Predicting (the unpredictable) vowel distributions in Egyptian Arabic verbs: a lexicon study

Lily X. Xu

MA thesis

Abstract:

This thesis presents novel results from a lexicon study on vowel alternation between the perfective and imperfective forms of Egyptian Arabic verbs. I ran logistic regression models to test the effects of various phonological predictors and compare the predictability of the two forms from each other. While the models show that the two forms can be predicted with roughly equal success, there is a striking asymmetry between the two paradigmatic directions such that consonant-vowel interactions governed by phonological naturalness are very effective when predicting the imperfective vowel but play no role when predicting the perfective vowel. I discuss several analyses of the organization of this perfective-imperfective paradigm, among which a serial derivation analysis arguably aligns best with the results from the lexicon study. However, more work, especially experiments on speakers' knowledge of these statistical trends, is needed to distinguish between the possible analyses.

Table of Contents

1. Introduction

1.1 Semitic morphology: the role of the consonantal root

It has long been recognized that the lexicons of Semitic languages are centered around consonantal roots, which are sequences of two to four consonants that convey some general semantic theme, e.g. the Arabic root *k-t-b* denotes the concept of writing. However, whether consonantal roots should be treated as morphemes that contribute directly in word-formation processes has been extensively debated in the literature.

Table 1 illustrates some common properties of Semitic morphology with the Egyptian Arabic verbal system. Here, the placeholder root *f-ʕ-l* is combined with different patterns. A pattern, traditionally called *wazn* (pl. *awzaan*; also called form, measure, or binyan), is a fixed prosodic template associated with certain morphosyntactic and semantic properties.[1] There are ten major awzaan for triconsonantal roots. Which of the awzaan a root may occur in is largely idiosyncratic, resulting in many paradigm gaps.[2]

Wazn	Perfective	Imperfective	Unifying property
I	**faʕal ~ fiʕil**	**-fʕal ~ -fʕil ~ -fʕul**	**non-derived/basic**
II	faʕːal ~ faʕːil	-faʕːal ~ -faʕːil	causative/transitive
III	faːʕil	-faːʕil	associative
IV	ʔa-fʕal	-fʕil	causative (rare)
V	t-faʕːal ~ t-faʕːil	-t-faʕːal ~ -t-faʕːil	reflexive of wazn II
VI	t-faːʕil	-t-faːʕil	reciprocal of wazn III
VII	t-faʕal	-t-fiʕil	passive of wazn I
VIII	f-t-aʕal	-f-t-iʕil	intransitive (rare)
IX	fʕalː	-fʕalː	color or defect (rare)
X	sta-fʕal ~ sta-fʕil	-sta-fʕal ~ -sta-fʕil	consideration or request

Table 1. EA triconsonantal verb patterns for sound roots; listed in stem (uninflected) form; f-ʕ-l are used as placeholder consonants (Harrell et al. 1963, Abdel-Massih et al. 1979)

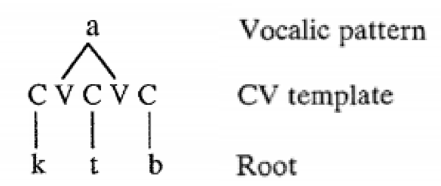
[1] The status of patterns as listed or emergent through prosodic constraints is under debate, where more recent literature tend to support an emergent approach (Ussishkin 2005, Tucker 2011, Kastner 2016).

[2] At least some gaps may be accidental, since speakers can sometimes assign a meaning to a novel combination of a root with one of the awzaan (Abdel-Massih et al. 1979), but I'm not aware of a comprehensive study on this.

Wazn I (in bold) is described as the non-derived, or basic wazn since it has the simplest morphology and has no unifying morphosyntactic or semantic properties. All other awzaan can be analyzed as deriving from wazn I, either directly or indirectly (e.g. McCarthy 1993). Nonconcatenative processes are involved in the derivation of some awzaan, e.g. wazn II by geminating the middle consonant. Wazn I is also unique in having idiosyncratic vowel alternations between the two tense/aspect forms (Abdel-Massih et al. 1979), whereas the other awzaan either have one unique vowel pattern, or in the case of wazn II, V, and X, two vowel patterns which do not alternate.

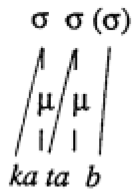
In McCarthy's (1979, 1981) analysis of Arabic verbal morphology, the consonantal root, along with the vocalic melody and the CV template, constitutes the basic components of lexical representations. Each component is realized on a distinct tier, as illustrated in (1).

(1) Tier representations of [katab] 'he wrote' (McCarthy 1981)



One benefit of this approach is that nonconcatenative processes can be straightforwardly accounted for with template modification. Another benefit is that co-occurrence restrictions on two root consonants with the same of place of articulation can be expressed on the consonant tier (McCarthy 1981, 1994). Later, McCarthy & Prince (1990) replaced the CV templates with prosodic templates, as in (2), which are composed solely of prosodic units such as syllables and moras and argued to be less stipulative than CV templates.

(2) Representation of [katab] 'he wrote' with prosodic templates (McCarthy & Prince 1990)



McCarthy and Prince (1990) derive the association of vowels and consonants through prosodic requirements, such as obligatory onset, and abstract away from the morphological decomposition of verbs and the notion of tiers. However, Tucker (2011) shows that the correct linearization can be derived from prosodic constraints given inputs that separate the consonantal root and the vocalic melody (e.g. [katab] from /ktb,a/); therefore, he argues that their status as independent morphemes should be maintained.

Two other lines of work offer support for the notion of consonantal roots as morphemes. Experimental research, including priming experiments (Frost et al. 2000 on Hebrew; Boudelaa and Marslen-Wilson 2000, 2001 on Arabic) and case studies on aphasic speakers (Prunet et al. 2000, Idrissi et al. 2008 on Arabic), have suggested that speakers have lexical storage of roots.[3] This thesis will not delve deep into this literature, but see Prunet (2006) for discussion.

Additional evidence comes from morphosyntactic works in the framework of Distributed Morphology (Halle & Marantz 1993). Arad (2005) shows that lexical-semantic division of labor in Hebrew in which the consonantal root denotes a core meaning and the vocalic melody denotes voice is best captured by treating them as different morphemes. Kastner (2016) shows that consonantal roots condition both semantic and phonological idiosyncrasy, in ways similar

[3] Berent et al. (2007) shows that Hebrew speakers' phonotactic judgements are affected by the frequencies of vowel and consonant combinations, demonstrating the necessity of whole word lexical storage. They suggest, however, that these results are compatible with additional levels of lexical representations which contain only the root.

to lexical roots in other languages.

Counter-evidence for the involvement of consonantal roots in morphological processes comes from works showing that they are inadequate in accounting for certain patterns. Bat-El (1994) and Ussishkin (1999) demonstrate output-output correspondence in the formation of Hebrew denominal verbs where both vowels and consonant clusters can be carried over from a nominal base. For example, *telegraf* 'telegraph' is the base to *tilgref* 'to telegraph', but *praklit* 'lawyer' turns into *priklet* 'to practice law', not **pirklet*, showing that preservation of a consonant cluster takes precedence over conforming to a template. They argue that retention of such information is not possible if these verbs are derived from consonantal roots.

The seemingly conflicting evidence may be reconciled by pursuing a hybrid approach in which certain types of words are formed with consonantal roots in the input, whereas others are derived from existing words. In McCarthy's (1993) analysis of the Arabic verb, most of the verb patterns are derived from wazn I through affixation and templatic modification. However, since wazn I verbs are still templatic in this account, a root-based analysis for this wazn in the style of Tucker (2011) is feasible. Arad (2005) independently argues for a distinction between root-derived and word-derived verbs in Hebrew, arguing that root-derived verbs are characterized by having more semantic and morphological idiosyncrasy, introduced by the root, whereas word-derived verbs are more restricted with respect to these properties, having to inherit them from their base. While she abstracts away from the morphophonological processes involved in these derivations, her work nonetheless demonstrates that root- or word-based derivation may not be the property of entire morphological systems. Instead, it is more appropriate to compare properties of different types of words and analyze them accordingly.

To address the debate on the role of consonantal roots in the Arabic verbal system, this thesis focuses on the vowel alternation pattern in wazn I verbs in Egyptian Arabic. Before presenting the details of this pattern, the next section discusses the morphological frameworks that are relevant for this thesis.

1.2 Morphological frameworks

Surface-based approaches to morphophonology focus on dependencies between related forms in a paradigm. The analyses of Semitic morphology based on output-output correspondence by Bat-El (1994) and Ussishkin (1999) are examples of these approaches. One debate in the literature centers around the directionality of paradigm relations. Albright (2002) proposes that language learners select one single form in a morphological paradigm as the base from which all other forms are derived. The base is selected by two criteria: 1) it should preserve the most contrasts, and 2) it should permit accurate productive generation of the greatest number of forms. A different proposal is made by Bochner (1993), who argues that morphological rules can be redundant and may encode multidirectional dependencies that can all be productive.

These surface-based approaches contrast with frameworks in which morphology is an extension of syntax, e.g. Distributed Morphology (Halle & Marantz 1993). In this view, syntactic operations underlie all word-formation processes, and dependencies between output forms are only possible when one is built from another syntactically.

This thesis will compare these two major approaches to morphology in how well they account for the vowel alternation patterns in Egyptian Arabic. One possibility is that Arabic speakers establish some purely surface-based mapping between the perfective and imperfective forms, either unidirectional or bidirectional. Alternatively, since the perfective and the imperfective have different morphosyntactic structure (see section 2.3), it is possible to test whether the phonological evidence lines up with the derivational relationship based on syntax.

1.3 Vowel alternation in Arabic wazn I verbs

As discussed in the previous section, wazn I verbs are unique in the Arabic verbal system in having idiosyncratic vowel alternation between the two tense/aspect forms, the perfective and the imperfective. Data from Egyptian Arabic are shown below.

	Perfective	Imperfective		
a/a	kasar	ji-ksar	'break'	
a/i	katab	ji-ktib	'write'	
a/u	χarag	ji/ju-χrug	'go out'	
i/a	rigiʕ	ji-rgaʕ	'return'	
i/i	libis	ji-lbis	'dress'	
i/u	sikin	ji/ju-skun	'live'	(rare)

Table 2.Vowel alternation in wazn I verbs from Egyptian Arabic (3sg masculine form)

Works on several different dialects report partial predictability of this pattern based on phonological factors.[4] One major factor that can affect vowel choice is the root consonants. A lexicon study on imperfective verbs in Hijazi Arabic by Ahyad (2019) and Ahyad & Becker (2020) found that the presence of a guttural (pharyngeal, glottal or uvular) leads to a preference for [a] and a pharyngealized coronal leads to a preference for [u]. In nonce word experiments, speakers were shown to mirror both effects. Similar gradient effects of gutturals have been found in Modern Standard Arabic (McCarthy 1991), and pharyngealized alveolars in Egyptian

[4] One instance where a semantic property (specifically stativity) allows identification of one alternation pattern has been noted in Modern Standard Arabic (McCarthy 1991), but most of the other sources of predictability are phonological.

Arabic (Abdel-Massih et al. 1979). These effects are argued to exist categorically in Palestinian Arabic (Herzallah 1990). A commonality of these studies is that they report consonant effects only in the imperfective, but not the perfective (the opposite pattern is reported in Muslim Baghdad Arabic; see Blanc 1964).

Predictive generalizations can also be made over the vowels of the two forms. Consider the Modern Standard Arabic data in Table 3.

Perf. V	Imp. V	Example	Count	Identifying property
a	u	katab/-ktub 'write'	1029	--
a	i	dˤarag/-dˤrib 'beat'	842	--
i	a	ʃarib/-ʃrab 'drink'	518	--
a	a	faʕal/-fʕal 'do'	436	contains a guttural
u	u	balud/-blud 'be stupid'	191	stative verbs

Table 3. Vowel alternation types in Modern Standard Arabic (McCarthy 1991).

The [a] to [a] and [u] to [u] classes are independently identifiable. McOmber (1995) shows that the three remaining classes are uniquely identifiable based on the imperfective vowel. Furthermore, he argues that the imperfective should be the base from which the perfective is derived, since using the perfective as base renders the [a]-perfective verbs unpredictable as to whether their imperfective should have [u] or [i]. However, there is some doubt as to whether such vowel-to-vowel correspondence are learned by speakers. In Hijazi Arabic, perfective [i] predicts imperfective [a] 90% of the time; however, this salient trend is not mirrored by speakers in nonce word experiments (Ahyad 2019).

This pattern of vowel alternation in wazn I verbs is theoretically interesting in two ways. First, the partial dependency of vowel choice on root consonants found in these verbs provides a good test case for whether lexical representations that separate vowels and consonants are motivated. Second, output-output correspondence relations may be established between the vowels in the perfective and the imperfective, which offer insight on paradigm organization.

