

The breakdown of abstract underlying representations in Maga Rukai

Jennifer Kuo

1 Introduction

In generative phonology, alternations are explained by positing abstract underlying forms (URs), with rules (or constraints) that derive the correct surface forms. Once we allow URs to deviate from surface forms, it also becomes important to ask how abstract can URs be. In other words, how much can URs diverge from their surface representations? Early work in generative phonology, centered around SPE (Chomsky and Halle, 1968), allowed for high degrees of abstractness in underlying forms. Since then, however, various scholars have questioned how abstract URs can be. In other words, it could be that speakers cannot learn overly abstract URs, and morphophonological patterns which require positing highly abstract URs will either be unproductive, or restructured by learners into more concrete patterns. Kiparsky's (1968; 1973; 1982) work was particularly influential in highlighting possible limits abstractness, drawing on data from language change.

When probing at the learnability of abstract URs, we should also make a distinction between levels of abstraction. At one end of the spectrum, learners might tolerate no abstractness at all, and only ever generalize from surface allomorphs of one particular slot in a paradigm. On the other end of the spectrum, learners may be able to learn phonemes that are never even observed in the surface forms of a language. In this paper, I adopt Kenstowicz and Kisseberth's (1977, ch. 1) taxonomy of UR abstractness, which defines levels of abstractness based on how divergent a UR is from its associated SRs.

Maga Rukai, a dialect of Rukai (Austronesian), is particularly suitable for addressing this issue because it exhibits rhythmic vowel deletion (syncope), where broadly speaking, every odd medial syllable counting from the left is deleted. Representative examples are given in (X), using the stem and negative forms, which are marked by both the prefix /i-/ and final vowel lengthening; nouns and stative verbs are additionally prefixed with /k-/. For ease of reading, the deleted vowel is indicated with an underline. As shown in this examples, prefixation shifts which vowels get deleted, resulting in $\emptyset \sim V$ alternations between the stem and negative forms.

As a consequence of rhythmic syncope, many stems require positing a composite UR, which are abstract in the sense that the UR is not identical to any one surface allomorph. For example, for the word 'salt', the UR must take its initial vowel from the negative form and its second vowel from the stem form. If the syncope pattern is stable over time, this would suggest that speakers are able to learn composite URs. On the other hand, restructuring of the pattern over time would suggest that composite URs are unlearnable (or highly dispreferred), and will be mis-learned in such a way that allows learners to posit less abstract URs.

(1) *Examples of $\emptyset \sim V$ alternations in Maga*

UR	STEM	NEGATIVE	GLOSS
/timusu/	t_musu	i-k-tim_su:	‘salt’
/ɖamari/	ɖ_mari	i-k-ɖam_ri:	‘moon’
/subuku/	s_buku	i-k-sub_ku:	‘eyeball’

In this paper, I track how the syncope pattern has been restructured over time. To preview the results, I find that while rhythmic syncope still persists in some Maga Rukai forms, many $\emptyset \sim V$ alternations have been removed through paradigm leveling. Where leveling has not occurred, forms have still been reanalyzed in such a way that makes $\emptyset \sim V$ alternations more predictable. These results provide evidence against the learnability of composite URs. Interestingly, however, the direction of leveling suggests that speakers are amenable to learning some level of abstraction.

There are paradigms in the modern language that require a composite UR, that is a UR that is not identical to any one surface form.

ADD: modeling?

2 Maga Rukai

Maga is one of the six dialects of Rukai, an Austronesian language spoken in Southern-Central Taiwan (the other five being Tanan, Budai, Tona, Labuan and Mantauran) (Li 1977a). The Maga dialect is spoken in Maolin Township; the population of the township is around 866 (Kaohsiung City Government Civil Affairs Bureau, 2023), but the number of fluent speakers is much less than that. The name Maga comes from the historic name of Maolin Township. The local community identifies themselves and the dialect by the name *tiɖilka*, but I use the term Maga to be consistent with prior literature.

Earlier descriptive work on the Maga dialect includes Li (1975) and Saillard (1995). Hsin (2000, 2003) provides a thorough description and analysis of Maga phonology and morphophonemic alternations; subsequent work by Chen (2008) focuses on the rhythmic syncope patterns. Additionally, Li (1977) has done extensive work comparing the Rukai dialects and working out their internal subgrouping. Li (1977), in particular, reconstructs and phonological system of Proto-Rukai and provides a list of around 500 cognates from dialects of Rukai. Other work comparing the Rukai dialects includes Zeitoun (1995), which is a cross-dialectal overview of Rukai phonology and morphosyntax.

In the rest of this paper, unless otherwise stated, modern Maga forms are taken from Hsin (2000), supplemented with data from Li (1975). The resulting corpus has 1049 stems; of these, a subset of 790 words are also given with the negative prefixed form. Proto-Rukai forms are taken from Li (1977), who provides a list of 532 Proto-Rukai forms with their cognates in the six Rukai dialects.

2.1 Basics of Maga phonology

The Maga consonant inventory is given in Table 1. In addition, the language has seven vowels, /i e u o i ə a/. The two central vowels /i/ and /ə/ are marginally contrastive (and actually descended from the same phoneme in Proto-Rukai); there are numerous transcription inconsistencies for the two vowels across authors, and the perceptual distinction between the two vowels is described as unclear, even for native speakers (Hsin, 2000, p. 37). However, there are minimal pairs between the two, such as [biki] ‘pig’ vs. [bəki] ‘nose running’.

	Labial	Dental	Alveolar	Retroflex	Dorsal
Stop	p b		t d	ɖ	k g
Affricate			c [tʃ]		
Fricative	v	θ (ð)	s (z)		
Nasal	m	n		ɳ	
Lateral			l		
Trill			r ~ ʀ		

Table 1: Consonant inventory of Maga Rukai

Stress in Maga is argued to be predictable, falling on either the penultimate syllable (e.g. [ábu] ‘ash’) or on final heavy (C)V: syllables (e.g. [adó:] ‘mud’). Maga syllables have a (C)(C)V(C) structure with some additional restrictions. First, no final consonants are observed. Additionally, consonant clusters are restricted to two segments, regardless of position (initial vs. medial).

