baːr-oːm 'my brother'	*baː-j-oːm* 'my wife'

Examples (5), (6), and (7) demonstrate Rule 5: Non-Application to Masculine Nouns. In example (5), the masculine noun *wat*, meaning 'stone', is followed by the predication marker *-i*. This makes a phonological environment the same as that of nouns with the feminine suffix *-i*. However, it retains the back vowel [ɑ] in the word stem. A similar pattern is observed in example (6) with nouns containing the close-mid vowel [oː]. Although the masculine noun *doːm*, meaning 'Dom man', precedes the predication marker *-i*, it maintains the back vowel in the word stem. Likewise, in example (7), the masculine noun *baːr,* meaning 'brother', is followed by the pronominal suffix *-im* but still retains the back vowel [ɑː] in the word stem.

(5)　a.　*aha*　　　　*wat-i.*
　　　　　this　　　　stone-PRED
　　　'This is a stone.' < *wat* 'stone' (Masculine)
　　　b.　*aha　　　　*wæt-i.*

(6)　a.　*ama*　　　　　*doːm-i.*
　　　　　1.sg.nom　　　Dom-PRED
　　　　　'I am a Dom man.' < *doːm* 'Dom man'
　　　b.　*ama　　　　　*døːm-i.*

(7)　*a.　aha*　　*botr-oːs*　　*baːr-im-ki.*
　　　　　This　　son-3.SG.POSS　brother-1.SG.POSS-*a*BL
　　　　　'This is my brother's son.' < *baːr* 'brother'
　　　b.　*aha　*botr-oːs*　　*bæːr-im-ki*

Based on Rules 3-5 (Application to a Consonant-Ending Feminine Noun; Application to Declined Feminine Noun with the Palatal Glide *-j-*; Non-Application to Masculine Noun), it can be concluded that this is not solely a phonological phenomenon; rather, the umlaut has morphologised.

In addition, the application of umlaut rules is observed in pre-Arabic loanwords, including those from Persian, Kurdish, and Turkish origins, but these rules are not applied to Arabic loanwords.

ex.   *zaːr-a* 'boy'   Kurdish or Persian
       zard 'gold'   Kurdish or Persian
       kaːb-i 'door'   Turkish or Kurdish

There are the following exceptions to these umlaut rules.
1.   Feminine nouns marked with the Arabic suffixes *-iːja*

ex.   *doːm-iːja* 'Dom woman'

2.   The masculine noun *kmɑːl-i*, meaning 'policeman', which is exceptionally marked with the feminine suffix *-i*
3.   Three masculine consonant-ending nouns: *kaːn* 'ear', *ag* 'fire', *løːn* 'salt'

While the exceptional cases typically involve Arabic loanwords or pre-Arabic words preceding the Arabic-origin feminine suffix *-iːja*, there are also a few exceptions among pre-Arabic nouns with the Indo-Aryan native gender system, as illustrated in exceptions 2 and 3. Specifically, from the observations made in this research, there are three exceptions among 60 pre-Arabic nouns or adjectives that feature front vowels [a(ː)] or back vowels [ɑ(ː)] in the last syllable. Additionally, there is one exception among 22 pre-Arabic nouns or adjectives that include back vowels [o(ː)] or central vowels [ø(ː)] in the last syllable.

Regarding exception 3, the reason why these three words are exceptional is not clear. However, it might be due to the fact that they are all consonant-ending nouns, which could somehow neutralize the umlaut rules.

5   Acoustic Analysis

5.1   Acoustic Analysis Based on Fieldwork Data in 2020

This section presents the results of an acoustic analysis of the umlaut. The audio data analysed consisted of around 28 minutes of elicitation recorded in 2020, with the aim of capturing the pronunciation of nouns and adjectives. The speaker was a 67-year-old male Jerusalem Domari speaker, the same speaker with whom I conducted the descriptive research introduced in section 4. During this session, I asked him to repeat carrier sentences in Domari such as "*aha __*" or "*ihi __*" (This is __) so that the gender of the words is indicated by the demonstrative *aha* 'This.ᴍ' or *ihi* 'This.ꜰ'. For example, "*aha ɑːn-eːk*" (This.ᴍ egg-PRED.SG.ᴍ) 'This is an egg' or "*ihi baːj-oːm-i*" (This.ꜰ mother-1.SG-PRED.SG) 'This is my mother'. After I said the sentence, the speaker repeated it with accurate Domari pronunciation. I recorded