In this thesis, I present a detailed lexicon study on this vowel alternation pattern in Egyptian Arabic, using logistic regression models to uncover statistical trends that could contribute to vowel predictability. The study explores the role of the consonantal root by comparing the effects of vowel-to-vowel correspondence and consonant-vowel effects and compares several analyses of the organization of the perfective-imperfective paradigm. Quantitative studies on vowel alternation in colloquial varieties of Arabic are rare, and this thesis is the first that has addressed both of these questions. I follow the line of work on probabilistic phonology (Zuraw 2000, Ernestus & Baayen 2003), which shows that speakers have the ability to internalize statistical patterns into their phonological grammars.

The rest of this thesis is structured as follows. Section 2 presents relevant background on Egyptian Arabic. Section 3 presents the results of lexicon study and modeling. Section 4 relates the results back to the theoretical questions and presents a preliminary analysis, and section 5 discusses future directions.

2. Language Background

Egyptian Arabic is mostly used in the literature to refer to the dialect of Cairo, which is the dominant colloquial dialect in Egypt (Versteegh 2014). As with other Arabic countries, the phenomenon of diglossia is found in Egypt (Ferguson 1959). Colloquial dialects are used in everyday life, whereas Modern Standard Arabic (MSA) is used in formal contexts. Mixed usage of the two is common in many contexts (Eid 2007). This thesis focuses on the colloquial Cairene dialect (discussed further in section 3.1.1).

2.1 Egyptian Arabic phonology

The phonemic inventory of Egyptian Arabic is shown below.

	Labial	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	 b	t tˤ d dˤ		k g	q		ʔ
Fricative	f	s sˤ z zˤ	ʃ		χ ʁ	ħ ʕ	h
Nasal	m	n					
Lateral		l					
Tap		r					
Glide	w		j				

Table 4. Egyptian Arabic consonant inventory (adapted from Broselow 1976).

i	u	iː	uː
		eː	oː
a		aː	

Table 5. Egyptian Arabic vowel inventory (adapted from Broselow 1976).[5]

The uvular /q/ is mostly restricted to words borrowed from MSA. The pharyngealized alveolars, also known as emphatics, are characterized by having a secondary constriction in the upper pharynx (Ghazeli 1977, Laufer & Baer 1988). They exert a lowering and backing effect on

[5] The long mid vowels /eː/ and /oː/ are derived historically from low vowel and glide sequences and therefore have more restricted distribution compared to the other vowel phonemes. They do not appear in the verbal forms analyzed in this thesis.

vowels within the same phonological word, known as *emphasis spreading* (Lehn 1963, Norlin 1987, Watson 2002). The tap /r/ generally patterns with pharyngealized alveolars in emphasis spreading (Younes 1994, Watson 2002).[6]

2.2 Types of wazn I verbs

Wazn I verbs can be divided into four types based on the prosodic structure of their stems: sound, defective, hollow and doubled.[7] Sound verbs (3a) have the prosodic shape CVCVC in the perfective and -CCVC in the imperfective.[8] Defective verbs (3b) have CVCV stems in the perfective and -CCV in the imperfective. Doubled verbs (3c) have CVCː stems in both forms, and hollow verbs (3d) have CVːC stems in both form.

(3) Types of wazn I verbs in Egyptian Arabic

a. Sound verbs

[katab]	'write-3sg.M-perfective'	[ji-ktib]	'write-3sg.M-imperfective'
[katab-t]	'write-2sg.M-perfective'	[ti-ktib]	'write-2sg.M-imperfective'

b. Defective verbs

[bana]	'build-3sg.M-perfective'	[ji-bni]	'build-3sg.M-imperfective'
[ban-eet]	'build-2sg.M-perfective'	[ti-bni]	'build-2sg.M-imperfective'

c. Doubled verbs

[ħabb]	'love-3sg.M-perfective'	[ji-ħibb]	'love-3sg.M-imperfective'
[ħabb-eet]	'love-2sg.M-perfective'	[ti-ħibb]	'love-2sg.M-imperfective'

d. Hollow verbs

[ʃaal]	'carry-3sg.M-perfective'	[ji-ʃiil]	'carry-3sg.M-imperfective'
[ʃ**i**l-t]	'carry-2sg.M-perfective'	[ti-ʃiil]	'carry-2sg.M-imperfective'
[ʔaal]	'say-3sg.M-perfective'	[ji-ʔuul]	'say-3sg.M-imperfective'
[ʔ**u**l-t]	'say-2sg.M-perfective'	[ti-ʔuul]	'say-2sg.M-imperfective'

[6] There are also contexts in which /r/ patterns like non-emphatics; whether a phonemic contrast of pharyngealized and non-pharyngealized /r/ exists is debated (Younes 1994).
[7] Verbs with an initial glide in the root are generally considered to be a separate type as well, since the glide is deleted in the imperfective in some dialects (MSA: wasˤal/ja-sˤil 'arrive'), but in Egyptian Arabic, these verbs pattern like sound verbs (wasˤal/ji-wsˤal 'arrive').
[8] Egyptian does not allow initial consonant clusters. Stems that begin with a consonant cluster always surface after a vowel.

While prefixes and suffixes carry person, number and gender inflections, the stem portion of the verb generally stays unaltered (see Appendix A for full inflectional paradigms). The discussion of vowel alternation and paradigmatic relationship in this thesis focuses on the stems. A unique feature of hollow verbs is that the perfective stem vowel (bolded) when followed by consonant-initial suffixes (1st and 2nd person) is a short high vowel that matches the imperfective vowel in frontness, while they have long [aa] in 3rd person perfectives. The vowel alternation patterns for these verbs are different from the sound verbs (see section 3).

2.3 Imperfective as infinitive

While the perfective and the imperfective in Arabic are often treated as two different tense/aspect forms, this section summarizes syntactic and acquisition studies which argue that the imperfective should be treated as the infinitive form and discusses how they may contribute to the investigation of the paradigmatic organization.

Benmamoun (1999) shows that in many Arabic dialects, the perfective form of the verb is always used in past tense clauses (4), whereas the imperfective form can occur in a wide range of contexts (5). These observations are illustrated below with data from Egyptian Arabic.

(4) Distribution of the perfective in Egyptian Arabic – past tense only
raʔasˤ
dance.PERF.3sg.M
'He danced.'

(5) Distribution of the imperfective in Egyptian Arabic

a. Present tense
bi-**ju-rʔusˤ**
PRES.INDICT-3sg.M-dance.IMP
'He dances/is dancing.'

b. Future tense
ha-**ju-rʔusˤ**
FUT-3sg.M-dance.IMP
'He will dance.'

c. After modals

laazim	**ju-rʔusˤ**
must	3sg.M-dance.IMP

'He must dance.'

d. After auxiliaries

kaan	**ju-rʔusˤ**
was	3sg.M-dance.IMP

'He was dancing/used to dance.'

e. In infinitive complement clauses

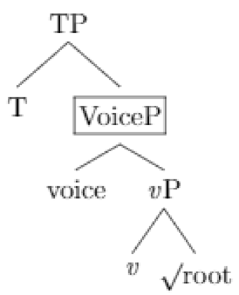
raaħ	in-naadi	ʕaʃaan	**ju-rʔusˤ**
dance.3sg.M.PERF	DEF-club	so.that	3sg.M-dance.IMP

'He went to the club so that he can dance.'

He notes that in the sentences containing an imperfective verb, tense information is always conveyed by some other particle (e.g. *ha-* for future) and the imperfective verb itself is not specified for tense; perfective verbs, on the other hand, are specified for past tense. For this reason, the so-called imperfective in Arabic may be misnamed, and infinitive or present participle are more appropriate terms for it. Acquisition studies offer additional support for this claim. Aljenaie (2010) found that Kuwaiti children in the age range of 1;8-3;1 use the bare imperfective stem as a non-finite form (see also Omar 1973).

In a morphological framework like Distributed Morphology (Halle & Marantz 1993), in which word-formation is syntactic, it follows from the above studies that the Arabic imperfective is used as the input to derive the perfective, which has more complex morphosyntactic structure, including the tense head and its projection. A sketch of Arabic clausal structure, adapted from Tucker (2011), is shown in (6).

(6)



The functional head *v* combines with a root to form a verb, which is then selected by a voice head. Since the imperfective verb always combines with other elements that carry tense information, its structure should not include T and only include VoiceP (boxed).[9] On the other hand, the perfective verb is always inflected for past tense, which means that its structure should also include a past T head.

[9] The structure of the imperfective should also include aspectual and agreement projections, which are not discussed here, but the crucial difference compared to the perfective is the inclusion of T.

3. Lexicon study and modeling

In this section, I report the results from a lexicon study based on 533 wazn I verbs in Egyptian Arabic and subsequent modeling using logistic regression models to explore the predictability of perfective and imperfective vowels. The goal is to uncover statistical patterns in the lexicon, specifically correspondence between the vowels in the two forms and consonant-vowel interactions, and test their strength. I will first present detailed data and modeling results for sound verbs, which form the predominant type and have the most complicated vowel alternation patterns, then briefly present the general patterns for all wazn I verbs.

The logistic regression models are set up so that they can be easily converted to a probabilistic phonological grammar (Goldwater & Johnson 2003, Hayes & Wilson 2008). They use a variety of consonant, vowel and verb type predictors to predict the vowels for each form. The coefficients that the models assign to each predictor are analogous to constraint weights. Additionally, the models can be used to directly compare the two possible paradigmatic directions under a surface-driven approach. The models that predict the perfective vowels represent the imperfective-to-perfective direction, since they include the identity of the imperfective vowel as one predictor and thus mimic a speaker who uses the imperfective as the base to derive the perfective. Similarly, the models that predict the imperfective vowels represent the perfective-to-imperfective direction. The models were fitted using the *nnet* package in R (Venables & Ripley 2002, R Core Team 2020).

3.1 Methodology

3.1.1 Data collection

A corpus of 533 wazn I verbs of Egyptian Arabic (Cairene) was compiled. An initial word list was made by extracting all wazn I verbs from the online dictionary Lisaan Masri (Green 2007) and later checked with a female native speaker, Fatema Shokr. She was born in Alexandria and is fluent in both Alexandrian and Cairene dialects. The data collection was carried in several sessions virtually. The speaker saw randomized lists of words in the Arabic script, which does not mark vowels, and was asked to pronounce them in both perfective and imperfective, provided that she recognized the verb. The author transcribed the vowels in both forms for each verb. The speaker pointed out the words that have divergent vowel choice in the two dialects. There were no dialectal differences in imperfective vowels, but the perfective vowel can only be [a] for Alexandrian but can be [a] or [i] for Cairene.

A small number of verbs in the dictionary was not included in the final corpus, for one of two reasons. One is that the speaker judged them to be very infrequent. The other is that the verb shows very clear features of MSA - typically if it has [a] as the imperfective prefix vowel instead of [i] or [u]. A verb was kept if the speaker felt it is a MSA borrowing but is commonly used by Egyptians and has the Egyptian prefix vowels.

The dictionary and the speaker had a fair amount of disagreement in vowel choice, which is assessed using Cohen's κ, as in Table 6. The disagreement is found overwhelmingly in the perfective of sound verbs ($\kappa = 0.37$), while the speaker and the dictionary had nearly perfect agreement in the imperfective of sound verbs ($\kappa = 0.91$) and were nearly identical on other verb types. The vast majority of disagreement involves cases in which the speaker has [i]-perfective for a verb whereas the dictionary has [a]-perfective. The reason for the discrepancies between the dictionary and the speaker is unclear. One possibility is that it reflects language change in progress. The analysis in this paper focuses on the speaker's judgements.[10]

Perfective: κ = 0.37

		Dictionary		
		a	i	Total
Speaker	a	163	5	168
	i	**97**	65	162
	Total	260	70	330

Imperfective: κ = 0.91

		Dictionary			
		a	i	u	Total
Speaker	a	154	6	0	160
	i	3	90	3	96
	u	4	2	68	74
	Total	161	98	71	330

Table 6. Speaker judgement vs. dictionary (Green 2007) on sound verb vowels

3.1.2 The corpus

In the corpus of 533 verbs, sound verbs take up more than half (330; 62%). Additionally, the corpus includes 59 defective verbs (11%), 61 hollow verbs (11%), and 83 doubled verbs (16%).

Vowel alternation patterns in Egyptian Arabic are different for each verb type. Table 7 shows these patterns with frequency information from the corpus. For sound verbs, the perfective vowel can be [a] or [i], while the imperfective vowel can be [a], [i] or [u]. A verb can have any combination of vowels in the two forms, but perfective-[i]/imperfective-[u] is very rare. Defective verbs can have [a] or [i] in both the perfective and the imperfective, and any combination is possible. For both hollow verbs and doubled verbs, [a] is the only possible perfective vowel. The imperfective vowel of hollow verbs may be [i], [u], or very rarely [a]. The imperfective vowel of doubled verbs may be [i] or [u].