2.2 Rhythmic syncope in Maga Rukai

In Maga, vowels are deleted in odd medial syllables counting from the left; as will be further discussed below, this means that initial (onsetless) vowels and final vowels do not delete. Affixation can shift the position of the deleted vowel, resulting in $\emptyset \sim V$ alternations between forms of the same paradigm. This is demonstrated in (2) for the stem and negative form of ‘shrimp’; negation is marked with both the prefix /i-/ and lengthening of the final syllable. Nouns and stative verbs take an additional prefix /k-/. More examples of alternations between the stem and negative forms are given in (3) below.¹ Note also that for underlyingly disyllabic stems like (3a-b), syncope is accompanied by lengthening of the remaining surface vowel (e.g. [smá:], cf. *[smá]). This could potentially be attributed to a typologically well-attested word-minimality requirement, where content words are minimally bimoraic (Hayes, 1995).

(2) Derivation for ‘shrimp’

UR	/ta ₁ ka ₂ su ₃ lu ₄ ɖu ₅ /	/i ₁ -k- ta ₂ ka ₃ su ₄ lu ₅ ɖu ₆ /
Syncope	t_kas_luɖu	ik-tak_sul_ɖu:
Stress	tkaslúɖu	iktaksulɖú:
SR	[tkaslúɖu]	[iktaksulɖú:]

¹add footnote about metrical incoherence

(3) *Examples of rhythmic syncope in Maga*

	UR	CITATION	NEGATIVE	GLOSS
a.	/sima/	smá:	ik-simá:	‘fat’
b.	/rana/	rná:	ik-raná:	‘creek’
c.	/ɖamari/	ɖmári	ik-ɖamrí:	‘moon, month’
d.	/timusu/	tmúsu	ik-timśu:	‘salt’
e.	/giŋigiŋi/	gŋígŋi	ik-giŋgiŋí:	‘longan’
f.	/takasuluɖu/	tkaslúɖu	ik-taksulɖú:	‘shrimp’

Rhythmic syncope, or the deletion of vowels in metrically weak positions, has been observed in many languages, including Macushi (Hawkins, 1950; Kager, 1997), Nishnaabemwin (Valentine, 2001; Bowers, 2012), and Southern Pomo (Kaplan, 2022). Notably, rhythmic syncope presents a challenge for parallel theories of phonology like OT, as syncope must happen after building the metrical structure which conditions it Kager (1997); McCarthy (2008). In the case of Maga Rukai, there is additional opacity because the location of surface stress does not always match the location of syncope; as pointed out by Kaplan (2022), who describes a similar pattern in Southern Pomo, this means that no clear surface-oriented explanation can account for both stress assignment and syncope (Kaplan, 2022). The conflict between stress and syncope is discussed below in Section X.

Overall, the opacity presented by rhythmic syncope makes it arguably less learnable, and therefore diachronically unstable. For example, Bowers (2012) shows that in Nishnaabemwin, extreme vowel reduction resulted in a rhythmic syncope pattern, but this pattern broke down within just a few generations of speakers. This characteristic makes rhythmic syncope a particularly good case study for understanding how paradigms are restructured over time.

In this paper, I focus on comparisons between the citation form and negative prefix, because these forms are the most well-attested and found across all lexical categories. However, note that not all affixes in Maga are in the domain of syncope (i.e. they do not trigger syncope). Affixes that trigger syncope include the negative /i-/ , plural /l-/ , and imperative suffix /-a/. On the other hand, Rukai also has a set of verb marking prefixes which obligatorily appear in the citation form but are not in the domain for syncope; these include /ma-/ ‘stative verb marker’, /u-/ ‘dynamic verb marker’, and a few other less productive verbalizing prefixes. This contrast in affix types is demonstrated in (4). For example, in the citation form of ‘surprised’, [ma-rkími], the stem still behaves as if it is at the left edge of the word, with syncope happening on the stem-initial vowel.

(4) *Verbal prefixes are not in the domain of syncope*

	UR	citation	negative	gloss
	/rikimi/	ma-rkími	i-k-rikmí:	‘surprised’
	/busuku/	ma-bsúku	i-k-busku:	‘get drunk’
	/kire/	u-kré:	i-kiré:	‘hand’

2.3 Blocking of syncope

There are several regular (phonologically-driven) exceptions to the syncope process. First, there is a near-exceptionless tendency for syncope to be blocked in onsetless syllables. Following Chen (2008), this can be characterized as the effects of a positional faithfulness constraint MAX-V/[σ], which privileges syllable-initial positions (Beckman, 1998).

Examples of hiatus blocking syncope are given in (5) for vowels in word-initial position (5a), medial position in the unaffixed stem (5b), and medial position in the negative form (5c). The vowel that should have been deleted (i.e. is underlyingly in an odd-numbered syllable counting from the left edge) is underlined. Overall, hiatus blocks syncope in 94% of vowels in onsetless syllables that should otherwise undergo deletion. Additionally, of the words counted as exceptions, 5 have variants where the vowel in hiatus is not deleted. For example, ‘broom’ /suəpi/ has stem variants [səpi] and [suəpi].

(5) Onsetless vowels do not delete (194/206; 94%)

a. Initial vowels (N=113/117; 97%)

STEM	UR	GLOSS
<u>á</u> bu (cf. *bú)	/abu/	‘ash’
<u>a</u> lápi (cf. *lápi)	/alapi/	‘slate’
<u>u</u> váci (cf. *váci)	/uvaci/	‘vein’

b. Medial onsetless stems (N=51/56; 91%)

STEM	UR	GLOSS
n <u>u</u> əŋi (cf. *nəŋi)	/nuəŋi/	‘cattle, cow’
ŋ <u>i</u> u (cf. *ŋú)	/ŋiu/	‘owl’
tlap <u>u</u> i (cf. *tlápi)	/talapui/	‘firefly’

c. Medial onsetless negative forms (N=30/33; 91%)

NETATIVE	UR	GLOSS
i-k-nu <u>ə</u> ŋi: (cf. *i-k-nuŋi:)	/nuəŋi/	‘cattle, cow’
i-ka <u>u</u> cú: (cf. *i-kacú:)	/kaucu/	‘mean (temper)’
i-kul <u>u</u> əŋə: (cf. *i-kuləŋə:)	/kuluəŋə/	‘sparrow’

Additionally, syncope is generally blocked if it would result in a geminate consonant; examples of this are given in (6). Note also that there is a contrast between initial and medial geminates; syncope is always blocked if it would result in an initial geminate (6a), but only variably blocked in the case of medial geminates (6b,c). The examples in (6d) show cases where syncope applies, even if this results in medial geminates in the surface form.

