Consonant-vowel interactions inform paradigm organization in Egyptian Arabic verbs

Lily Xu
UCLA

Introduction. This paper presents a novel corpus study of Egyptian Arabic verbal paradigms, showing evidence for bidirectional predictability between the imperfective (tenseless) forms and the perfective (tensed) forms. In particular, imperfective forms are predictable based primarily on consonant features of the root, while perfective forms are predictable based solely on vowel information from the imperfective. The vowels in both forms are idiosyncratic, as illustrated in (1), so all predictors are probabilistic. In comparison to paradigmatic models in which a single base can be used to predict all other forms (Albright 2002), the Egyptian Arabic corpus data suggest a paradigm structure that is more complex.

(1) Perfective (CVCVC)	Imperfective (-CCVC)	Gloss
kasar	-ksar	'break'
katab	-ktib	'write'
χarag	-χrug	'leave'
rigiʕ	-rgaʕ	'return'
libis	-lbis	'wear'

Methods. A corpus of 330 verbs in Egyptian Arabic was collected by working with a native speaker. Subsequently, two logistic regression models were run to investigate the predictability of vowel choice for the two aspectual forms based on consonant and vowel information. Both consonant (*labial, plain alveolar, pharyngealized alveolar, palatal, velar, uvular, pharyngeal, glottal*) and vowel (*imperfective: [i], [u]; perfective: [a]*) predictors were used in each regression. The coefficients that the models assign to the predictors are analogous to constraint weights in a MaxEnt grammar (Goldwater & Johnson 2003, Hayes & Wilson 2008).

Results. In the model predicting the perfective vowel, only the imperfective vowel predictors were significant. Imperfective [u] introduces a strong bias for perfective [a] ($\beta = -2.47$, $p < 0.001$), whereas imperfective [i] introduces a moderate preference for perfective [i] ($\beta = 0.59$, $p = 0.04$). These trends are illustrated in Fig. 1. None of the consonant predictors came out significant or had a large coefficient in this model.

Figure 1

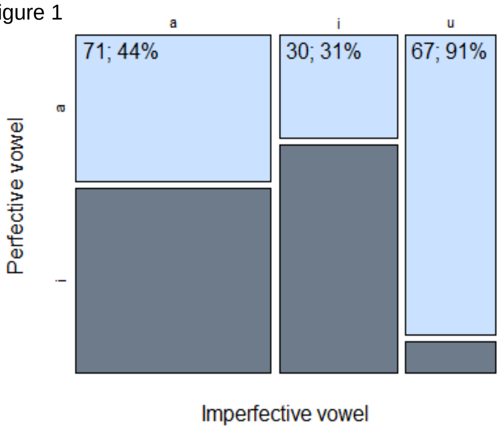
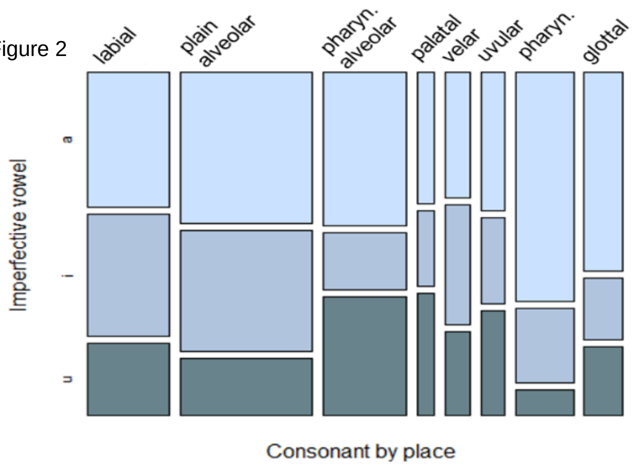