Verb type	Template	Perf. V/Imp. V	Count	Examples
Sound	CVCVC/-CCVC	a/a	71	kasar/-ksar 'break'
		a/i	30	katab/-ktib 'write'
		a/u	67	χarag/-χrug 'go out'

[10] Though the corpus was based on the judgements of one speaker, another female speaker from Cairo was surveyed with an arbitrarily-selected list of 40 sound verbs. Her judgements for perfective vowels on this subset of words agreed substantially with the main speaker (κ = 0.70), and much less with the dictionary (κ = 0.41). For imperfective vowels, both speakers and the dictionary were identical.

		i/a	89	rigiʕ/-rgaʕ 'return'
		i/i	66	libis/-lbis 'dress'
		i/u	7	sikin/-skun 'live'
		Total:	*330*	
Defective	CVCV/-CCV	a/a	10	baʔa/-bʔa 'become'
		a/i	26	bana/-bni 'build'
		i/a	11	sˤiħi/-sˤħa 'wake up'
		i/i	12	miʃi/-mʃi 'walk'
		Total:	*59*	
Hollow	CV:C/-CV:C	a/a	4	naam/-naam 'sleep'
		a/i	27	ʃaal/-ʃiil 'carry'
		a.u	30	ʔaal/-ʔuul 'say'
		Total:	*61*	
Doubled	CVC:/-CVC:	a/i	46	ħabb/-ħibb 'love'
		a/u	37	marr/-murr 'pass'
		Total:	*83*	

Table 7. Types of non-derived verbs and vowel alternation patterns in Egyptian Arabic, with counts based on speaker judgement.

3.1.3 Modeling

The logistic regression models presented in this thesis use three main classes of predictors: vowel, consonant, and verb type. Table 8 lists all vowel predictors. Since there are only two perfective vowels, one predictor with binary values was used. Two binary predictors, namely impV_i and impV_u, were used for imperfective vowels, which have three choices. Words with imperfective [a] receive the value of "0" on both predictors and therefore serve as the baseline.

Predictors	Value	Note
PerfV_a	Words which have [a] as the perfective vowel get the value "1", otherwise "0"	For models predicting imperfective vowels only
ImpV_i	Words which have [i] as the imperfective vowel get the value "1", otherwise "0"	For models predicting perfective vowels only
ImpV_u	Words which have [u] as the imperfective vowel get the value "1", otherwise "0"	For models predicting perfective vowels only

Table 8. Vowel predictors in logistic regression models.

The second class is consonant predictors, which assess whether a consonant belonging to a particular natural class is present in the word. I employed 8 natural class predictors based on place of articulation, listed in Table 9.[11]

labial	plain alveolar	emphatic alveolar	palatal	velar	uvular	pharyngeal	glottal
b, f, m	t,d,s,z,n,l	tˤ,dˤ,sˤ,zˤ,r	ʃ	k,g	q,χ,ʁ	ħ,ʕ	ʔ,h

Table 9. Consonant natural classes by place of articulation.

The glides {j,w} are not included since their distributions are predictable based on verb type. Place of articulation was chosen to be the main consonant property investigated because the consonant effects on vowel alternations in other dialects generally involve place features (section 2.3). The tap /r/ patterns phonologically more like emphatic alveolars instead of the plain alveolars, and AIC comparison also show that the models where /r/ is listed as an emphatic perform better (Appendix B).These predictors are all binary, where a word receives the value of "1" just in case they contain that consonant/class of consonant anywhere in the word.

In the imperfective of sound verbs with the pattern -$C_1C_2VC_3$, the second and the third root consonants are directly adjacent to the vowel, while the first root consonant is not, so one might expect stronger effects for the second and third consonants in this case. To test this, I ran the models with positional consonant predictors, which specifies whether a given natural class is present in each of the three positions in the consonantal root. However, because of the relatively small sample size, the models with positional predictors overfitted and thus were not informative. I will briefly return to this issue of consonant position in section 3.2.2.

Lastly, verb type predictors were used in the models run on all wazn I verbs. Since possible vowel choice differ by verb type, distinguishable based on prosodic shape, one would expect

[11] Models using natural class consonant predictors were compared with models using individual consonant predictors (e.g. [b]). The results favor models using natural class predictors (Appendix C). For breakdown of vowel distribution by individual consonants, see Appendix B.

speakers to use them to predict vowels.

Predictors	Value
Defective	Defective verbs (marked by the prosodic shape CVCV/-CCV) get the value "1", otherwise "0"
Hollow	Hollow verbs (marked by the prosodic shape CVVC/-CVVC) get the value "1", otherwise "0"
Doubled	Doubled verbs (marked by the prosodic shape CVCC/-CVCC) get the value "1", otherwise "0"

Table 10. Verb type predictors in logistic regression models.

Models with different combinations of these three types of predictors are compared using ANOVA tests when they are nested and using AIC (Akaike Information Criterion) when they are not. The results show that that vowel, consonant and verb type predictors all improve model fit (see Appendix C for a full list of models).

3.2 Sound verbs

This section focuses on the models' ability to predict vowel alternation in sound verbs. The distribution of vowels in sound verbs is shown below.

		Imp. V			
		a	i	u	Total
Perf. V	a	71 (44%)	30 (31%)	67 (91%)	168 (51%)
	i	89 (56%)	66 (69%)	7 (9%)	162 (49%)
	Total	160 (48%)	96 (29%)	74 (22%)	

Table 11. Perfective and imperfective vowel frequencies in sound verbs.

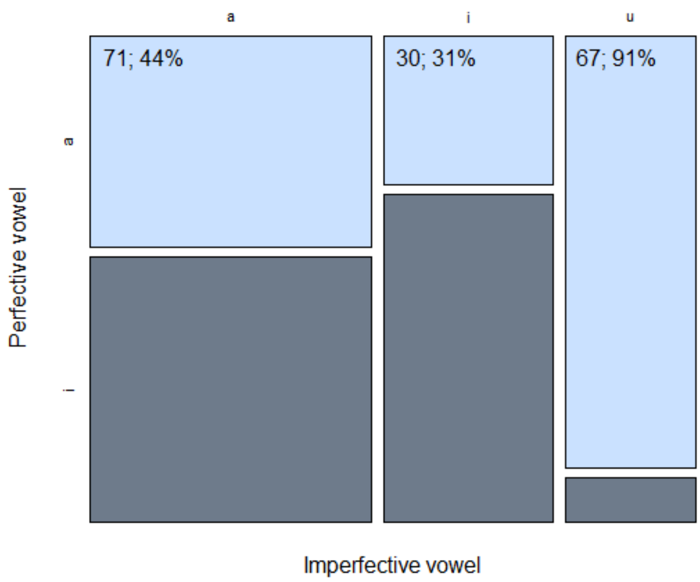


Figure 1. Breakdown of perfective vowel by imperfective vowel for sound verbs.

The distribution of [a]- and [i]-perfectives is even. For the imperfective, [a] is the most common (48%), followed by [i] (29%), and [u] (22%). The breakdown of perfective vowel choice for each imperfective vowel also differs. For [a]-imperfectives, the number of [a] vs. [i] in the perfective is close with a tiny preference for [i] (56%); [i]-imperfectives have a preference for [i] in the perfective (69%); [u]-imperfectives strongly prefer [a] in the perfective (91%).

3.2.1 Predicting the perfective vowel in sound verbs

Effects of the place of articulation of root consonants on the perfective vowel are shown below.

	a	i
labial	90 (51%)	88 (49%)
alveolar_plain	128 (45%)	155 (55%)
alveolar_emphatic	107 (60%)	71 (40%)
palatal	23 (66%)	12 (34%)
velar	24 (44%)	31 (56%)
uvular	34 (68%)	16 (32%)
pharyngeal	59 (47%)	66 (53%)
glottal	39 (48%)	42 (52%)
glide	0 (0%)	5 (100%)

Total	(51%)	(49%)

Table 12. Effects of consonant natural classes on perfective vowel distribution in sound verbs.

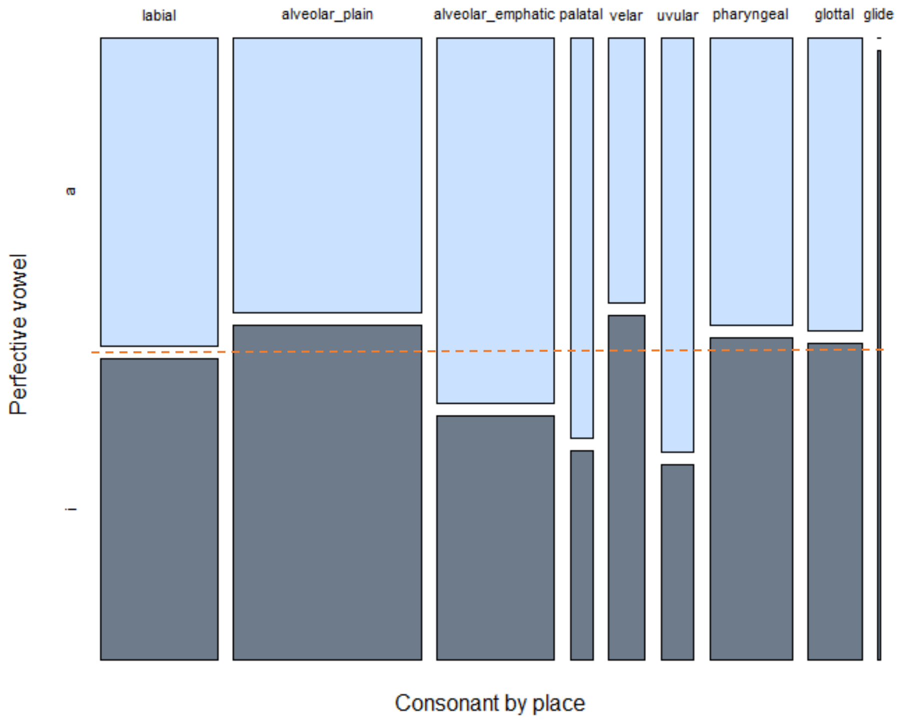


Figure 2. Effects of consonant natural classes on perfective vowel distribution in sound verbs.[12]

The presence of a certain class of consonant is recorded for each of the three root positions, so each verb is represented three times. The perfective vowel distribution for most consonant natural classes is close to the 51%/49% overall distribution (dotted line). Notable exceptions are emphatic alveolars, palatals and uvulars, which show moderate preferences for [a]. While the effects of emphatic alveolars and uvulars may have phonologically natural explanation – they have lowering effects on vowels, the same cannot be said for the effects of palatals.

The model in Table 13 predicts perfective vowels in sound verbs from imperfective vowel and consonant natural class predictors. Positive coefficients indicate a preference for [i]-perfective, whereas negative coefficients indicate a preference for [a]-perfective.

[12] There are only five sound verbs with a glide, which all have perfective [i].

Predictors	Coefficients	Std.Err.	z	p	
ImpV_i	0.5962	0.2945	2.0242	0.0429	*
velar	0.4674	0.3446	1.3567	0.1749	
alveolar_plain	0.2552	0.2337	1.0923	0.2747	
alveolar_emph	0.1837	0.2360	0.7785	0.4363	
glottal	0.1391	0.2916	0.4772	0.6332	
labial	-0.0234	0.2344	-0.0999	0.9204	
pharyngeal	-0.1282	0.2421	-0.5294	0.5965	
palatal	-0.3234	0.4172	-0.7750	0.4383	
uvular	-0.6530	0.3667	-1.7807	0.0750	.
ImpV_u	-2.4678	0.4452	-5.5425	0.0000	***
Residual Deviance: 378.3495	AIC: 398.3495				
Pseudo R^2 measures:					
McFadden: 0.173	CoxSnell: 0.213	Nagelkerke: 0.284			
Cross Validation:					
Accuracy: 0.639					

Table 13. Imperfective-to-perfective model for sound verbs.

Consistent with the lack of obvious consonant influence in the lexical data above, the consonant predictors did not come out significant and all had relatively small effect sizes. The uvular predictor had the highest effect size among them. The only predictors that came out significant in this model are the imperfective vowel predictors: having [u] in the imperfective strongly predicts [a]-perfective, while having [i] in the imperfective leads to a moderate preference for [i]-perfective. These are also consistent with lexical statistics.

3.2.2 Predicting the imperfective vowel in sound verbs

Effects of root consonants on imperfective vowel choice in sound verbs are shown below.

	a	i	u
labial	73 (41%)	66 (37%)	39 (22%)
alveolar_plain	130 (46%)	104 (37%)	49 (17%)
alveolar_emphatic	83 (47%)	31 (17%)	64 (36%)
palatal	14 (40%)	8 (23%)	13 (37%)

velar	21 (38%)	20 (36%)	14 (25%)
uvular	21 (42%)	13 (26%)	16 (32%)
pharyngeal	87 (70%)	28 (22%)	10 (8%)
glottal	49 (60%)	15 (19%)	17 (21%)
glide	2 (40%)	3 (60%)	0 (0%)
Total	(49%)	(29%)	(22%)

Table 14. Effects of consonant natural classes on imperfective vowel distribution in sound verbs.