(6) *Deletion is blocked by geminates*

a. *Initial geminates (N = 25/25; 100%)*

STEM	UR	GLOSS
t <u>u</u> túku (cf. *ttúku)	/tutuku/	‘rabbit’
m <u>u</u> muóli (cf. *mmuóli)	/mumuəli/	‘snail’
r <u>a</u> rámi (cf. *rrámi)	/rarami/	‘bird’

b. *Medial geminates in stems (N = 33/51; 65%)*

STEM	UR	GLOSS
θev <u>a</u> va (cf. *θlévva)	/θelevava/	‘rainbow’
u-m <u>u</u> múdu (cf. *u-mmúdu)	/mumuðu/	‘to bud/flower’
vlir <u>u</u> ru (cf. *vlírru)	/valiruru/	‘glutinous rice’

c. *Medial geminates in negative forms (N = 28/32; 88%)*

NEGATIVE	UR	GLOSS
i-k-porirí: (cf. *i-k-porrí:)	/poriri/	‘shade’
i-k-tagagá: (cf. *i-k-taggá:)	/tagaga/	‘crow’
i-pasísí: (cf. *i-passí:)	/pasisi/	‘to sift’

d. *Examples where syncope is not blocked by medial geminates*

STEM	UR	GLOSS
brúccu (cf. *bruc <u>ú</u> cu)	/barucucu/	‘egg’
u-ssípi (cf. *u-sisípi)	/sisipi/	‘count’
ma-nnǫ́ki (cf. *ma-nanǫ́ki)	/naniki/	‘breathe’

2.4 Other phonological alternations

In this section, I detail two morphophonemic alternations that interact with rhythmic syncope, but are not of direct interest to the current paper. First, in positions where vowels would otherwise delete, high vowels in preceding another vowel (i.e. in hiatus) become glides. Examples of glide formation are shown in (7).

(7) *Glide formation in VV sequences*

STEM	UR	GLOSS
kmus <u>y</u> á:	/kamusia:/	‘sugar’
ma-ly <u>a</u> ci	/liaci/	‘blind’
arib <u>w</u> á:	/aribua:/	‘sweat’
u-k <u>w</u> ǎŋi	/kuani/	‘shoot (with firearm)’

Alternations between the high vowels /i, u, ɨ/ and mid vowels /e, o, ə/ are also observed, as demonstrated by the examples in (8). Following Hsin’s (2003) description, all occurrences of /i, u, ɨ/ following an /a/ are lowered to [e, o, ə] respectively, but *only* if the /a/ is deleted. In

words like (8a-b), this causes lowering of /i/ in the stem form, while in words like (8c), lowering happens in the negative form. Words like [amici] (*[amece]) ‘tree root’ and [pagu] (*[pago]) ‘gallbladder’ demonstrate that lowering does not happen if the preceding /a/ is not deleted.

(8) *Vowel lowering (alternating vowels are underlined)*

	UR	STEM	NEGATIVE	GLOSS
a.	/damili/	dme <u>le</u>	i-k-damli:	‘hemp’
b.	/caki/	cke:	i-k-caki:	‘excrement’
c.	/alapi/	alapi	i-k-alpe:	‘slate’
d.	/valu/	vlo:	i-k-valu:	‘bee’
e.	/taludu/	tlodo	i-k-taldu:	‘bridge’
f.	/tesavunju/	pa-tsavɲu	i-p-tesvɔŋo:	‘bridge’
g.	/θadi/	u-θdǎ:	i-θadǎ:	‘headdress for women’
h.	/θadi/	u-θdǎ:	i-θadǎ:	‘to raise (animal, children)’
i.	/akanji/	akánjǐ	i-k-akŋǎ:	‘fish’

2.5 Opacity in the interaction of syncope and stress assignment

In Rukai, deletion is assigned on odd syllables counting from the left, but stress surfaces on the penultimate syllable (i.e. second, counting from the right). As a result, for words of even syllable parity, syncope is inconsistent with surface stress. This inconsistency between stress and syncope is schematized in (9). Under the UR column, the position that should have been stressed (i.e. is in penultimate position) is underlined. For stems that have underlying even parity, syncope targets the position that should have been stressed, had deletion not occurred.

This is a type of *metrical incoherence*, where a metrically conditioned process (in this case syncope) does not appear to refer to surface stress (Dresher and Lahiri, 1991). As a result, there is no straightforward surface-oriented solution that can derive both stress and syncope. Prior analyses of Maga syncope have therefore either opted from derivational solutions (Hsin, 2000), or include phonemic stress in underlying representations (Chen, 2008).

(9) *Interaction of syncope and stress assignment by syllable parity*

Parity	UR	SR	Example		
Even	/σ ₁ σ ₂ /	[∅ ₁ 'σ ₂]	/t̪i̯ba/	tbá:	‘wild yam’
	/σ ₁ σ ₂ σ ₃ σ ₄ /	[∅ ₁ 'σ ₂ ∅ ₃ σ ₄]	/tavanu/	tvánlu	‘chili pepper’
Odd	/σ ₁ σ ₂ σ ₃ /	[∅ ₁ 'σ ₂ σ ₃]	/ɖamari/	ɖmári	‘moon’
	/σ ₁ σ ₂ σ ₃ σ ₄ σ ₅ /	[∅ ₁ σ ₂ ∅ ₃ 'σ ₄ σ ₅]	/takasulu/	tkaslúdu	‘shrimp’

There is evidence that metrical incoherence makes a pattern less learnable (Kaplan, 2022). If this is the case, we might expect that words of even syllable parity are more likely to be mislearned and reanalyzed than ones of odd syllable parity.

2.6 Historical origins of the rhythmic syncope pattern

The Maga pattern of rhythmic syncope is likely to have developed from severe phonetic reduction of unstressed vowels (followed by reassignment of stress to be penultimate). Similar patterns of diachronic change have been reported from languages like Nishnaabemwin (Valentine, 2001; Bowers, 2012). In general, when Maga stem-negative paradigms are compared against cognates in other Rukai dialects, the vowels which surface match that of cognate words.

(10) Cross-dialectal comparison of Rukai forms

	Maga paradigms	Proto-Rukai	Tanan	Budai	Tona	Mantauran
‘tongue’	rdámə~ikrídṁṁ:	*lídámə	lídámə	lídámə	idámə	lídámə
‘moon’	ḑmári~ikḑamrí:	*ḑamarə	ḑamárə	ḑamarə	ḑamáə	ḑamarə

The picture that emerges is that Maga Rukai, had a point earlier in history, paradigms with no Ø~V alternations, and therefore did not require positing abstract URs. The subsequent loss of metrically weak vowels then resulted in a synchronic pattern which requires composite URs. The (in)stability of the syncope pattern over time can then give us insight into how well abstract representations are learned and tolerated.