some additional words on another day in the same manner and setting, in the same room. I selected nouns and adjectives from the recorded file and manually extracted vowel areas from them, measuring the mean values of the first formant and second formant of the extracted vowel areas using Wavesufer. There were 50 tokens of 31 words, including multiple repetitions of the same words with long open vowels [ɑː] or [aː]; 31 tokens of 22 words with short open vowels [ɑ] or [a]; 17 tokens of 7 words with long close-mid vowels [oː] or [ɵː]; and 9 tokens of 5 words with short close-mid vowels [o] or [ɵ]. As evident from the number of tokens, the number of nouns or adjectives including close-mid vowels is considerably less than those with open vowels.

In the following charts, I have plotted the first formant (F1) and the second formant (F2) of all tokens of those vowels in the last syllable of nouns and adjectives. A higher F1 value indicates a more open vowel, while a higher F2 value means a more frontal position of the vowel. 'F' represents feminine words, and 'M' indicates masculine words [charts 1-2].
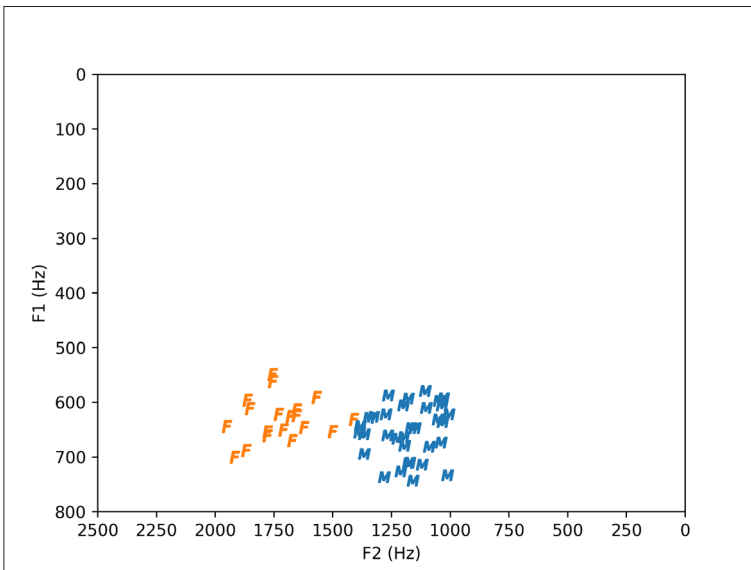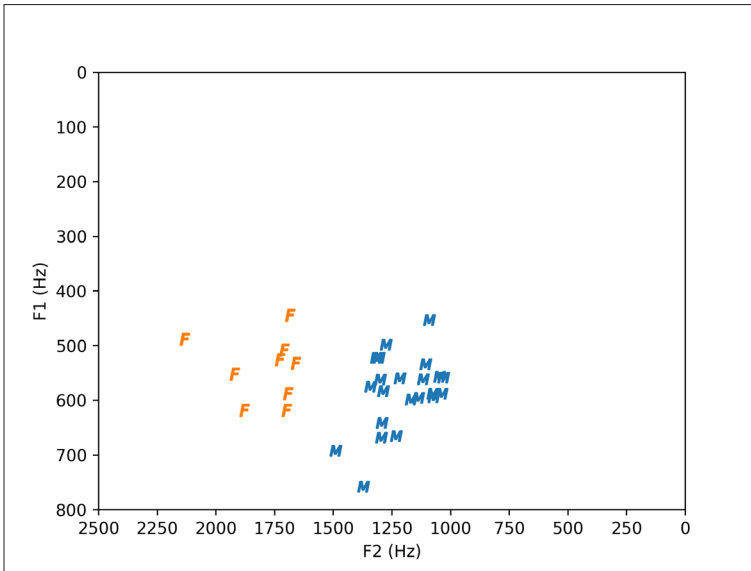

Chart 1 Umlaut in Long Open Vowels [ɑː]/[aː]

Umlaut in Short Open Vowels [ɑ]/[a]

Chart 1 illustrates the umlaut phenomenon in long open vowels. It shows that the open vowels in feminine words exhibit a higher F2 than those in masculine words, indicating that the open vowels in feminine words are more fronted.