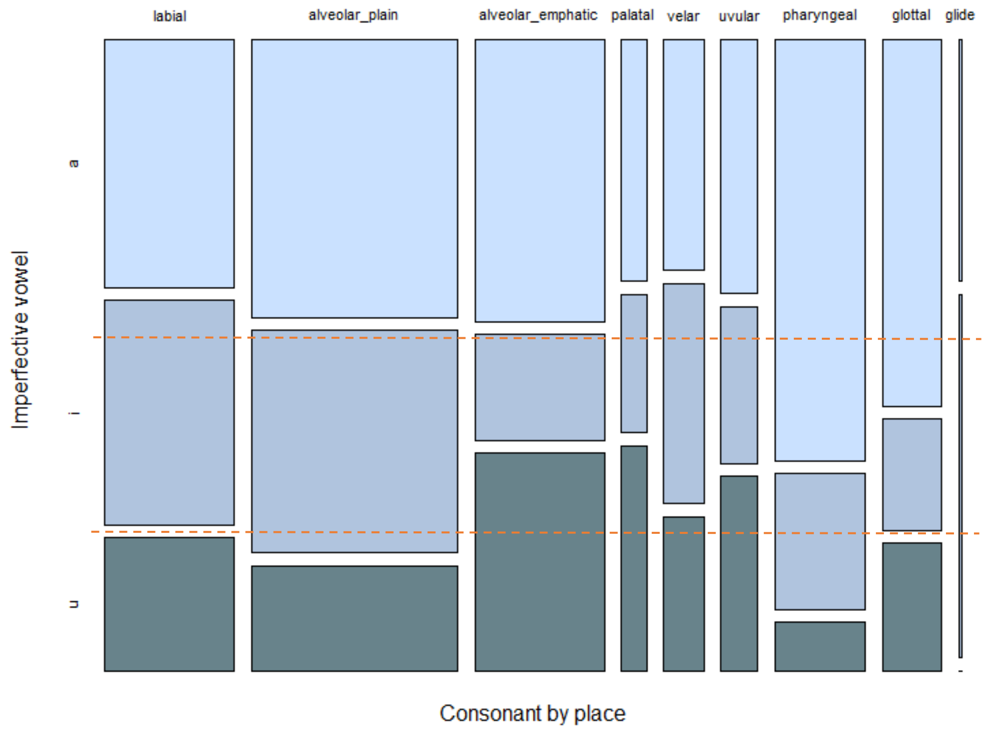


Figure 3. Effects of consonant natural classes on imperfective vowel distribution in sound verbs.

The imperfective vowel distributions for specific consonant natural classes show a great deal of divergence from the 49%/29%/22% overall distribution (dotted lines). Verbs with an emphatic alveolar have a large preference for [u]-imperfectives (36%) and dispreference for [i]-imperfectives (17%). Verbs with a palatal also have a preference for [u] (37%), whereas [a] (40%) and [i] (23%) are both moderately dispreferred compared to the overall distribution. The presence of a uvular also leads to a moderate preference for [u] (32%). Strong preferences of [a] is found in the presence of pharyngeals (70%) and glottals (60%). The presence of velars, plan alveolars, and labials seem to cause a moderate preference for [i] (36%, 37%, 37%).

The model in Table 15 predicts imperfective vowels in sound verbs from perfective vowel and consonant natural class predictors. Since there are three possible imperfective vowels, the model uses [a] as the reference and shows pairwise comparisons for [i] vs. [a] and [u] vs. [a]. Positive coefficients indicate a preference for [i]-perfective or [u]-perfective, depending on which one is compared against [a]-perfective, which is associated with negative coefficients.

	Predictors	Coefficients	Std.Err.	z	p	
i vs. a	labial	0.8542	0.2640	3.2354	0.0012	**
	alveolar_plain	0.5593	0.2637	2.1212	0.0339	*
	velar	0.3557	0.3691	0.9637	0.3352	
	palatal	-0.0136	0.5166	-0.0262	0.9791	
	uvular	-0.3641	0.4333	-0.8403	0.4008	
	PerfV_a	-0.4961	0.3085	-1.6081	0.1078	
	alveolar_emph	-0.8460	0.2835	-2.9839	0.0028	**
	glottal	-1.3146	0.3669	-3.5827	0.0003	***
	pharyngeal	-1.6010	0.3075	-5.2072	0.0000	***
u vs. a	PerfV_a	2.5937	0.4412	5.8786	0.0000	***
	velar	0.0070	0.4623	0.0152	0.9879	
	alveolar_emph	-0.1184	0.3248	-0.3647	0.7154	
	palatal	-0.2754	0.5020	-0.5486	0.5833	
	labial	-0.8113	0.3391	-2.3923	0.0167	*
	alveolar_plain	-1.0219	0.3216	-3.1776	0.0015	**
	uvular	-1.0432	0.4499	-2.3189	0.0204	*
	glottal	-1.5301	0.4251	-3.5996	0.0003	***
	pharyngeal	-2.8840	0.4361	-6.6133	0.0000	***

Residual Deviance: 515.219 AIC: 551.219

Pseudo R^2 measures:

McFadden: 0.253 CoxSnell: 0.411 Nagelkerke: 0.469

Cross Validation:

Accuracy: 0.606

Table 15. Perfective-to-imperfective model for sound verbs

This model shows that consonant predictors make a substantial contribution in predicting the imperfective vowel, consistent with the various effects by consonants shown above. Most, but not all statistical patterns highlighted above are found to be significant. Moreover, these effects have phonologically natural explanations. The natural classes pharyngeal and glottal strongly favor [a]-imperfective relative to both [i] and [u]. Uvulars strongly favor [a] over [u]. These are consistent with McCarthy's (1994) finding that pharyngeals, glottals and uvulars in Semitic languages often induce vowel lowering and favor [a] in the imperfectives of wazn I verbs in MSA. These consonants involve retraction of the tongue root, which creates an affinity for low vowels. Emphatics strongly favor [a] over [i] but are neutral between [a] and [u]. Since emphatics also involve tongue root retraction and have the effect of lowering the F2 of surrounding vowels (Norlin 1987, Laufer & Baer 1988, McCarthy 1994), it is more natural for [a] and [u] to occur in their proximity than [i]. Additionally, labials and plain alveolars favor [i] over [a], and [a] over [u]. The affinity of plain alveolars for front vowels can be seen as backness agreement, but the patterning of labials have no obvious phonological explanations. In terms of the effects of the perfective vowel, the presence of [a] in the perfective significantly increases the probability of having [u] over [a] in the imperfective. The perfective vowel does not bias the choice between [i] and [a] in the imperfective significantly.

As discussed in section 3.1.3, the differential effects of consonants by position could not be modeled due to small sample size. However, a gross inspection of vowel distribution based on the presence of consonants in specific positions show striking positional effects for pharyngeals and glottals. The table below shows that the preference for [a]-imperfective is very strong in verbs that have a pharyngeal or a glottal immediately next to the imperfective vowel (as C_2 or C_3) but is absent in verbs that have these consonants in C_1.

		a	i	u
pharyngeal	C_1	15 (33%)	21 (47%)	9 (20%)
	C_2	29 (78%)	7 (19%)	1 (3%)
	C_3	43 (100%)	0 (0%)	0 (0%)
	Total:	*87 (70%)*	*28 (22%)*	*10 (8%)*
glottal	C_1	12 (38%)	10 (31%)	10 (31%)
	C_2	18 (75%)	3 (13%)	3 (13%)
	C_3	19 (76%)	2 (8%)	4 (16%)
	Total:	*49 (60%)*	*15 (19%)*	*17 (21%)*
uvular	C_1	11 (44%)	10 (40%)	4 (16%)
	C_2	6 (46%)	3 (23%)	4 (31%)
	C_3	4 (33%)	0 (0%)	8 (67%)
	Total:	*21 (42%)*	*13 (26%)*	*16 (32%)*
emphatic alveolar	C_1	23 (53%)	8 (19%)	12 (28%)
	C_2	35 (45%)	15 (19%)	27 (35%)
	C_3	25 (43%)	8 (14%)	25 (43%)
	Total:	*83 (47%)*	*31 (17%)*	*64 (36%)*

Table 16. Effects of consonant natural classes by position in sound verb imperfectives; pharyngeals, glottals and emphatic alveolars.

Similar positional effects are found for uvulars. Even though the cases are few, there is a shift from preferring [i] with a C_1 uvular to preferring [u] with a C_2 or C_3 uvular. Interestingly, strong effects of consonant position are not found in verbs with an emphatic alveolar. The vowel distribution, which shows preference for [a] and [u] over [i], remains fairly stable regardless of the position of the emphatic alveolar.

Recall that emphatic alveolars exert lowering and backing effects on vowels in the same phonological word in Egyptian Arabic and other varieties of Arabic, which is termed emphasis spreading. Pharyngeals affect vowels in similar ways, with two differences: they are sometimes argued to have less lowering effect and only affect vowels that are strictly adjacent (Watson 2002). Interestingly, these effects are mirrored at a phonemic level in the vowel choice in the imperfective of Wazn I verbs. Pharyngeals favor [a] over [i] or [u], whereas emphatics prefer [a] or [u] over [i]. While the effects of emphatics seem to be stable regardless of position, the effects of pharyngeals are extremely strong when the vowel is adjacent and nearly absent when the vowel is not. Although the allophonic variation is generally not sensitive to morphological contexts, this phonemic-level effect only occurs in the imperfective. Ahyad & Becker (2020) reports a similar observation in Hijazi Arabic.

3.2.3 Comparison

An important finding so far is that there is an asymmetry between the two models in the role played by the consonant natural class predictors. In the model predicting the perfective vowel, the two imperfective vowel constraints had large effect sizes and were the only ones that came out significant, whereas the consonant natural class predictors had little contribution to the model. In the model predicting the imperfective vowel, while the perfective vowel predictor still played an important role in the [a] vs. [u] comparison, many of the consonant natural class predictors were significant with large effect sizes. The directions of the effects in the perfective-to-imperfective model also favor phonologically natural vowel-consonant interactions.

This asymmetry is further illustrated in Figure 4. The leftmost plot shows the counts of each type of vowel alternation for all sound verbs, with perfective vowel distribution on the left and imperfective on the right. The other two plots show the same information for verbs containing a pharyngeal and verbs containing an emphatic. For verbs that contain a pharyngeal, even though the frequency of [a] in the imperfective is much higher than that in the perfective, around half of the words with imperfective [a] has perfective [i], showing that the consonant-vowel effects only hold in the perfective. Similarly, in verbs that contain an emphatic, over half of the verbs with imperfective [a] are mapped to perfective [i].

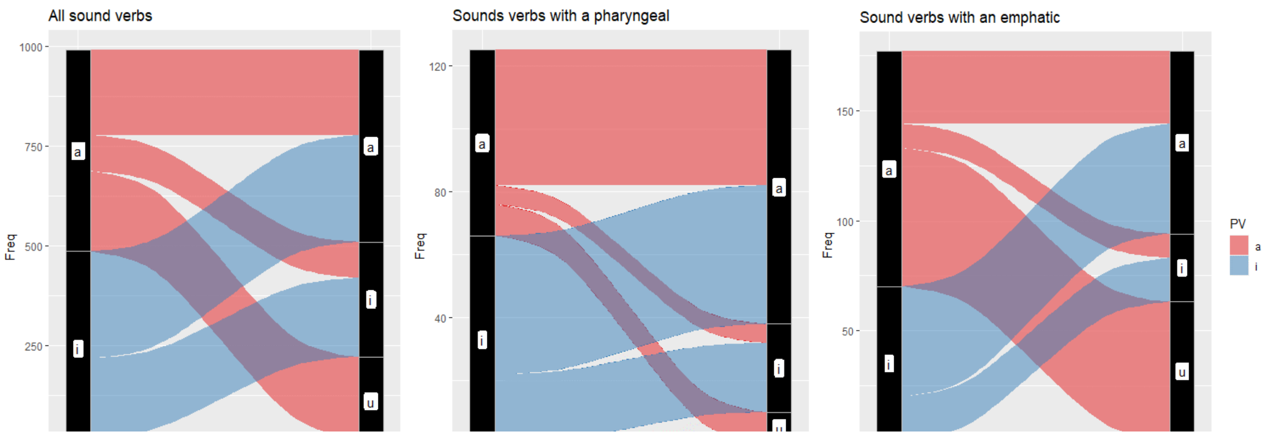


Figure 4. Vowel alternation between perfective and imperfective. All sound verbs; with a pharyngeal; with an emphatic alveolar.

The two models' goodness-of-fit is then compared with k-fold cross-validation (k=5). The dataset was randomly divided into 5 parts, and each model was run on 4 of the parts and tested on the other. This process was repeated for all 5 parts, and the average model accuracy from all the trials was calculated by comparing the model predictions on the testing data in each run with the corpus. The imperfective-to-perfective model had an higher average accuracy (0.639) than the perfective-to-imperfective model (0.606), but the difference is very small.