Figure 2



In the opposite-direction model, predicting the imperfective vowel, consonant predictors play a major role (Fig. 2). These predictors show preference toward phonologically natural consonant-vowel cooccurrences, as listed in (2). For example, for both pharyngeals and glottals, [a] is the most common imperfective vowel, with no difference between [i] and [u]. Pharyngealized alveolars prefer both [a] and [u] over [i]. The affinity for low vowels in both cases is not surprising since these consonants all involve tongue root retraction. Pharyngealized alveolars exert additional backing effects on neighboring vowels (Norlin 1987), resulting in the affinity for [u]. Similar consonant-vowel interactions in the imperfective verb forms have been noticed for Palestinian Arabic (Herzallah 1990), Modern Standard Arabic (McCarthy 1994), and Hijazi Arabic (Ahyad & Becker 2020). The perfective vowel plays a smaller, yet still significant role compared to the consonants. Perfective [a] increases the chance of imperfective [u] ($\beta = 2.59$, $p < 0.001$), but has no significant effect for [i] ($\beta = -0.50$, $p = 0.11$).

(2) Consonant predictors	Imperfective vowel preferences	[u] vs. [a] *coeff.*	[i] vs. [a] *coeff.*
pharyngeal	[a] >> [u] = [i]	-2.88, p < 0.001	-1.60, p < 0.001
glottal	[a] >> [u] = [i]	-1.53, p < 0.001	-1.31, p < 0.001
uvular	[a] = [i] >> [u]	-1.04, p = 0.02	-0.36, p = 0.40
pharyn. alveolar	[a] = [u] >> [i]	-0.12, p = 0.72	-0.84, p < 0.01
plain alveolar	[i] >> [a] >> [u]	-1.02, p < 0.01	0.56, p = 0.03
labial	[i] >> [a] >> [u]	-0.81, p = 0.02	0.85, p < 0.01
palatal	[a] = [i] = [u]	-0.27, p = 0.58	-0.01, p = 0.98
velar	[a] = [i] = [u]	0.01, p = 0.99	-0.36, p = 0.40

The models reveal comparable predictability in the two paradigmatic directions. Cross-validation results show that perfective vowels can be predicted at a slightly higher accuracy (63.9%) than imperfective vowels (60.6%). Yet when compared to chance-level performance, the model predicting the imperfective vowel shows greater improvement, as supported by Nagelkerke's Pseudo R^2 (0.469), higher than the model predicting perfective (0.284).

Discussion. The modeling results show that the imperfective vowel can be predicted based primarily on the place of articulation of consonants, whereas the perfective vowel can be predicted based only on the imperfective vowel. The roughly equal predictability for both forms suggests that speakers might be able to learn bidirectional generalizations on this paradigm, which does not align with models in which one base is used to predict the rest of the paradigm (Albright 2002). 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, thus 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 and imperfective forms.

One analysis is that imperfective verbs are derived from the consonantal root alone, while perfective verbs are derived from the imperfective. Such an analysis follows the approach of treating the consonantal roots as morphemes in Semitic languages, which has been pursued by several lines of work in the phonological (Brame 1970, McCarthy 1979), morphosyntactic (Arad 2007, Kastner 2016), and experimental (Boudelaa & Marslen-Wilson 2001, 2011) domains. An advantage of this analysis is that it aligns with syntactic evidence suggesting that imperfective verbs in Arabic have simpler morphosyntactic structure than perfective verbs (Benmamoun 1999). This analysis also provides two explanations for the absence of consonant-vowel interactions in the perfective. First, it instantiates a widely attested pattern in which phonotactic restrictions are stronger across smaller morphological domains (Gouskova 2018). Second, there is evidence for strict syntactic locality constraints which disallow phonological interactions between any non-adjacent elements in the morphosyntactic structure (Embick 2010). Since the consonantal root is the most deeply embedded, these constraints can account for why the perfective derivation is blind to its effects. This analysis, however, cannot account for why the perfective vowel has significant contribution when predicting the imperfective vowel. Nonetheless, Ahyad's (2019) experimental work on Hijazi Arabic shows that speakers fail to learn a similar pattern, so a model on speaker's knowledge of Arabic verbal morphology may be right to not account for this pattern.

In sum, the **bidirectional predictability** and the **asymmetry of consonant-vowel interactions** demonstrated in this paper suggest a complex paradigmatic structure of the Egyptian Arabic verb. Experimental work on speakers' knowledge of this system will be crucial for theories of representations of morphological paradigms.