Chart 2 displays the umlaut phenomenon in short open vowels. It shows that the open vowels in feminine words exhibit a higher F2 compared to those in masculine words, suggesting that the open vowels in feminine words are more fronted. Furthermore, the border between front and back vowels in short open vowels appears to be more distinct than in long open vowels.

Chart 3 illustrates the umlaut phenomenon in long close-mid vowels. The close-mid vowels in feminine words exhibit a higher F2 compared to those in masculine words, suggesting that the close-mid vowels in feminine words are more centralized. Additionally, the close-mid vowels in feminine words exhibit a slightly lower F1 than those in masculine words, indicating that they are more closed.

Chart 4 demonstrates the umlaut phenomenon in short close-mid vowels. Observation shows that the close-mid vowels in feminine words have a higher F2 than those in masculine words, indicating centralization in the former **[charts 3-4]**.
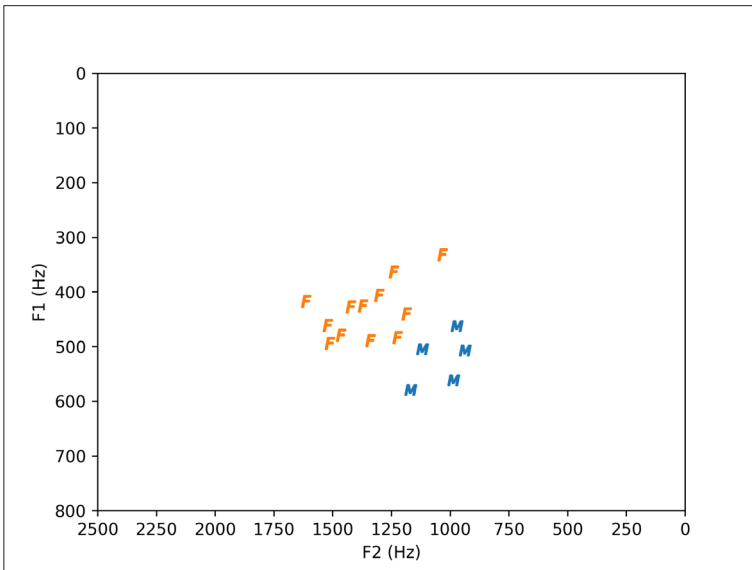
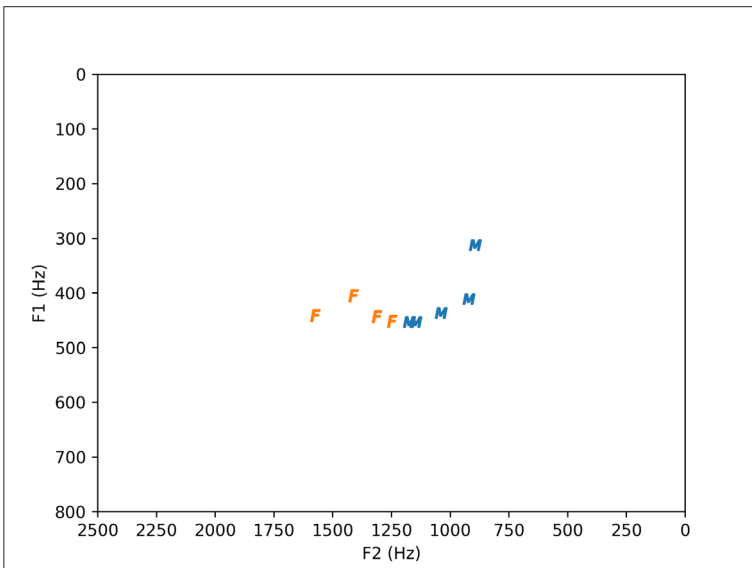**Chart 3** Umlaut in Long Close-mid Vowels [oː]/[ɵː]