Furthermore, these two models can be compared to null models, which have no predictors and guess the vowel by chance. The perfective-to-imperfective model should be compared to a baseline accuracy rate of 0.33, since there are three imperfective vowel choices, whereas the imperfective-to-perfective vowel should be compared to a baseline accuracy rate of 0.5, because it is choosing between two perfective vowels. In this view, the perfective-to-imperfective model clearly has greater improvement in predictive power (0.33 to 0.606), compared to the imperfective-to-perfective model (0.5 to 0.639). This intuition is supported by

Pseudo R^2 measures. Since the two models predict different dependent variables, they cannot be directly compared with AIC or likelihood measures, and Pseudo R^2 measures are appropriate. I report three Pseudo R^2 measures, McFadden, Cox & Snell and Nagelkerke.[13] All three measures were higher for the perfective-to-imperfective model, which suggests that it is superior in terms of model fit.

To better understand the behavior of the two models, their predictions for the entire set of sound verbs were examined in closer detail. Table 17 and Table 18 list the models' accuracy for each type of vowel alternation.

Perf.V/Imp.V	Freq in lexicon	# correct prediction	Accuracy
a/a	71	16	23%
a/i	30	2	7%
a/u	67	67	100%
i/a	89	74	83%
i/i	66	66	100%

Table 17. Imperfective-to-perfective model predictions by type of vowel alternation.

Perf.V/Imp.V	Freq in lexicon	# correct prediction	Accuracy
a/**a**	71	53	75%
a/**i**	30	10	33%
a/**u**	67	50	75%
i/**a**	89	70	79%
i/**i**	66	32	48%

Table 18. Perfective-to-imperfective model predictions by type of vowel alternation

The performance of the two models for each type of vowel alternations reflects the differing role of consonant and vowel predictors. The imperfective-to-perfect model achieves

[13] These measures should be interpreted differently. While Cox & Snell and Nagelkerke both measure the improvement from null model to fitted model, McFadden also takes into consideration the proportion of variability explained by the model.

100% accuracy rate in the categories a/u (perf/imp), which can be attributed to the high weighted imperfective [u] predictor. Additionally, it is very successful for i/a and i/i, while having very low accuracy rates for a/a and a/i. Indeed, the vast majority of the errors made by this model is when the actual perfective vowel in the lexicon is [a], but the model predicts [i] (Table 19). Taken together, this shows that this model uses a rather simple strategy: the perfective vowel is [a] if the imperfective vowel is [u], otherwise [i].

Predicting the perfective:

		Model		
		a	i	Total
Lexicon	a	85	**83**	168
	i	22	140	162
	Total	107	223	330

Predicting the imperfective:

		Model			
		a	i	u	Total
Lexicon	a	123	22	15	160
	i	43	47	11	96
	u	17	7	50	74
	Total	183	71	76	330

Table 19. Model prediction vs. lexicon on sound verb vowels

On the other hand, the perfective-to-imperfective model's performance on different vowel alternation categories are not as polarized. It is reasonably successful for a/a, a/u and i/a, while having chance-level performance on a/i and i/i. This more gradient range of accuracy rates across categories is consistent with the model's reliance on consonant predictors.

3.3 All Wazn I verbs

The number of verbs of the other verb types are much smaller compared to sound verbs (83 doubled verbs, 61 hollow verbs, 59 defective verbs), so it was difficult to test the statistical patterns in these verbs. Therefore, this thesis will not discuss them in as much detail as for the sound verbs. See Appendix D for discussion on these verb types.

The distribution of vowels in all wazn I verbs in the corpus is shown below.

		Imp. V a	i	u	Total
Perf. V	a	85 (46%)	129 (62%)	134 (95%)	348 (65%)
	i	100 (54%)	78 (38%)	7 (5%)	185 (35%)
	Total	185 (35%)	207 (39%)	141 (26%)	

Table 20. Perfective and imperfective vowel frequencies in all verbs.

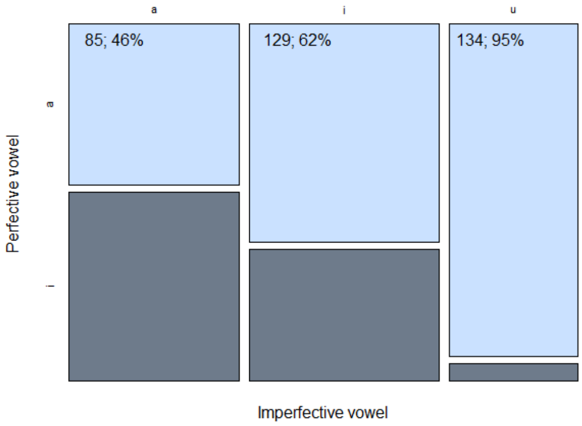


Figure 5. Breakdown of perfective vowel by imperfective vowel for all verbs.

Pooling together all verb types, 65% of the verbs have [a] in the perfective, and 35% have [i]. For the imperfective, [i] is the most common (39%), closely followed by [a] (35%), next followed by [u] (26%). Regarding the breakdown of perfective vowel choice for each imperfective vowel, the number of [a] vs. [i] in the perfective is very close for [a]-imperfectives; [i]-imperfectives have a perfective with [a] 62% of the time; [u]-imperfectives strongly prefer [a] in the perfective (95%).

Despite the difference in overall vowel distributions, the distributions as affected by consonants seem qualitatively very similar to those for the sound verbs. The place of articulation of consonants seems to have little influence of the distribution of perfective vowels but do affect the distribution of imperfective vowels. The directions of the effects found for the entire corpus also in general match those found for sound verbs only (Appendix E).

The models run on all Wazn I verbs are very similar with the models run on sound verbs with respect to the vowel and consonant predictors (Appendix F). As expected, the verb type predictors contribute significantly to these models. Cross-validation was also run for these models in order to assess their goodness-of-fit. The imperfective-to-perfective model has better average accuracy (0.732) than the perfective-to-imperfective model (0.651).

A close inspection of the models' predictions show the main reason for the lower accuracy of the perfective-to-imperfective model is that it has below chance performance for hollow verbs, with an accuracy rate of 42%. This is likely because hollow verbs do not have the phonologically natural consonant-vowel interactions found in the other types of verbs, and in some cases even have preferences in the opposite direction. For example, hollow verbs with an emphatic seem to show a weak preference for imperfective [i]. Ahyad and Becker (2020) found a similar pattern in Hijazi Arabic. They suggest that this is a case of dissimilation but did not discuss why it only occurs in hollow verbs.

There are two possible accounts for the special behavior of hollow verbs. The first, according to the traditional analysis, is that these verbs have an underlying glide as the second root consonant which matches the imperfective vowel in frontness and does surface in other related forms, such as causatives (Brame 1970). If speakers indeed have this glide in their mental representation of hollow verbs, they can easily predict the imperfective vowel. Another possibility is that speakers actually use the verbs in the 1st or 2nd person perfective forms as the base. Recall that perfective vowels for hollow verbs in these forms match the imperfective vowel in frontness, as in [ji-ʔuul] 'he says', [ʔaal] 'he said', but [ʔul-t] 'I/you said' (section 2.2). Future work should try to implement a perfective-to-imperfective model in which the perfective vowel predictors represent the vowel choice in either 1st or 2nd person forms.

4. Discussion and conclusion

4.1 Summarizing modeling results

Two key findings emerged from the models. The first is that different types of phonological factors are relevant when predicting the vowel in different forms. The imperfective vowel can be predicted based primarily on the consonant predictors, in ways governed by the phonological naturalness of consonant-vowel combinations, while a minor role is played by mapping to the perfective vowel. The perfective vowel can be predicted based solely on the perfective vowel, and consonant predictors played no role. The other key finding is that there is substantial vowel predictability in both of the paradigm directions tested. The two models had comparable overall accuracy in their predictions, though the perfective-to-imperfective model shows greater improvement from chance-level performance. This section will focus on the models on sound verbs since the models on all wazn I verbs did not differ with respect to these two points.

Since all the models attempted achieve only partial predictability, both the imperfective and the perfective forms must be memorized for most verbs. However, the asymmetry in the presence of phonologically natural consonant-vowel interactions suggests that there are different mechanisms at play in the formation of imperfectives compared to perfectives, rendering a memorization-only account unsatisfactory. This asymmetry is puzzling given that the phonological environments with respect to consonants and vowels are very similar across both the perfective (CVCVC) and imperfective (-CCVC) forms.[14] In this section, I assess various proposals on the organization of the perfective-imperfective paradigm based on their

[14]Similar observations can be made for various other dialects of Arabic, as discussed in section 1.3.

compatibility with the lexicon data and modeling results and show that the results align best with a serial derivation analysis, in which the consonantal root is used to derive the imperfective and the imperfective is used to derive the perfective.

4.2 Bidirectional vs. unidirectional analysis

The modeling results show that both the perfective and the imperfective can be predicted with roughly equal accuracy (around 60%). Notably, in both models, the correspondence between perfective [a] and imperfective [u] is salient and has significant contribution. These results suggest the possibility that speakers learn bidirectional mappings, consistent with the proposal by Bochner (1993).

However, there is a reason to consider a single-base analysis, consistent with the hypothesis by Albright (2002), namely that the two models have different strategies in their predictions. The prediction of the perfective vowel relies on a rather simple strategy, selecting [i] as the default unless the imperfective is [u] (section 3.2.3). As a result, the model prediction is largely biased, having perfect accuracy rates in some categories of vowel alternations but extremely poor performance in others. The prediction of the imperfective vowels, on the other hand, relies on phonologically natural constraints on consonant-vowel interactions. Taken together, these results suggest that the perfective is a better candidate for inflectional base compared to the imperfective. In terms of model fit (shown by Pseudo-R^2 metrics), the perfective-to-imperfective models trained on both datasets were also shown to result in more improvement from the null model.

4.3 Consonant effects and the serial derivation analysis

Neither the bidirectional nor the perfective-as-base account offers a satisfying explanation for the finding that consonant-vowel interactions governed by phonological naturalness only affect vowel choice in the imperfective. As noted above, this asymmetry cannot be explained by differences in phonological environments, since the perfective (CVCVC) and the imperfective (-CCVC) are very similar in this respect.[15]

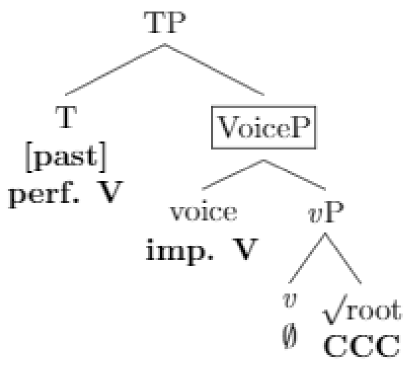
Recall that in the perfective-to-imperfective model, the perfective vowel predictor played a minor role compared to the consonant predictors, suggesting that it may be possible to predict the imperfective from the consonantal root alone. On the other hand, predicting the perfective seems only to be sensitive to the imperfective vowel. Note that, since the vowel predictors contributed significantly to the perfective-to-imperfective model, predicting the imperfective from consonantal roots alone will surely yield less accurate predictions. However, this sacrifice of predictability may be justified, since it aligns with wug test results in Ahyad (2019), which showed that Hijazi speakers actually do not utilize salient distributional information about the perfective vowel when forming the imperfective (section 1.3). Consonant-vowel interactions, such as the preference of pharyngeals for imperfective [a], on the other hand, were mirrored in wug test responses.

I will show that an analysis under Distributed Morphology (Halle & Marantz 1993) aligns with the modeling results and also accounts for the issue discussed above. Under this analysis, the consonantal root, the imperfective vowel and the perfective vowel are all treated as separate

[15] Ahyad & Becker (2020) argue for a word-based approach based on the evidence that consonants condition imperfective vowels in Hijazi Arabic, but their argument also does not account for the lack of similar effects in the perfective.

morphemes. This analysis is illustrated in (7).

(7)



In this structure, the consonantal root first combines with the functional head *v*, then it combines with imperfective vowel to form the imperfective. According to the view that the imperfective in Arabic is just the default form of the verb not specified for tense (section 2.3), this form would originate fairly low in the syntactic structure. The imperfective vowel is merged at the voice head, following Arad's (2005) proposal for Hebrew. The perfective vowel, however, merges at the T head, consistent with the observation that the perfective always conveys past tense (Benmamoun 1999).

With this structure, the absence of consonant-vowel interactions in the perfective form follows from independently proposed syntactic locality constraints, which disallow allomorphy selection between any two elements that are separated by other over material in the morphosyntactic structure (Embick 2010). Because the consonantal root merges with the imperfective vowel first, it is possible for phonological interactions between consonants and vowels to influence the imperfective form. And since the perfective vowel is structurally closer to the imperfective vowel than to the consonantal root, it follows that vowel predictors were the only ones that contributed in the imperfective-to-perfective models.

This analysis also aligns with the cross-linguistic pattern in which phonotactic restrictions seem to be stricter in smaller morphological domains, for example, in roots as opposed to morphologically complex words (Gouskova 2018). The consonant-vowel interactions can be seen as phonotactic restrictions. Analyzing the imperfective, where these consonant-vowel interactions are much stronger, as more morphologically simple than the perfective, is consistent with this fact.