3 Formalizing degrees of UR abstraction in Maga

In this paper, I adopt Kenstowicz and Kisseberth’s (1977, ch. 1) formalization of UR abstractness; following Wang and Hayes (under revision), I refer to this hierarchy as the KK-hierarchy. The KK-hierarchy not only allows for a more fine-grained evaluation of how abstract URs can be in Maga, but also allows for comparison with recent work by Wang and Hayes (under revision), who propose a model of UR learning that is situated using the same hierarchy. In this section, I first give an overview of the KK-hierarchy, characterizing the predicted types of restructuring, if learners are restricted to each level of the hierarchy. I then characterize the amount of abstraction required in Maga.

As it turns out, if syncope regularly applies in Maga, most of the lexicon would require positing relatively abstract composite URs. However, many stems fail to undergo the expected syncope pattern; in the following section (Section X), I describe the types of restructuring that have occurred, all of which suggest that Maga has changed in concretist directions.

3.1 The KK-hierarchy

KK-A: only invariant properties of allomorphs

Under KK-A, the UR of a morpheme “consists of all and only the invariant phonetic properties of that morpheme’s various SRs” (Kenstowicz and Kisseberth, 1977, p. 8). This is highly restrictive, and disallows standard phonemicization. For example, English “pan” [pæn], where the vowel

is predictably nasalized in pre-nasal position, would have the UR /pǣn/. KK-A', which is a less restrictive variant of KK-A, states that the UR consists only of invariant features of corresponding SRs, but that alternations may be treated as unspecified (rather than listed in the UR).

Both KK-A and KK-A', when applied to Maga Rukai, would not allow any of the alternating vowels to be present in the UR. A root like 'moon', given again in (11), would have the UR /ɖmri/, which includes only the stem-final vowel. If KK-A is a hard restriction on UR learning, then URs should be reanalyzed in a way that renders the alternating vowels predictable from the non-alternating parts of the stem. Hypothetically, this could result in a grammar like (12a), where only the non-alternating parts of a root has been retained. Alternatively, Maga might be restructured so that the vowel(s) that surface are always predictable (e.g. epenthesize a default 'least marked' vowel), as in (12b).

(11) *Maga URs under KK-A*

stem	negative	gloss	UR-A
ɖmari	i-k-ɖamri:	'moon'	/ɖmri/
tmsu	i-k-timsu:	'salt'	/tmsu/
smá:	ik-simá:	'fat'	/sma:/

(12) a. Grammar A1: only retain non-alternating parts of the paradigm.

UR	STEM	NEGATIVE
/ɖmri/	→ [ɖmri]	[i-k-ɖmri:]
/tmsu/	→ [tmsu]	[i-k-tmsu:]

b. Grammar A2: always epenthesize the same vowel (e.g. [i])

UR	STEM	NEGATIVE
/ɖmri/	→ [ɖm̥iɾi]	[i-k-ɖ̥im̥ri:]
/tmsu/	→ [tm̥isu]	[i-k-t̥imsu:]

KK-B: single allomorph

Broadly, all variants of KK-B have the restriction that all URs must be based on the same surface allomorph (i.e. the same slot in a paradigm). Under KK-B, the UR includes idiosyncratic properties of a morpheme, but must be based on the surface allomorph "that appears in isolation (or as close to isolation as the grammar of the language permits)" (Kenstowicz and Kisseberth, 1977, p. 11) (p. 11). In other words, KK-B requires that the UR be based solely on the isolation allomorph. The less restrictive KK-B' states that the learner should base the UR on the allomorph that appears in the most contexts, while KK-B", which has been extensively explored in work by Albright (2002a,b, 2010, et seq.), restricts URs to be based on the most informative slot in a paradigm.

The subtypes of KK-B, when applied to Maga, predict grammars like (13a), where paradigms have leveled towards the stem form, or grammars like (13b), where paradigms have leveled towards the negative allomorph.² These two example grammars here have no vowel alternations, but KK-B also allows for grammars like (13c), where there are $\emptyset \sim V$ alternations, but the alternating vowels are completely predictable from one surface allomorph; in this hypothetical grammar, the vowel that surfaces in the negative form is a copy of the stem vowel.

- (13) a. Grammar B1: leveling towards the stem allomorph.

UR	STEM	NEGATIVE
/ɖmari/	→ [ɖmari]	[i-k-ɖmari]
/tmusu/	→ [tmusu]	[i-k-tmusu:]

- b. Grammar B2: leveling towards negative allomorph.

UR	STEM	NEGATIVE
/ɖamri/	→ [ɖamri]	[i-k-ɖimri:]
/timsu/	→ [timsu]	[i-k-timsu:]

- c. Grammar B3: $\emptyset \sim V$ alternations are predictable from the stem allomorph

UR	STEM	NEGATIVE
/ɖmari/	→ [ɖmari]	[i-k-ɖamri]
/tmusu/	→ [tmusu]	[i-k-tumsu:]

KK-C: choosing among allomorphs

In the KK-C level of the hierarchy, the UR of a morpheme is always based on some surface allomorph, but this allomorph does not need to come from the same paradigm slot for every morpheme. For Maga, this allows for grammars like (14), where some words have leveled towards the stem allomorph while others have leveled towards the negative allomorph.

- (14) Grammar C: leveling towards either the stem or negative allomorph.

UR	STEM	NEGATIVE	
/ɖmari/	→ [ɖmari]	[i-k-ɖmari]	(base = stem)
/timsu/	→ [timsu]	[i-k-timsu:]	(base = negative)

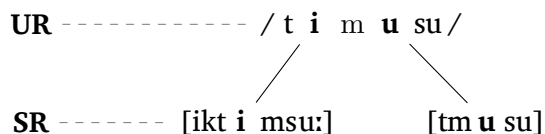
KK-D: segmentally composite URs

On this level of the hierarchy, every segment of the UR must be based on some surface allomorph, but not necessarily the same one. A typical analysis of the Maga syncope pattern would require this level of abstraction. This is because, at least for a subset of roots, no single surface allomorph has information about all the underlying vowels. This is schematized in (15) for the word ‘salt’

²Both allomorphs also surface in other morphological contexts

/timusu/; the negative allomorph shows that V1 is /i/, while the stem allomorph shows that the V2 is /u/. Crucially, however, no single allomorph can tell us about both V1 and V2. If the KK-D level of the hierarchy is learnable, then we would predict the syncope pattern to be stable over time, with few to no exceptions.