Chart 4 Umlaut in Short Close-mid Vowels [o]/[ɵ]

Chart 3 and 4 have differences in the F1 values. The F1 values of the short close-mid vowels in Chart 4 do not show lower values in feminine nouns. This observation suggests that the defining aspect of the umlaut phenomenon would relate to the F2 values, which reflect the front and back positions of vowels, and variations in F1 seem to be influenced by other phonological factors. However, a more detailed analysis is needed for an extensive discussion of this topic.

5.2   Acoustic Analysis in Audio Samples from Matras (2012)

Recently, I had the opportunity to read the work of Schubert (2007), compiled under the supervision and direction of Professor Matras, which offers a detailed phonetic analysis of audio data of a female Domari speaker recorded by Professor Matras in 2000. This analysis helped inform the phonetics and phonology sections in Matras's 2012 publication, *A Grammar of Domari*. In Schubert (2007), there is a report of an umlaut phenomenon similar to the one described in this study, a detail that was prompted by Professor Matras, but not further discussed in Matras 2012.

> A phoneme /æ/ was suggested to occur in particular lexical items (*næmos, gæmos*), and as an umlaut vowel in feminines (e.g. *māmyom*), with a fronted raised quality compared to the vowel in their masculine counterparts (*māmom*). The feminine umlaut is probably better analysed as fronting from /ɑ/ to /a/. (Schubert 2007, 2)

This led me to apply a methodology similar to that used with my audio data to analyse the audio data of narratives that were transcribed in Matras (2012, 391-425), which were kindly made available to me by Professor Matras (personal communication). Due to the limited number of tokens in the sample audio, the analysis was restricted to long vowels.

There were 27 tokens of 7 words, including multiple repetitions of the same words with long open vowels [ɑː]/[aː] in the final syllable of their stems (*bɑːj-* 'father', *bɑːd-* 'grandfather', *xɑːl-* 'maternal. uncle', *daːd-j-* 'grandmother-ꜰ', *maːm-j-* 'paternal.aunt-ꜰ', *baːj-* 'mother-ꜰ', *maːs-i* 'meat-ꜰ'), from Sample 2, 'Life after retirement', a 3-minute, 46-second narrative, and 12 tokens of two words with long close-mid vowels [oː]/[ɵː] in the final syllable of their stems (*ʃoːn-ɑ* 'non-Dom.boy-ᴍ' and *ʃoːn-i* 'non-Dom.girl-ꜰ'), from Sample 3, 'A love tale', an 8-minute, 56-second narrative. The two audio files were recorded from distinct male Jerusalem Domari speakers, each of whom was in his sixties during the period from 1997 to 2000, when the fieldwork was conducted.