Despite its success in accounting for consonant/vowel separation, this analysis is arguably inferior in terms of learnability compared to the surface-based analyses. It assumes that children are able to extract the consonantal root and vocalic melodies as separate lexical entries, whereas the surface-based analyses are free of such assumptions. Furthermore, comparing these two types of analyses lead to insight on the relationship between morphology and syntax.

4.4 Conclusion

In this section, I have outlined several possible analyses that are compatible with the modeling results. The first two analyses both involve mapping between surface forms and do not invoke abstract consonantal root morphemes. Their predictions differ, with regard to whether speakers learn predictive generalizations in one direction or two directions. The last analysis predicts that speakers first form the imperfective form by combining the imperfective vowel with the consonantal root, whereas the perfective vowel can only combine with these two at a higher level. Under this analysis, speakers are predicted to internalize the consonant-vowel interactions found in the imperfective form as well as the influence of the imperfective vowel on perfective vowel choice. More work needs to be done in order to get a more thorough understanding of paradigm organization in the minds of Egyptian Arabic speakers. Some possible future directions are discussed in the next section.

5. Future directions

The analyses discussed in this paper can crucially be distinguished with wug testing, since they make distinct predictions on which statistical trends in the lexicon speakers should learn. Table 21 lays out the predictions by various analyses with regard to all major types of statistical effects discovered by the modeling work. In the perfective-as-base analysis, since it supposes that speakers use the perfective form as the base, speakers should be able to use both the perfective vowel and root consonants when they predict the imperfective vowel but should be mostly guessing when asked to predict the perfective vowel. The bidirectional analysis predicts all three types of generalization to be learned. The serial derivation analysis, on the other hand, predicts that only the root consonants should help speakers predict the imperfective vowel, since they should not have access to the perfective vowel at this level of lexical representation. When asked to predict the perfective vowel, speakers should be able to generalize the effects of imperfective vowels.

Statistical effects in the lexicon	Perfective-as-base	Bidirectional	Serial derivation
Root consonants on imperfective vowel choice	Yes	Yes	Yes
Perfective vowel on imperfective vowel choice	Yes	Yes	No
Imperfective vowel on perfective vowel choice	No	Yes	Yes

Table 21. Predictions on wug test results by various analyses.

Two questions regarding the consonant-vowel interactions found in the imperfective are worth pursuing further. The first is whether they are restricted to the imperfective and not found in other parts of the Egyptian Arabic lexicon, or they represent general phonotactic restrictions true of the language. We have seen that the perfective form does not have these interactions. Other cases in the Arabic lexicon where vowel choice is not uniquely determined by the prosodic template can be found in nominals. For example, monosyllabic nouns of the form CVCC do not have a unique vowel pattern. However, no consonant-vowel interactions of the sort found in the imperfective verb have been discussed for these nouns. Some examples of monosyllabic nouns are shown in (8). These words have similar consonant environments, and all three vowels can be found in this environment. It would be useful to do a similar quantitative study on the nominals.

(8) Egyptian Arabic monosyllabic nouns

a. [ʕafr]	dust
b. [ʕibʔ]	burden, load
c. [ʕubtˁ]	hug

Additionally, across Arabic dialects, this asymmetry between the perfective and the imperfective, namely that consonant-vowel interactions are only significant in the imperfective, seems to be true. Previous works that have identified consonant-vowel interactions for determining the vowel in non-derived verbs generally report them in the imperfective but not in the perfective, with the exception of Iraqi Arabic (Blanc 1964). Quantitative studies that specifically test this in other dialects would be informative.

References

Abdel-Massih, Ernest T., Zaki N. Abdel-Malek & El-Said M. Badawi. 1979. *A comprehensive study of Egyptian Arabic.* (Vol. 3: A reference grammar of Egyptian Arabic). Ann-Arbor.

Ahyad, Honaida Yousuf. 2019. *Vowel distribution in the Hijazi Arabic root*. Doctoral dissertation, Stony Brook University.

Ahyad, Honaida Yousuf & Michael Becker 2020. Vowel unpredictability in Hijazi Arabic monosyllabic verbs. *Glossa: a journal of general linguistics,* 5(1).

Albright, Adam C. 2002. *The identification of bases in morphological paradigms.* Doctoral dissertation, University of California, Los Angeles.

Aljenaie, Khawla. 2010. Verbal inflection in the acquisition of Kuwaiti Arabic. *Journal of Child Language* 37(4).

Arad, Maya. 2005. *Roots and patterns: Hebrew morpho-syntax.* Springer.

Badawi, El-Said M. 1985. Educated spoken Arabic: a problem in teaching Arabic as a foreign language. In K. R. Jankowsky (ed.), *Scientific and humanistic dimensions of language*, 15-22. Georgetown University Press.

Bat-El, Outi. 1994. Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory* 12, 571-596.

Benmamoun, Elabbas. 1999. Arabic morphology: The central role of the imperfective. *Lingua* 102(2), 175-201.

Berent, Iris, Vered Vaknin, & Gary F. Marcus. 2007. Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition* 104(2), 254-286.

Blanc, Haim. 1964. *Communal dialects in Baghdad.* Harvard University Press.

Bochner, Harry. 1993. *Simplicity in generative morphology*. De Gruyter Mouton,

Boudelaa, Sami & William D. Marslen-Wilson. 2000. Non-concatenative morphemes in language processing: Evidence from Standard Arabic. *Proceedings of the workshop on spoken word access processes 1,* 23-26. Nijmegen, The Netherlands.

--- 2001. Morphological units in the Arabic lexicon. *Cognition* 81, 65-92.

Brame, Michael K. 1970. *Arabic phonology: implications for phonological theory and historical Semitic.* Doctoral dissertation, Massachusetts Institute of Technology.

Broselow, Ellen I. 1976. *The phonology of Egyptian Arabic*. Doctoral dissertation, University of Massachusetts Amherst.

Eid, Mushira. 2007. Arabic on the media: Hybridity and styles. In *Approaches to Arabic linguistics*, 403-434. Brill.

Ernestus, Mirjam & R. Harald Baayen. 2003. Predicting the unpredictable: interpreting neutralized segments in Dutch. *Language*, 5-38.

Ferguson, Charles A. 1959. Diglossia. *Word* 15, 325-340.

Frost, Ram, Avital Deutsch, & Kenneth I. Forster. 2000. Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26, 751-765.

Ghazeli, Salem. 1977. *Back consonants and backing coarticulation in Arabic*. Doctoral dissertation, The University of Texas at Austin.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm workshop on variation within Optimality Theory,* 111-120.

Gouskova, Maria. 2018. Morphology and Phonotactics. In *the Oxford Research Encyclopedia of Linguistics.*

DOI: 10.1093/acrefore/9780199384655.013.613.

Green, M. 2007. *Lisaan Masri*. https://www.lisaanmasry.org/

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection, in Kenneth Hale and Samuel Jay Keyser, eds., *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge, MA, 111–176.

Harrell, Richard S., Laila Tewfik, & George Selim. 1963. *Lessons in Colloquial Egyptian Arabic*. Georgetown University Press.

Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39(3), 379-440.

Herzallah, Rukayyah S. 1990. *Aspects of Palestinian Arabic phonology : A nonlinear approach.* Doctoral dissertation, Cornell University.

Idrissi, Ali, Jean-François Prunet, & Renée Béland. 2008. On the mental representation of Arabic roots. *Linguistic Inquiry* 39(2), 221-259.

Kastner, Itamar. 2016. *Form and meaning in the Hebrew verb.* Doctoral dissertation, New York University.

Laufer, Asher & Thomas Baer. 1988. The emphatic and pharyngeal sounds in Hebrew and in Arabic. *Language and speech* 31(2), 181-205.

Lehn, Walter. 1963. Emphasis in Cairo Arabic. *Language* 39(1), 29-39.

McCarthy, John J. 1979. *Formal problems in Semitic phonology and morphology.* Doctoral dissertation, Massachusetts Institute of Technology.

--- 1981. A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* 12, 373–418.

--- 1986. OCP effects: gemination and antigemination. *Linguistic Inquiry* 17(2), 207-263.

--- 1991. Semitic gutturals and distinctive feature theory. *Perspectives on Arabic linguistics* 3, 63-91.

--- 1993. Template form in prosodic morphology. In *Papers from the Third Annual Formal Linguistics Society of Mid-America Conference,* Indiana University Linguistics Club, 187–218.

--- 1994. The phonetics and phology of Semitic pharyngeals. *Papers in laboratory phonology III: Phonological structure and phonetic form* 86, 191-233.

McCarthy, John J. & Alan Prince. 1990. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8(2), 209-283.

McOmber, Michael L. 1995. Morpheme edges and Arabic infixation. *Perspectives on Arabic linguistics* 7, 173-189.

Norlin, Kjell. 1987. A phonetic study of emphasis and vowels in Egyptian Arabic. *Working papers/Lund University, Department of Linguistics and Phonetics* 30.

Omar, Margaret K. 1973. *The acquisition of Egyptian Arabic as a native language.* Walter de Gruyter.

Prunet, Jean-François. 2006. External evidence and the Semitic root. *Morphology* 16(1), 41-67.

Prunet, Jean-François, Renée Béland, & Ali Idrissi. 2000. The mental representation of Semitic words. *Linguistic Inquiry* 31(4), 609-648.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

Tucker, Matthew A. 2011. The morphosyntax of the Arabic verb: Toward a unified syntax-prosody. *UC Santa Cruz: Festschrifts*.

Ussishkin, Adam. 1999. The inadequacy of the consonant root: Modern Hebrew denominal verbs and output-output correspondence. *Phonology* 16, 401-442.

--- 2005. A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language and*

Linguistic Theory 23(1), 169-218.

Venables, Wiliam & Brian Ripley. 2002. Modern Applied Statistics with S, Fourth edition. Springer, New York. ISBN 0-387-95457-0, https://www.stats.ox.ac.uk/pub/MASS4/.

Versteegh, Kees. 2014. *Arabic language.* Edinburgh University Press.

Watson, Janet C. E. 2002. *The phonology and morphology of Arabic.* Oxford University Press.

Younes, Munther. 1994. On emphasis and /r/ in Arabic. *Perspectives on Arabic linguistics* 6, 215-215.

Zuraw, Kie. 2000. *Patterned exceptions in phonology.* Doctoral dissertation, University of California, Los Angeles.

Appendix

A. Full inflectional paradigms for Egyptian Arabic wazn I verbs

Inflections for Wazn I sound verbs in Egyptian Arabic; [katab]/[-ktib] 'write', [fihim]/[-fham] 'understand' (Harrell et al. 1963)

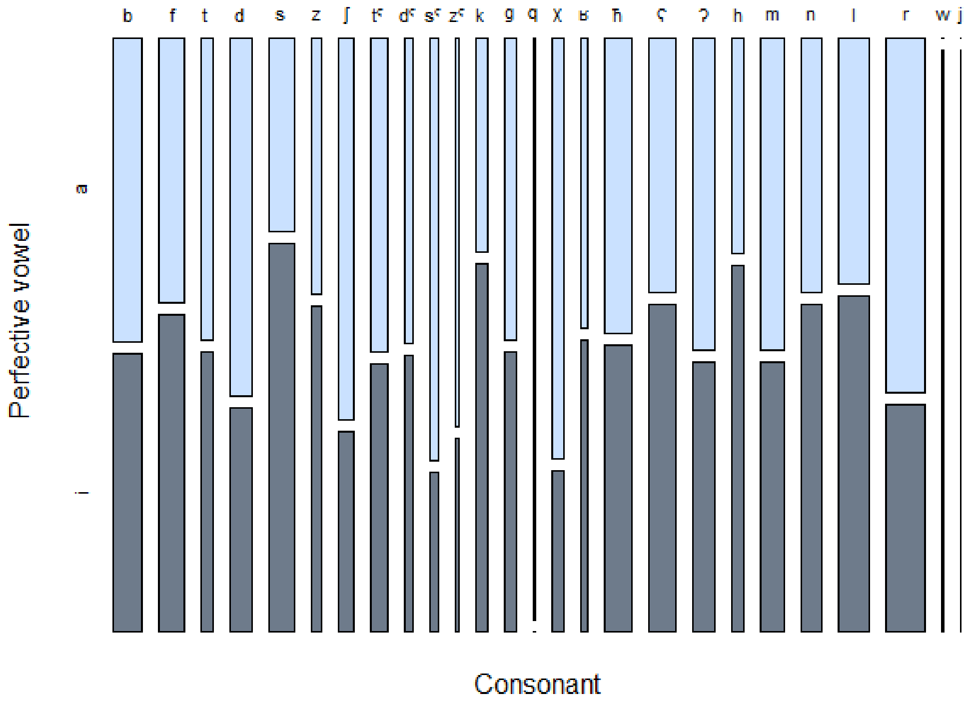
	Perfective	Imperfective	Perfective	Imperfective
1sg	katab-t	ʔa-ktib	fihim-t	ʔa-fham
1pl	katab-na	ni-ktib	fihim-na	ni-fham
2sg.m	katab-t	ti-ktib	fihim-t	ti-fham
2sg.f	katab-ti	ti-ktib-i	fihim-ti	ti-fham-i
2pl	katab-tu	ti-ktib-u	fihim-tu	ti-fham-u
3sg.m	katab	ji-ktib	fihim	ji-fham
3sg.f	katab-it	ti-ktib	fihm-it	ti-fham
3pl	katab-u	ji-ktib-u	fihm-u	ji-fham-u