(15) *Composite UR of ‘salt’*



KK-E: featurally composite URs

Under KK-E, the only restriction on each underlying segment /S/ is that it must have feature values that occur in some surface allomorph, but not necessarily the same allomorph. Consequently, URs are able to contain segments never present in the surface allomorphs (as long as every feature occurs in some allomorph). Kenstowicz and Kisseberth (1977) give the example of English /ng/ as an analysis that requires KK-E levels of abstraction. Specifically, if we take the underlying form of *sing* to be /sing/, with surface [ŋ] derived from /ng/, the /ng/ sequence is abstract in the sense that neither segment surfaces in any allomorphs (cf. *sing* [sɪŋ], *singer* [sɪŋə]). This level of the hierarchy does not directly concern the current paper, since the Maga pattern does not require a higher level of abstraction than KK-D.

3.2 Predictions of opacity-avoidance vs. abstractness-avoidance

As discussed in Sections 2.2 and 2.5, rhythmic syncope is opaque, as deletion must occur after the building of the metrical structure which conditions it. In most cases, rhythmic syncope is not analyzable using parallel frameworks like Classical OT. This is especially so for a language like Maga, where surface stress and syncope are inconsistent. The opaque nature of rhythmic syncope could make it unstable and prone to restructuring over time; this is in fact what Bowers (2012) proposes has happened in Nishnaabemwin. The consequence is that if restructuring has occurred in Maga, it could be attributed to either abstractness avoidance or the avoidance of opaque syncope processes.

However, the syncope pattern could be restructured in a way which renders it transparent, but still abstract in the sense of requiring composite URs. The example in (16) shows a hypothetical Maga grammar, where vowel reduction to [ə] has replaced syncope. In this new grammar, vowel reduction is driven by a constraint (or rule) that is surface satisfied (e.g. “only [ə] is licensed in metrically weak positions”). This new grammar would be more faithful to the old Maga pattern, as it maintains all the underlying vowel qualities. However, because no surface allomorph tells

us about all the underlying stem vowels, a morphophonemic analysis of this “new Maga” would still require composite URs.

(16) *Hypothetical Maga: composite URs without opaque syncope*

UR	STEM	NEGATIVE
/ɖamari/	→ [ɖə́mári]	[i-k-ɖamə́rí:]
/timusu/	→ [tə́músu]	[i-k-timə́sú:]

In sum, if reanalysis has occurred in a way that makes Maga less abstract, due to some combination of leveling towards surface allomorphs, this cannot be explained purely as opacity-avoidance.

3.3 Degree of abstraction required in the Maga lexicon

As a result of the Maga syncope pattern, a subset of words require positing segmentally composite URs (KK-E). For other words, however, the UR is still predictable from one or both surface allomorphs. Table (17) provides examples of paradigms that vary in their degree of abstractness.

For example, roots that are underlying /V1 CV2/ (i.e. disyllabic with an initial vowel) can be accounted for using KK-A; there are no $\emptyset \sim V$ alternations because syncope prevents V1 from deleting in the stem, while in the negative form, neither vowel is in targeted positions for syncope. Some roots are predictable from one surface allomorph, so would require positing URs of levels KK-B/C. For example, roots of the shape /CV1 CV2/ never undergo syncope in the negative allomorph, so are fully predictable from this allomorph. Likewise, roots of the shape /V1 CV2 CV3/ are fully predictable from the stem allomorph. In general, blocking of syncope from vowel hiatus and/or geminates can also result in surface allomorphs that contain information about all the underlying vowels.

(17) *Examples of UR shapes that vary in their required degree of abstractness*

UR shape	Predictable allomorph	KK level	UR	stem	negative	gloss
/V1 CV2/	both	A	/abu/	ábu	i-abú:	‘ash’
/CV1 CV2/	negative	B/C	/tupe/	tpe:	i-k-tupé:	‘gourd dipper’
/V1 CV2 CV3/	stem	B/C	/amici/	amíci	i-k-amcí:	‘tree root’
/CV1 CV2 CV3/	neither	D	/dukaci/	dkáci	i-k-dukcí:	‘mud’

Table (18) summarizes the degree of abstractness required in the Maga corpus, *if* the syncope pattern is completely regular (i.e. assuming no reanalysis has occurred). If this corpus is taken to reflect the general distributions of the Maga lexicon, we see that the majority of the lexicon requires positing abstract URs of some kind. Only 8% of the corpus can be accounted for under KK-A. On the other end of the spectrum, over 60% of words require a composite UR, and can only be accounted for under KK-D.

In the rest of the lexicon (~31%), the UR is predictable from either the stem or negative allomorph. Under KK-B (isolation allomorph only), paradigms are expected to level towards the stem form. However, under KK-B” (most informative allomorph only), we predict leveling towards the negative allomorph, since it the predictable allomorph for a larger proportion of the lexicon. Recall that KK-C permits leveling in both directions; if this leveling is driven by informativeness, then we can predict leveling towards the stem allomorph for words of type (18ii), but leveling towards the negative allomorph for words of type (18iii).

(18) *Summary of UR types required in the lexicon (assuming regular application of syncope)*

	Predictable allomorph	KK-level	N	P
i.	both	A	84	0.08
ii.	stem	B/C	96	0.09
iii.	negative	B/C	230	0.22
iv.	neither	D	639	0.61

3.4 Zooming out: restructuring in Maga

Assuming regular application of syncope, over half of the Maga lexicon would require composite URs. However, it turns out that in total, about 24% (250/1049) of stem forms and 23% (185/790) of negative forms fail to undergo the regular syncope pattern. If we just look at the subset of 790 forms where the full stem/negative paradigm is available, 40% (314/790) paradigms have at least one member that does not undergo syncope in the expected positions. This irregularity suggests that many Maga paradigms have been restructured

In one subset of paradigms, although the regular syncope pattern has been maintained, URs have been reanalyzed in such a way that makes the negative allomorph more predictable from the stem allomorph; these cases are discussed in Section 4. In Section 5, I discuss cases where paradigms have leveled towards either the stem or negative allomorph. The data I use is the same corpus of 1049 words above; note that for some of these words, the negative allomorph is not available. In total, a subset of 790 words are listed with the stem/negative paradigm; for these, we can infer the direction of restructuring.

Note that throughout the rest of this paper, unless otherwise noted, I use ‘UR’ to refer to the hypothetical UR that would be posited under a traditional morphophonemic analysis, where representations of the type KK-D are permitted. The actual representation that speakers learn, as inferred by directions of restructuring, may not exactly match these URs.