Applying the same methodology as in the preceding subsection, I charted the first and second formants of both the long open vowels and long close-mid vowels located in the final syllables of nouns or adjectives.
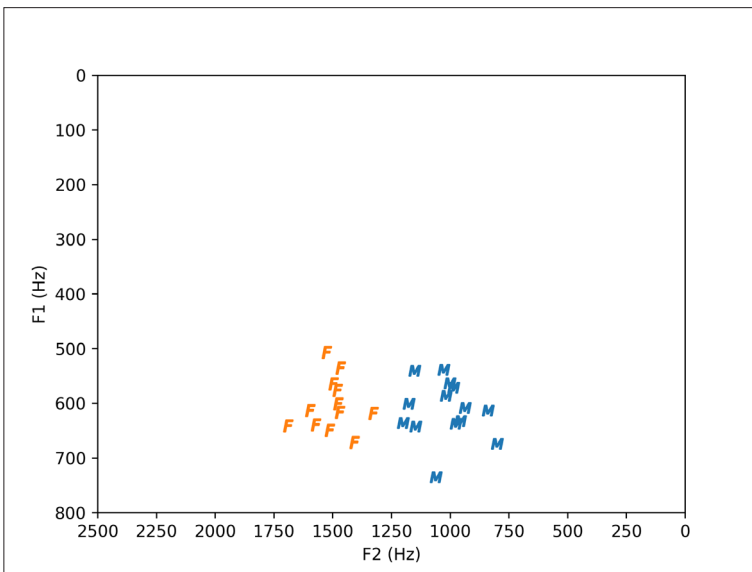
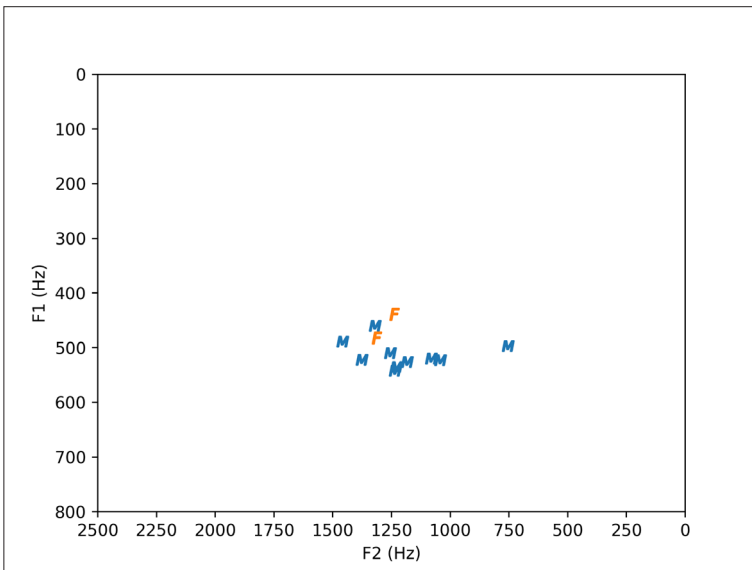**Chart 5** Umlaut in Long Open Vowels [ɑː]/[aː]

Chart 6 Umlaut in Long Close-mid Vowels [oː]/[ɵː]

Chart 5 shows that the open vowels in feminine words possess higher F2 values compared to those in masculine words, indicating a more fronted articulation of open vowels in feminine words.

However, in Chart 6, the presence of umlaut in long close-mid vowels was not observed: the F2 value in long close-mid vowels does not vary significantly with gender. However, the number of tokens is too low to draw definitive conclusions. Tentatively, I conclude that while umlaut in open vowels is a prevalent feature in Jerusalem Domari, it is not observed prevalently in the close-mid vowels [o(ː)] and [ɵ(ː)], or at least not for all speakers **[charts 5-6]**.

6 Discussion: A New Perspective on the Vowel System of Jerusalem Domari

In the preceding section's acoustic analysis, a phonetic distinction between open vowels in masculine nouns/adjectives and those in feminine nouns/adjectives was observed in both speakers. This distinction is evidenced by lower F2 values in masculine words, indicating a more back tongue position, transcribed as [ɑ(ː)]. In contrast, feminine words exhibit higher F2 values, suggesting a more front tongue position, represented as [a(ː)]. As for the close-mid vowels, one speaker displayed a pattern of lower F2 values in masculine words and higher F2 values in feminine words, while this pattern was not observed in the other speaker.

Considering these results, it seems reasonable to suggest a revised phonemic system for Jerusalem Domari.

- There is a contrast between [+BACK] and [-BACK] in open vowels, resulting in two distinct open vowel phonemes: /ɑ(ː)/ and /a(ː)/. The following examples illustrate some minimal pairs. As discussed in Section 3, these two vowel phonemes also occur independently of the umlaut context.

ex. *baːj-om* 'father-1.SG' vs *baː-j-om* 'wife-F-1.SG'
   qar-i 'stupid.M-PRD' vs. *qar-i* 'stupid-F.NOM'
   maːs-i 'month-PRD' vs. *maːs-i* 'meat-F.NOM'

This poses a typological difference between the vowel system of Romani, which has one open vowel phoneme /a/ in most of its dialects, and that of Domari, which has two open vowel phonemes: /ɑ(ː)/ and /a(ː)/.