Inflections for other types of Wazn I verbs in EA; defective: [bana]/[-bni] 'build'; doubled: [ħabb]/[-ħibb] 'love'; hollow: [ʃaal]/[-ʃiil] 'carry', [ʔaal]/[-ʔuul] 'say' (Harrell et al. 1963)

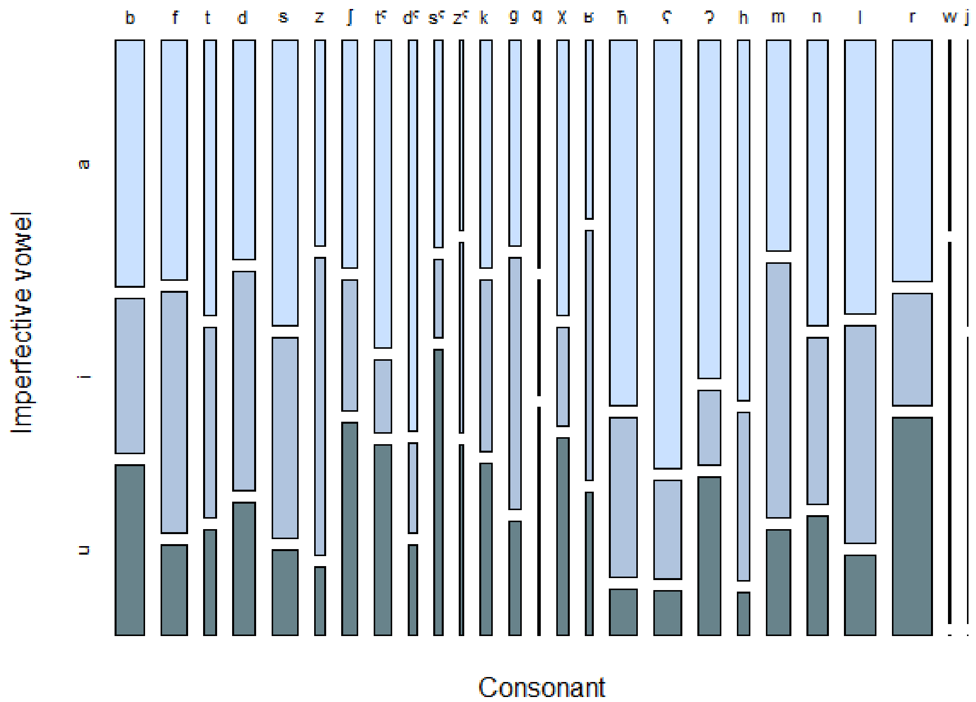
	Defective		Doubled		Hollow			
	Perf.	Imp.	Perf.	Imp.	Perf.	Imp.	Perf.	Imp.
1sg	ban-eet	ʔa-bni	ħabb-eet	ʔa-ħibb	ʃ**i**l-t	ʔa-ʃiil	ʔ**u**l-t	ʔa-ʔuul
1pl	ban-eena	ni-bni	ħabb-eena	ni-ħibb	ʃ**i**l-na	ni-ʃiil	ʔ**u**l-na	ni-ʔuul
2sg.M	ban-eet	ti-bni	ħabb-eet	ti-ħibb	ʃ**i**l-t	ti-ʃiil	ʔ**u**l-t	ti-ʔuul
2sg.F	ban-eeti	ti-bn-i	ħabb-eeti	ti-ħibb-i	ʃ**i**l-ti	ti-ʃiil-i	ʔ**u**l-ti	ti-ʔuul-i
2pl	ban-eetu	ti-bn-u	ħabb-eetu	ti-ħibb-u	ʃ**i**l-tu	ti-ʃiil-u	ʔ**u**l-tu	ti-ʔuul-u
3sg.M	bana	ji-bni	ħabb	ji-ħibb	ʃaal	ji-ʃiil	ʔaal	ji-ʔuul
3sg.F	ban-it	ti-bni	ħabb-it	ti-ħibb	ʃaal-it	ti-ʃiil	ʔaal-it	ti-ʔuul
3pl	ban-u	ji-bn-u	ħabb-u	ji-ħibb-u	ʃaal-u	ji-ʃiil-u	ʔaal-u	ji-ʔuul-u

B. Vowel distribution by individual consonants

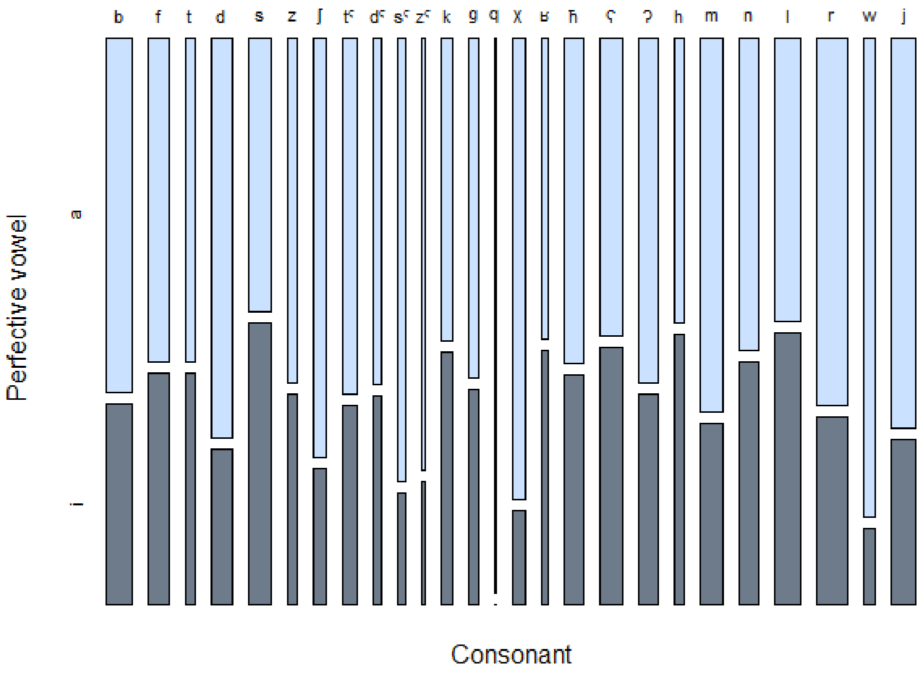
Sound verbs – perfective



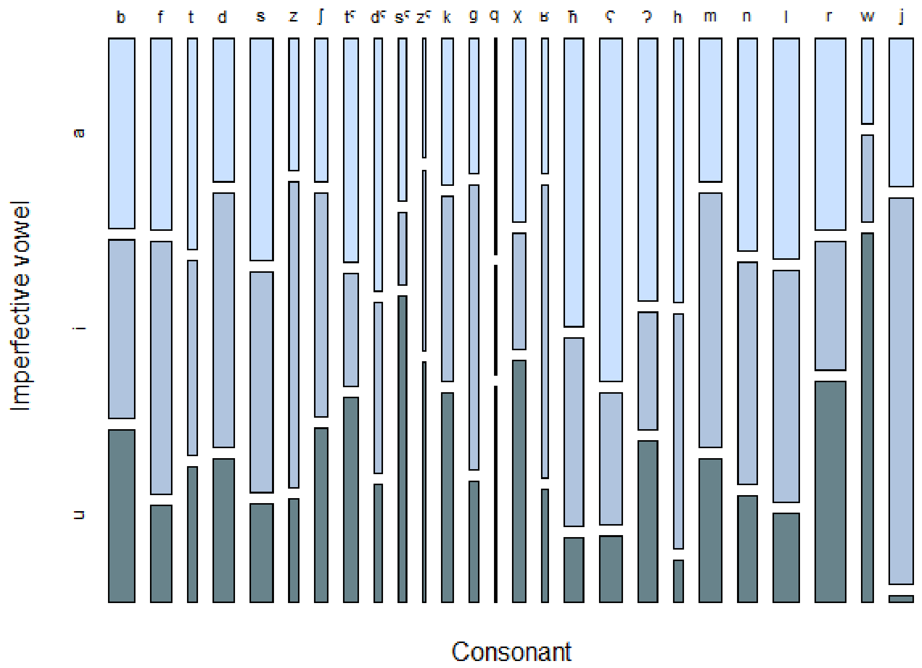
Sound verbs – imperfective



All verbs – perfective



All verbs - imperfective



C. Model selection

Imperfective-to-perfective models fitted on all verbs

Model	Predictors	AIC
1	Consonant (natural class) only	707.05
2	Consonant (natural class), imperfective vowel	595.92
3	**Consonant (natural class), imperfective vowel, verb type**	**508.24**
4	Consonnant (individual), imperfective vowel, verb type	516.12

Perfective-to-imperfective models fitted on all verbs

Model	Predictors	AIC
1	Consonant (natural class) only	1073.61
2	Consonant (natural class), perfective vowel	989.96
3	**Consonant (natural class), perfective vowel, verb type**	**884.75**
4	Consonnant (individual), perfective vowel, verb type	890.64

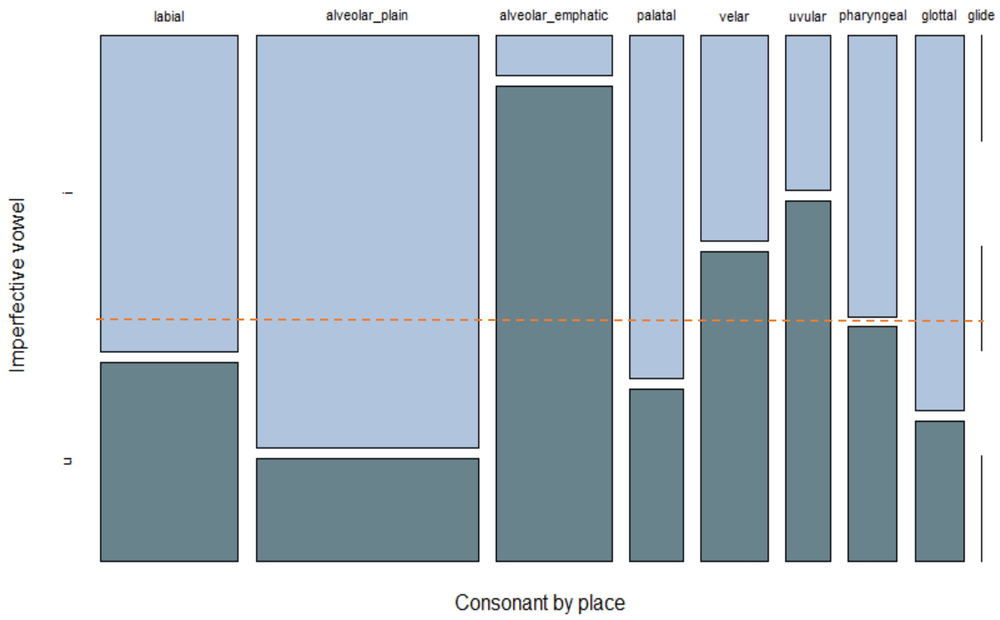
D. Breakdown for doubled, hollow and defective verbs

DOUBLED VERBS

The perfective vowel for doubled verbs is always [a], whereas the imperfective vowel can be [i] or [u]. The table and figure below show the effect of root consonants on imperfective vowel choice in doubled verbs. The corpus contains only 83 doubled verbs, so the data might suffer from more noise. There are, however, two patterns that are especially salient. Emphatic alveolars show a strong preference for [u]-imperfectives (92%) compared to the overall distribution. Plain alveolars show a strong preference for [i]-imperfectives (80%).

	i	u
labial	19 (61%)	12 (39%)
alveolar_plain	40 (80%)	10 (20%)
alveolar_emphatic	2 (8%)	24 (92%)
palatal	8 (67%)	4 (33%)
velar	6 (40%)	9 (60%)
uvular	3 (30%)	7 (70%)
pharyngeal	6 (55%)	5 (45%)
glottal	8 (73%)	3 (27%)
Total	(55%)	(45%)

Effects of consonant natural classes on imperfective vowel distribution in doubled verbs.