4 Vowel-matching alternations

In the Maga Rukai lexicon, vowels in alternating positions tend to match each other; examples of this are given in (19), where the vowels that undergo syncope (in one of the paradigm slots) are underlined. In fact, in Hsin’s corpus, of the 322 paradigms that still require a composite UR, 61% (n = 196/322) exhibit this vowel-matching pattern. Crucially, if **vowel-matching** is a strong enough tendency in the Maga lexicon, it would actually render paradigms with vowel alternations predictable from just one surface allomorph.

(19) *Vowel matching examples*

UR	STEM	NEGATIVE	GLOSS
/s <u>i</u> mitu/	s <u>m</u> itu	i-k-s <u>i</u> mtu:	‘lips’
/g <u>o</u> romasi/	g <u>r</u> omasi	i-k-g <u>o</u> rmasi:	‘centipede’
/l <u>a</u> m <u>a</u> t <u>a</u> ki/	l <u>m</u> atki	i-k-l <u>a</u> mt <u>a</u> ki:	‘leech’

For example, given a paradigm like [smitu~i-k-simtu:], assuming vowel-matching is predictable, we could derive the negative allomorph from the stem, as schematized in (20). For illustrative purposes, I show a rule-based derivation involving a combination of vowel copying and deletion, but the mapping from the stem to negative allomorphs could happen in other ways. Additionally, though I show a derivation from the stem to negative allomorph, the input could just as well be the negative allomorph. What matters is that vowel-matching renders one allomorph predictable from the other, so that the paradigm no longer requires a composite UR. Instead, forms will be predictable from one surface allomorphs, in line with predictions of KK-B/C.

(20) *Deriving [i-k-simtu:] from a stem base*

Input	smitu (= stem allomorph)
Morphology	i-k-smitu:
Vowel copying	i-k-s <u>i</u> mitu:
Vowel deletion	i-k-simtu:
Stress Assignment	i-k-simtú:
Output	i-k-simtú:

In the rest of this section, I show that the proportion of vowel-matching forms in Maga Rukai has increased over time. In other words, while vowel-matching might have already been present as a tendency in an earlier stage of the language, speakers have reanalyzed Maga forms in a way that has strengthened this tendency. As a result, a large proportion of Maga paradigms that would otherwise require composite URs are now predictable from surface allomorphs. To infer the direction of vowel changes over time, I compare the distribution of vowel-matching forms in Proto-Rukai and Maga.

4.1 Proto-Rukai data

Data on Proto-Rukai comes from Li's (1977) list of 532 protoforms and their cognates in the six Rukai dialects. Table (21) lists Proto-Rukai vowels and their regular reflexes in Maga; in the rest of this section, for comparability with the Maga data, Proto-Rukai vowels are presented and organized by their reflexes in Maga.

In general, there is a straightforward correspondence between Proto-Rukai and Maga, with the exception that mid vowels *o and *ə have raised to Maga [u] and [i]. Proto-Rukai diphthongs like *aw and *ay are reflected as monophthongs in Maga; Maga reflexes of Proto-Rukai diphthongs can be short or long, but notably resist syncope.

The one wrinkle is that *i, *o, *ə are reflected as mid vowels [e], [o], and [ə] if they follow a *deleted* low vowel (*a). This lowering process generally does not apply if [a] has not been deleted (e.g. *aŋato > aŋatu 'tree'), though there are a few exceptions where vowels adjacent to *a exceptionally lower (e.g. *aθo > aθə: 'dog'). The result is that in Maga, we observe vowel-lowering alternations that interact with syncope (see Section 2.4 for more details).

(21) Vowel correspondences, Proto-Rukai and Maga

PROTO-RUKAI	MAGA	EXAMPLE
*a	a	*kav <u>a</u> > kv <u>a</u> : 'hoe'
*i	i	*rig <u>i</u> > rg <u>i</u> : 'horse'
*o, oa	u	*tim <u>oso</u> > tm <u>úsu</u> 'salt'; *oalopo > ulúpu 'hunt'
*ə	i	*pəl <u>əŋə</u> > pl <u>íŋi</u> 'ghost'
*(a...)i	e	*cak <u>i</u> > ck <u>é</u> : 'excrement'
*(a...)o	o	*val <u>o</u> > vl <u>ú</u> 'bee'
*(a...)ə	ə	*bal <u>ə</u> bal <u>ə</u> > bl <u>ə</u> bl <u>ə</u> 'bamboo'
*ay, *ai	e, e:	*acil <u>ay</u> > ac <u>íle</u> 'water'; *pag <u>ay</u> > pg <u>é</u> : 'rice plant'
*aw, *ao	o, o:	*maram <u>aw</u> > marámo 'same'; *ko <u>aw</u> > ku <u>o</u> : 'eagle'

Note that other dialects of Rukai have largely maintained vowel contrasts, undergoing much less vowel reduction or deletion than Maga. As a result, Proto-Rukai reconstructions can be reliably used to infer what Maga vowels should look like, if no restructuring has occurred.

4.2 Degree of vowel-matching in Proto-Rukai and Maga

To approximate the degree of vowel matching in the Maga Rukai lexicon, I focus on just underlying V1 and V2 (e.g. /u/ and /a/ in /dukaci/ 'mud'). This dataset, though slightly pared down, should still be a good reflection of vowel-matching patterns across Maga, since most stems have three or fewer underlying vowels (70%, N = 735/1049), and for these words, only V1 and V2 are alternating.

Additionally, forms are categorized as having either **composite** or **non-composite** URs. For the Proto-Rukai forms, where we do not have direct access to the negative allomorphs, the nature

of the UR is inferred from word shape. For example, a protoform with the shape CVCVCV, such as *timoso ‘salt’, should correspond to a Maga paradigm that requires composite URs (in fact it does, cf. [tmúsu]~[i-k-timsú:]). Additionally, some forms in Maga should require composite URs, but have leveled towards one of the surface allomorphs. For example, the verb ‘to tattoo’ has the paradigm [u-ckícki]~[i-ckickí:]; the negative allomorph does not show the expected syncope pattern (cf. *[i-cikciki:]), suggesting that it has leveled towards the stem allomorph. Leveling is discussed more in Section 5; for now, we categorize the leveled paradigms as having non-composite URs.

Going into the results, Fig. 1 shows the proportion of vowel-matching forms in Proto-Rukai vs. Maga Rukai. We see that there has generally been an increase in the proportion of vowel-matching forms between Proto-Rukai and Maga. Notably, this increase is much larger for the composite UR forms. This suggests that restructuring of UR vowels has occurred specifically in the paradigms where vowel-matching would result in less abstract URs.