- Variation between two speakers is observed regarding the umlaut in close-mid vowels [o(ː)] and [ɵ(ː)], suggesting a phonological contrast between the two vowels for one speaker but not the other. Given the low number of tokens for close-mid vowels, especially for the second speaker, this conclusion is tentative.

ex.   *ʃoːn-a* 'non-Dom.boy-M.NOM'

vs. *ʃeːn-i* 'non-Dom.girl-F.NOM' in one speaker

ʃoːn-a 'non-Dom.boy-M.NOM'

vs. *ʃoːn-i* 'non-Dom.girl-F.NOM' in the other speaker

This proposal addresses the challenges associated with identifying vowel phonemes in Jerusalem Domari, as explained in the previous section 2.3.

Figure 2 in § 2.3 illustrates the vowel system as presented by Matras (2012), while Figure 3 shows the newly proposed vowel system in this study **[fig. 3]**.
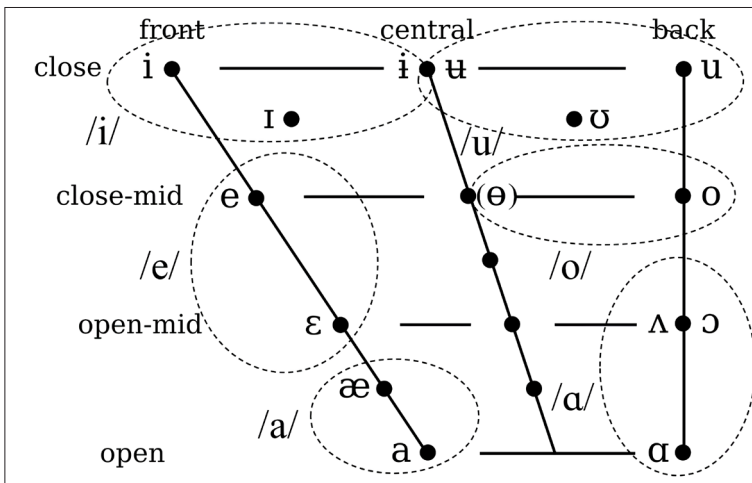

Figure 3 Proposed Vowel System

The primary differences involve the open vowel, as well as the back close-mid and open-mid vowels. As discussed in 2.3, the phonemic statuses of [ʌ] and [ɔ] have been regarded as ambiguous. I propose to consider the open-mid vowels [ʌ] and [ɔ] as allophones of the back open vowel phoneme /ɑ/. Establishing /ɑ/ as a phoneme contributes to solving the problem of the unclear phonemic status of [ʌ] and [ɔ] mentioned in Matras (2012).

Furthermore, the representation of the central close-mid vowel [ɵ] is enclosed in brackets. Considering the inter-speaker variation in the umlaut of close-mid vowels, it is challenging to decisively determine whether the central close-mid vowel [ɵ] functions as an allophone of the back close-mid vowel /o/ or is an independent phoneme.

In addition, the phenomenon of particular interest is the functioning of the open vowels [a] and [ɑ] as allophones in Arabic loanwords. Matras mentioned this point:

The realisation of /a/ as [ɑ] is in fact consistent or obligatory in the immediate environment of pharyngeals, and so one might speak of a pharyngealizing effect on the vowel, similar to that found in Arabic. (Matras 2012, 39)

ex.   [ˈzˤaːbɪt] Arabic *ẓābiṭ* 'officer' (Matras 2012, 47)
      [tˤɑwˈlɛ] Arabic *ṭawle* 'table' (Matras 2012, 44)

This effect is widely common in Arabic. Cowell (1964), in his descriptive research on "Syrian Arabic", the colloquial Arabic of what is called "Greater Syria", including Syria, Lebanon, Jordan, and Palestine, explains as follows:

Velarization is usually not limited to a single sound in a word, but commonly affects whole syllables and often whole words: *ḍạḷḷ, ṃaḅṣụ̄ṭ, ẓ̣ạ̄ḅẹṭ.* (Cowell 1964, 7)