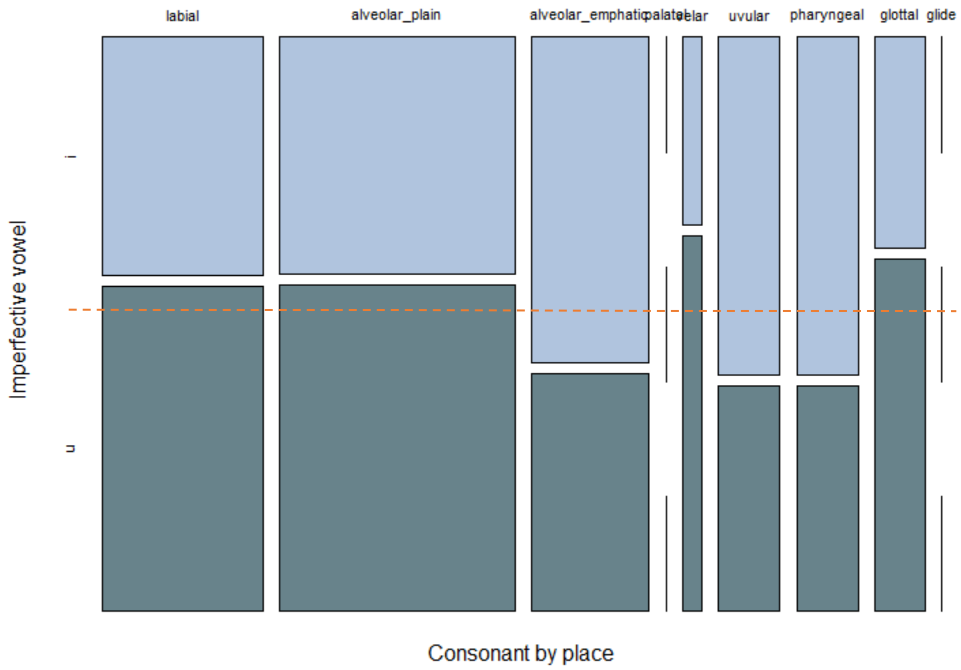
Effects of consonant natural classes on imperfective vowel distribution in doubled verbs.

HOLLOW VERBS

The perfective vowel for hollow verbs is always [a], whereas the imperfective vowel can be [a], [i] or [u]. Verbs with [a]-imperfective are very rare (4) and are excluded from the discussion below. The table and figure below show the effect of root consonants on imperfective vowel choice in hollow verbs. The data on hollow verbs likewise might suffer from more noise, since there are only 61 hollow verbs. Notably, emphatic alveolars have a slight preference for [i]-imperfectives, and plain alveolars have a slight preference for [u]-imperfectives. Both of these trends are opposite to what has been attested in sound verbs and doubled verbs.

	i	u
labial	11 (42%)	15 (58%)
alveolar_plain	16 (42%)	22 (58%)
alveolar_emphatic	11 (58%)	8 (42%)
velar	1 (33%)	2 (67%)
uvular	6 (60%)	4 (40%)
pharyngeal	6 (60%)	4 (40%)
glottal	3 (38%)	5 (63%)
Total	(47%)	(53%)

Effects of consonant natural classes on imperfective vowel distribution in hollow verbs.



Effects of consonant natural classes on imperfective vowel distribution in hollow verbs.

DEFECTIVE VERBS

The distribution of vowels in the 59 defective verbs is shown in the table below. There are more [a]-perfectives (61%) than [i]-perfectives (39%). For the imperfective, [i] (64%) is more common than [a] (36%). The alternation pattern [a]-perfective/[i]-imperfective is more common than other patterns. Consonant effects in defective verbs are difficult to analyze because of the small sample size, and there is no obvious patterns that stand out like the ones noted above in doubled and hollow verbs.

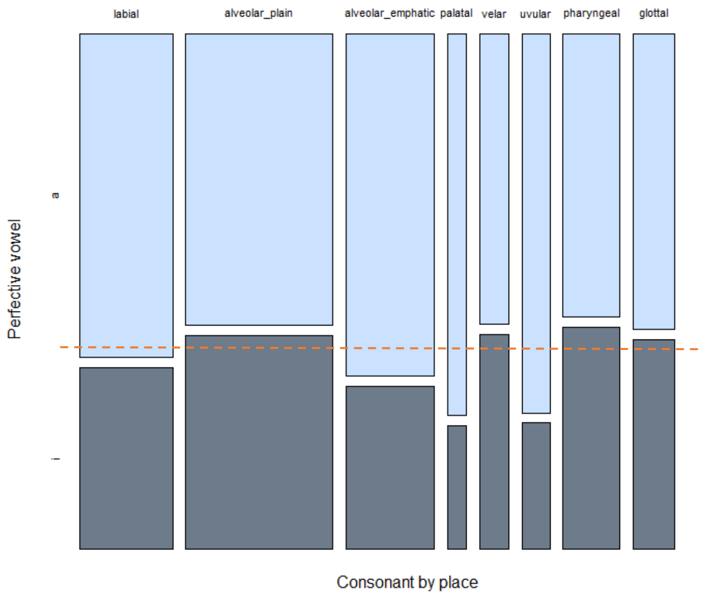
		IV		
		a	i	Total
PV	a	10 (48%)	26 (68%)	36 (61%)
	i	11 (52%)	12 (32%)	23 (39%)
	Total	21 (36%)	38 (64%)	

Perfective and imperfective vowel frequencies in defective verbs.

E. Breakdown for all verbs

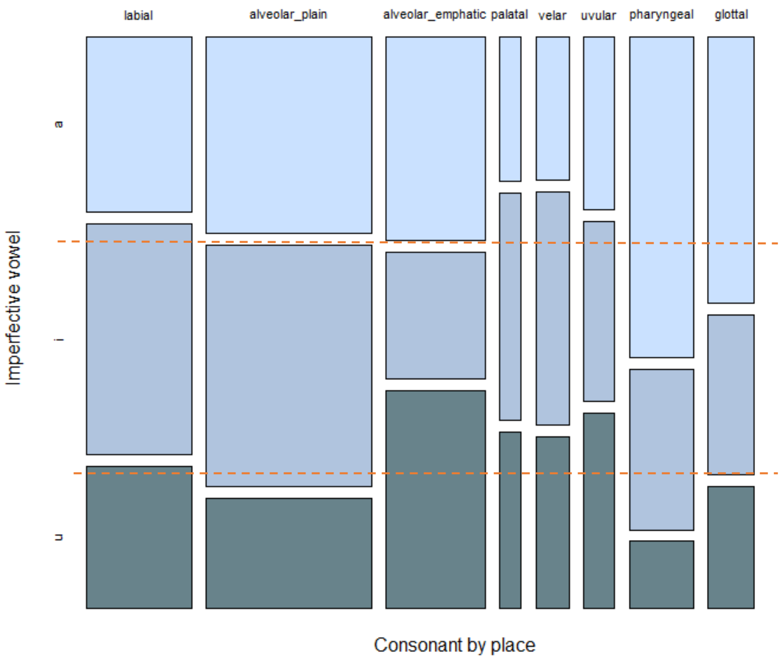
Effects of consonant natural classes on perfective vowel distribution in all verbs.

	a	i
labial	164 (64%)	92 (36%)
plain alv.	233 (58%)	171 (42%)
emph. alv.	164 (68%)	78 (32%)
palatal	40 (75%)	13 (25%)
velar	46 (58%)	34 (42%)
uvular	57 (75%)	19 (25%)
pharyngeal	88 (56%)	69 (44%)
glottal	66 (58%)	47 (42%)
Total	(63%)	(37%)



Effects of consonant natural classes on imperfective vowel distribution in all verbs.

	a	i	u
labial	82 (32%)	108 (42%)	66 (26%)
plain alv.	145 (36%)	178 (44%)	81 (20%)
emph. alv.	90 (37%)	56 (23%)	96 (40%)
palatal	14 (26%)	22 (42%)	17 (32%)
velar	21 (26%)	34 (43%)	25 (31%)
uvular	24 (32%)	25 (33%)	27 (36%)
pharyngeal	92 (59%)	46 (29%)	19 (12%)
glottal	55 (49%)	33 (29%)	25 (22%)
Total	(37%)	(38%)	(25%)



F. Models ran on all Wazn I verbs

Imperfective-to-perfective. Positive coefficients indicate preference for [i]-perfective, negative for [a]-perfective.

Predictors	Coefficients	Std.Err.	z	p	
velar	0.5270	0.3139	1.6789	0.0932	.
alveolar_plain	0.4668	0.2084	2.2402	0.0251	*
ImpV_i	0.1824	0.2556	0.7136	0.4755	
alveolar_emphatic	0.1684	0.2137	0.7879	0.4307	
glottal	0.0368	0.2617	0.1408	0.8880	
labial	0.0113	0.2093	0.0538	0.9571	
pharyngeal	-0.2332	0.2226	-1.0476	0.2948	
palatal	-0.4173	0.3786	-1.1022	0.2704	
uvular	-0.4850	0.3354	-1.4463	0.1481	
Defective	-0.7507	0.3044	-2.4656	0.0137	*
ImpV_u	-2.4563	0.4200	-5.8477	0.0000	***
Hollow	-3.8673	1.0302	-3.7538	0.0002	***
Doubled	-4.4084	1.0309	-4.2762	0.0000	***

Residual Deviance: 482.2377 AIC: 508.2377

Psudo R Square measures:

McFadden: 0.302 CoxSnell: 0.324 Nagelkerke: 0.446

Cross Validation:

Accuracy: 0.732

Perfective-to-imperfective. Pairwise comparison for [i] vs. [a] and [u] vs. [a]. Positive coefficients indicate preference for [i]- or [u]-imperfective, negative for [a]-imperfective.

	Predictors	Coefficients	Std.Err.	z	p	
i vs. a	Doubled	4.1349	1.0367	3.9884	6.65E-05	***
	Hollow	2.3126	0.5859	3.9470	7.91E-05	***
	Defective	1.0263	0.3184	3.2228	0.0013	**
	velar	0.4962	0.3229	1.5367	0.1244	
	palatal	0.4404	0.4091	1.0764	0.2817	
	labial	0.4143	0.2133	1.9421	0.0521	.
	alveolar_plain	0.3286	0.2077	1.5822	0.1136	
	PerfV_a	-0.0727	0.2598	-0.2799	0.7796	
	uvular	-0.4856	0.3546	-1.3695	0.1708	
	alveolar_emphatic	-0.6340	0.2302	-2.7538	0.0059	**
	glottal	-1.0473	0.2877	-3.6402	0.0003	***
	pharyngeal	-1.3223	0.2482	-5.3282	9.92E-08	***
u vs. a	Doubled	2.8120	1.0436	2.6945	0.0070	**
	PerfV_a	2.4528	0.3932	6.2389	4.41E-10	***
	Hollow	1.5542	0.5916	2.6270	0.0086	**
	velar	0.3118	0.3887	0.8021	0.4225	
	alveolar_emphatic	-0.0540	0.2703	-0.1997	0.8417	
	palatal	-0.3214	0.4649	-0.6914	0.4893	
	labial	-0.7248	0.2730	-2.6549	0.0079	**
	uvular	-0.9519	0.3837	-2.4808	0.0131	*
	alveolar_plain	-1.0434	0.2672	-3.9051	9.42E-05	***
	glottal	-1.4885	0.3555	-4.1868	2.83E-05	***
	pharyngeal	-2.6297	0.3517	-7.4773	7.59E-14	***
	Defective	-3.6824	1.0589	-3.4777	0.0005	***
Residual Deviance: 836.7546		AIC: 884.7546				
Psudo R Square measures:						
McFadden: 0.278		CoxSnell: 0.454		Nagelkerke: 0.512		
Cross Validation:						
Accuracy: 0.651						

G. Detailed predictions by the models ran on all verbs

Imperfective-to-perfective model predictions by rooty type and type of vowel alternation.

Verb type	Perf.V/Imp.V	Freq in lexicon	# correct prediction	Accuracy	Freq in lexicon	# correct prediction	Accuracy
Sound	**a**/a	71	17	24%	330	217	66%
	a/i	30	3	10%			
	a/u	67	67	100%			
	i/a	89	68	76%			
	i/i	66	62	94%			
Defective	**a**/a	10	10	100%	59	37	63%
	a/i	26	25	96%			
	i/a	11	0	0%			
	i/i	12	2	17%			
Doubled	**a**/i	46	46	100%	83	83	100%
	a/u	37	37	100%			
Hollow	**a**/i	27	27	100%	57	57	100%
	a/u	30	30	100%			

Perfective-to-imperfective model predictions by rooty type and type of vowel alternation.

Verb type	Perf.V/Imp.V	Freq in lexicon	# correct prediction	Accuracy	Freq in lexicon	# correct prediction	Accuracy
Sound	a/**a**	71	54	76%	330	217	66%
	a/**i**	30	10	33%			
	a/**u**	67	49	73%			
	i/**a**	89	76	85%			
	i/**i**	66	28	42%			
Defective	a/**a**	10	5	50%	59	39	66%
	a/**i**	26	21	81%			
	i/**a**	11	1	9%			
	i/**i**	12	12	100%			
Doubled	a/**i**	46	41	89%	83	65	78%
	a/**u**	37	24	65%			
Hollow	a/**i**	27	11	41%	57	24	42%
	a/**u**	30	13	43%			