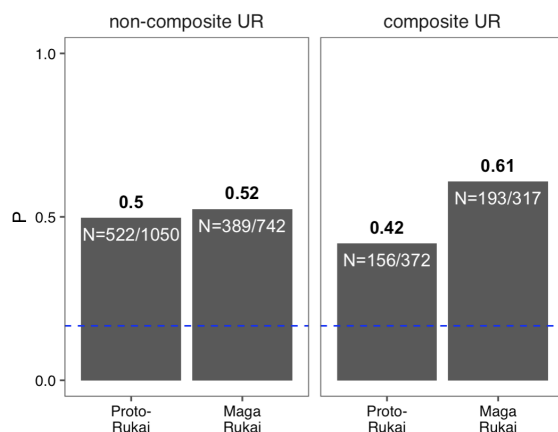


Figure 1: Proportion of vowel-matching (between V1 and V2) in Proto-Rukai vs. Maga Rukai. Dashed line indicates the chance-level rate of vowel-matching alternations.

Fig. 2 shows the same results, but separated by identity of underlying V1. For example, the subfigure for /u/ represents forms like /dukaci/ ‘mud’ ([dkáci]~[i-k-dukci:], non-vowel-matching) and /surugu/ ‘liver’ ([srúgu]~[i-k-surgú:], vowel-matching). Forms where V1 is /ə/ are omitted because there just two of them. Interestingly, based on Fig. 2, rates of vowel-matching has increased even for vowels that were not historically matching. For example, in Proto-Rukai, there was no vowel-matching with /e, o/³, but vowel-matching involving the mid vowels is observed in Maga. These results suggest that speakers learned a general vowel-matching pattern, extending it past environments observed in the lexicon. Similar trends are observed if we group forms by identity of underlying V2; this figure is given in the Appendix.

Fig. 3 shows the same data as Fig. 2 above, but visualized in terms of the *change* in proportion of vowel-matching between Proto-Rukai and Maga Rukai. A positive value on the y-axis indicates that vowel-matching increased over time, while a negative value indicates a decrease in vowel-

³This is expected given that /e, o/ developed either from diphthongs or lowering of high vowels after deleted *a.

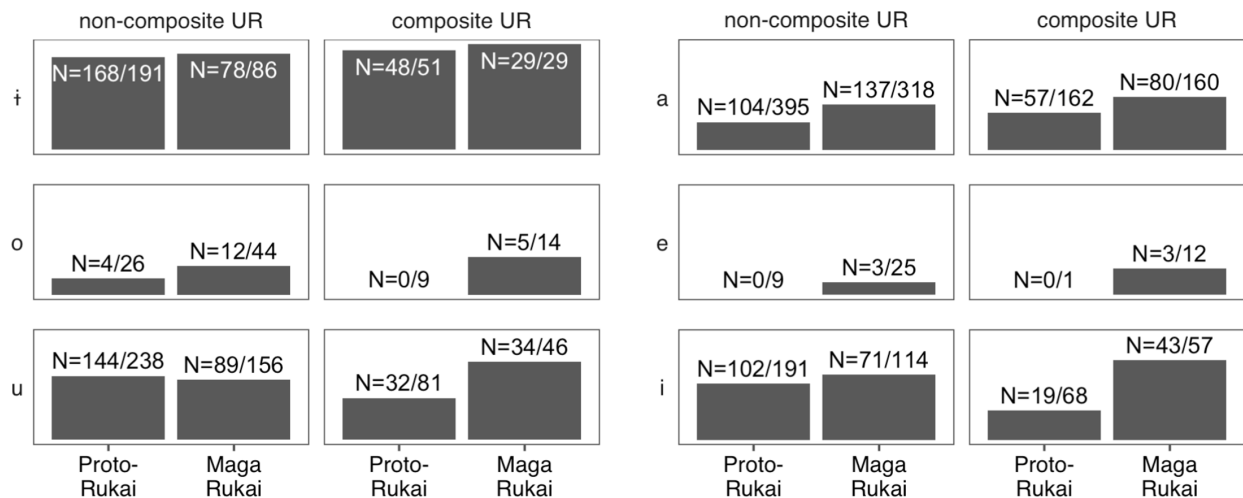


Figure 2: Proportion of vowel-matching in Proto-Rukai vs. Maga Rukai, by /V1/

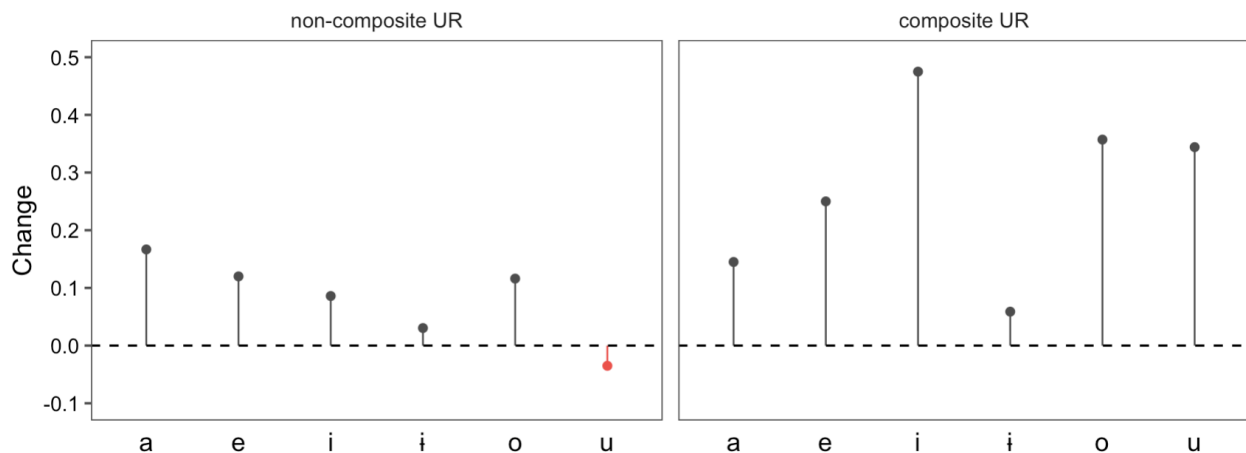


Figure 3: Change in proportion of vowel matching (Maga – Proto-Rukai)

matching. This figure makes it evident that across almost every vowel and UR type, there has been increased rates of vowel-matching. In the one exception (V1 = /u/ in non-composite URs), the decrease in vowel-matching is very small. Importantly, the increase in vowel-matching is much larger for composite URs, again supporting an analysis where vowels have been restructured specifically where doing so can remove composite URs.