In addition, Cowell (1964) mentions the velarized glottal stop sound [ʔˤ], showing an example of a minimal pair: *ʔáššaṛ* 'he signalled' vs. *ʔ̣áššaṛ* 'he peeled' (Cowell 1964, 7). There is a view that attributes the pharyngeal effect to the open vowel in these words rather than to the glottal stop and establishes the pharyngeal vowel phoneme /aˤ/ as a distinctive vowel from /a/, as in Obégi (1971, 25-8), the analysis of phonemic systems of Lebanese Arabic. This phenomenon is, however, observed as a marginal case limited to some areas, as Cowell (1964, 8) points out:

In a large part of the central area, including Damascus, and most of Lebanon, the distinction between ʔ and ʔ̣ is likewise obliterated, and is likewise subject to much vacillation elsewhere.

Thus, the view in which the back open vowel /aˤ/ is treated as a distinct phoneme from /a/ in Arabic seems unrelated to the Domari open vowels discussed in this paper.

Consequently, the phenomenon of open vowels in Domari is summarized as follows: These open vowels /a/ and /ɑ/ behave as distinct phonemes in pre-Arabic Domari lexicon, yet they function as allophones with complementary distribution in Arabic loanwords. This indicates that Domari possesses complex phonological layers, exhibiting dual phonological systems within its structure.

Cross-linguistically, this kind of accurate phonological borrowing without phonological adaptation, as seen in Arabic loanwords in Domari, occurs under an intense bilingual setting (Matras 2009, 342). In that situation, speakers have full knowledge of the donor language and make an effort to replicate the original phonology. In most cases, this type of phonological borrowing causes enrichment of the

phonemic system in the recipient language by adding new phonemes to its inventory, without causing conflict between the phonological environments of native words and loanwords. In the case of Domari discussed in this study, however, the phonological rule for the open vowel [ɑ] as an allophone of /a/ in Arabic loanwords, modelled on Arabic, adds another phonological layer to the existing phonemic system of pre-Arabic words, where the back open vowel /ɑ/ is a distinct phoneme from the front open vowel /a/.

In the work of Kossmann (2010, 459), the phenomenon called Parallel System Borrowing is explored. This refers to the type of borrowing which leads to "a coexistence of borrowed and native paradigms in one and the same language".

Kossmann describes two key typological or sociolinguistic factors that influence Parallel System Borrowing. The first is defined as

> contact situations with a high-prestige language, which is used for purposes related to religious and scientific learning, and which is formally taught. (480)

The second factor is described as a bilingualism setting where

> relatively small language communities in a setting with a foreign language, which is the dominant language in most communicative domains that extend outside the community. (481)

Although Kossmann's study focuses on morphology, the latter sociolinguistic setting is notably similar to that of Domari, and the results of this study imply Parallel System Borrowing in the field of phonology, particularly in the phonemic system.

7   Conclusion

In conclusion, through descriptive research and acoustic analyses of fieldwork data from one Domari speaker recorded in Jerusalem during 2019–2020, along with acoustic analyses of sample audio data from two additional speakers as presented in Matras (2012), this study can be summarized by the following four points:

1. In Jerusalem Domari, umlaut in nouns and adjectives with open vowels in their final syllables is observed. Consonant-ending nouns and adjectives, which have lost the final vowels that would have triggered umlaut, also exhibit umlaut.
2. Pre-Arabic loanwords, including those from Persian, Kurdish, and Turkish, are assimilated into the same morpho-phonological rules of umlaut as Indo-Aryan native words, whereas loanwords from Arabic, the latest contact language, do not follow these rules.
3. The two open vowels /ɑ(:)/ and /a(:)/ are recognized as distinct phonemes in pre-Arabic words. It should be noted, however, that the distinction between these open vowels has existed independently of the umlaut phenomenon.
4. In Arabic loanwords, the open vowels [ɑ(:)] and [a(:)] function as allophones of /a(:)/ with complementary distribution. This implies the possibility of Parallel System Borrowing in a phonemic system in the bilingual setting of Domari.

There remains a challenge to be addressed in understanding this phenomenon. The umlaut on the close-mid vowels [o(:)] and [ɵ(:)] exhibits apparent inter-speaker variation.

In addition, it is essential to address the historical explanation of this phenomenon: is the umlaut in question a recent innovation or an archaic retention? The umlaut phenomenon has not been documented in Northern Domari or Romani, despite the observed phonemic distinction between the back open vowels /ɑ(:)/ and /a(:)/ in the former (Herin 2012; 2014). However, there has been no research specifically concentrating on the phonetics or phonology of Northern Domari. It means that there is still a possibility of related phenomena in this dialect.

To investigate the inter-speaker variation regarding close-mid vowels [o(:)] and [ɵ(:)] in detail and to determine whether this umlaut constitutes the retention of an archaic feature or represents a recent innovation, it is essential to collect additional fieldwork data from other dialects.

Bibliography

Boeston, K.; Koni Muluwa, J. (2014). "Umlaut in the Bantu B80 Languages of the Kwi-lu (DRC)". *Transactions of the Philological Society*, 112(2), 209-30.
https://doi.org/10.1111/1467-968x.12033

Cowell, Mark W. (1964). *A Reference Grammar of Syrian Arabic.* Washington, D.C.: George Town University Press.

Herin, B. (2012). "The Domari Language of Aleppo (Syria)". *Linguistic Discovery*, 10(2), 1-52.
https://doi.org/10.1349/ps1.1537-0852.a.412

Herin, B. (2014). "The Northern Dialects of Domari". *Zeitschrift der Deutschen Morgenländischen Gesellschaft*, 164(2), 407-50.
https://www.jstor.org/stable/10.13173/zeitdeutmorggese.164.2.0407

Hock, H.H. (1991). *Principles of Historical Linguistics.* Berlin: Mouton de Gruyter.

Janda, R.D. (2003). "'Phonologization' as the Start of Dephoneticization – or, on Sound Change and Its Aftermath: Of Extension, Generalization, Lexicalization, and Morphologization". Joseph, B.D.; Janda, R.D. (eds), *The Handbook of Historical Linguistics.* Oxford: Blackwell, 401-22.

Kossmann, M. (2010). "Parallel System Borrowing: Parallel Morphological Systems Due to the Borrowing of Paradigms". *Diachronica,* 27(3), 459-87.
https://doi.org/10.1075/dia.27.3.03kos

Koul, O.N. (2003). "Kashmiri". Cardona, G.; Jain, Dh. (eds), *The Indo-Aryan Languages*. London: Routledge, 891-948.
https://doi.org/10.4324/9780203945315

Liljegren, H. (2019). "Gender Typology and Gender (In)Stability in Hindu Kush Indo-Aryan Languages". Di Garbo, F.; Olsson, B.; Wälchli B. (eds), *Grammatical Gender and Linguistic Complexity*. Vol. 1, *General Issues and Specific Studies*. Berlin: Language Science Press, 279-328.
https://doi.org/10.5281/zenodo.3462772

Matras, Y. (2009). *Language Contact.* Cambridge: Cambridge University press.

Matras, Y. (2012). *A Grammar of Domari.* Berlin: De Gruyter Mouton.

Masica, C.P. (1991). *The Indo-Aryan Languages.* Cambridge: Cambridge University Press.

Obégi, M. (1971). *The Phonemic System of a Lebanese Arabic Dialect* [Master's thesis]. Burnaby: Simon Fraser University.
https://summit.sfu.ca/item/2958

Schubert, C. (2007). *Domari Results of Phonetic Analysis* [Unpublished].
https://domari.org/wp-content/uploads/2024/08/domari_results_of_phonetic_analysis.pdf

Turner, R. (1926). "The Position of Romani in Indo-Aryan". *Journal of the Gypsy Lore Society*, 3(5), 145-89.

Wiese, R. (1996). "Phonological Versus Morphological Rules: on German Umlaut and Ablaut". *Journal of Linguistics*, 32(1), 113-35.
https://doi.org/10.1017/s0022226700000785