Finally, it should also be noted that vowel-matching appears to already have been present as a statistical tendency in Proto-Rukai. Maga has 6-7 phonemic vowels (depending on whether /ə/ is treated as phonemic), meaning that if vowel-matching occurred at chance, it should occur about 16.7% (1/6) of the time; this value is shown for reference in Fig. 1. Even in Proto-Rukai, vowel-matching happened at well above this rate. This is corroborated by a Monte Carlo test of significance, where I find that vowel-matching was already overrepresented for most vowels in Proto-Rukai (all vowels excluding /e/ and /o/), and that this tendency was further exaggerated

in Maga Rukai. Details of the Monte Carlo test are given in Appendix 2. These results suggest that when faced with resolving the ambiguity caused by syncope, speakers utilized a tendency that was already present in the lexicon, causing it to become stronger over time.

4.3 Direct examples of vowel-matching changes

In this section, I look at the subset of Maga Rukai paradigms that have known Proto-Rukai forms. These forms can provide more direct evidence for vowel changes, and also give us insight into the direction of restructuring. The examples in (22) show the types of changes we are interested in, where forms have been restructured to result in more vowel-matching. In these examples, the changed vowel is underlined.

(22) *Examples of vowel-matching changes*

Proto-Rukai	expected UR	actual UR	paradigm	gloss
*samito	/s <u>a</u> mitu/	/s <u>i</u> mitu/	smítu~i-k-sím <u>tú</u> :	‘lips’
*tobakə	/tu <u>b</u> aki/	/ta <u>b</u> aki/	tbáki~i-k-ta <u>b</u> kí:	‘shell’
*sinaw	/sin <u>o</u> sino/	/sin <u>i</u> sino/	u-sn <u>i</u> snó:~i-sinsinó:	‘wash (clothes)’

Overall, there were 56 forms where Proto-Rukai and Maga Rukai vowels mismatched. Table (23) summarizes the types of changes observed. The majority of changes (~70%) have resulted in increased vowel-matching, consistent with the findings of the previous section. Of the remaining changes, most neither increase nor decrease the rate of vowel-matching; for example, Proto-Rukai **ma-riday* should correspond to Maga [ma-rdé:]~[i-k-rdé:], but the observed paradigm is instead [ma-rdé:]~[i-k-rudé:].

(23) *Summary of vowel changes*

CHANGE RESULTS IN...	N	P	EXAMPLE
MORE vowel-matching	39	0.70	*s <u>a</u> mito → /s <u>i</u> mitu/ ‘lips’
LESS vowel-matching	5	0.09	*v <u>i</u> li → /v <u>u</u> li/ ‘leech’
Neither	12	0.21	*ma-r <u>i</u> day → /r <u>u</u> dé:/ ‘fast’

For a few words (N=5), vowel-matching forms were changed to be *non*-matching. Of these cases, two could potentially be explained as the result of speakers extending a vowel-lowering alternation, where high vowels lower following a deleted /a/; this pattern is discussed in Sections 2.4 and 4.1. For example, Proto-Rukai **ma-pilay* ‘tired’ should correspond to Maga [ma-plé:]~[i-k-pilé:] (UR = /pile/), but instead we observe [ma-plé:]~[i-k-parí:] (UR = /pari/). This could be because speakers, when given the stem allomorph [ma-ple:], interpreted the final [e] as the result of /i/-lowering after a deleted /a/.

Direct comparison of Proto-Rukai and Maga forms can also give us insight into whether vowel-matching changes were based on the stem or negative allomorph. Given a protoform like **samitu*,

where V1 and V2 mismatch, there are two ways it could be reanalyzed as vowel-matching, schematized in (24). One option would result in the paradigm [smitu]~[i-k-simtú:], with the UR /simitu/ (in this example, and examples below, the restructured allomorph is shaded in). This is a case of restructuring based on the stem allomorph, as speakers are extending the vowel that is present in the stem allomorph to the negative allomorph. The opposite scenario, where speakers extend the vowel present in the negative allomorph to the stem allomorph, would result in the paradigm [smatu]~[ik-samtú:].

(24) *Hypothetical directions of vowel-matching changes for *samito*

	UR	stem	negative
a. Expected Maga form	/samitu/	smitu	i-k-samtú:
b. Restructured from stem base	/simitu/	smítu	i-k-simtú:
c. Restructured from negative base	/samatu/	smatu	i-k-samtú:

Table (25) below summarizes the observed directions of vowel changes, with representative examples for each direction. Overall, restructuring has largely been based on the stem allomorph; this accounts for 45/56 (≈80%) of all vowel changes observed. A much smaller number of changes (N=6, 0.11%) were based on the negative allomorph. Finally, for few words, both V1 and V2 have changed, making it unclear whether restructuring was from the stem or negative allomorph.

(25) *Directions of vowel-matching reanalysis*

BASE	N (P)	PROTO-RUKAI	EXPECTED UR	ACTUAL UR
Stem	45 (0.80)	* <u>s</u> amito	/s <u>a</u> mitu/	/s <u>i</u> mitu/ (smitu~ik-s <u>i</u> mtú:)
Negative	6 (0.11)	*sin <u>a</u> w	/sin <u>o</u> sino/	/sin <u>i</u> sino/ (u-sn <u>i</u> sno~i-sinsino:)
Unclear	5 (0.09)	*tol <u>a</u> rə	/tul <u>a</u> ri/	/tal <u>u</u> ra/ (tl <u>u</u> ra~i-k-tal <u>a</u> ra:)

4.4 Interim summary: vowel-matching

In comparing Proto-Rukai and Maga Rukai, we find that an earlier stage of Rukai (represented by Proto-Rukai) already had a statistical tendency towards vowel-matching. In Maga Rukai, specifically for stems that require composite URs, forms were restructured in a way that increased the degree of vowel-matching. In other words, speakers appear to have restructured forms in a way that reduced the need for composite URs. This vowel-matching pattern is particularly interesting in that while it does not remove the opaque syncope alternations, it does remove the need for composite URs. This suggests that vowel-matching changes were really motivated by the avoidance of abstract URs, rather than by the avoidance of opaque patterns.

In comparing the directions of restructuring, I also find that most vowel changes have been based on the stem allomorph; in other words, the vowel quality present in the stem allomorph was extended to the negative allomorph. Restructuring in the other direction (negative→stem) was attested, but much less frequent. Taken together, these results suggest that speakers can learn

URs of type KK-C (choosing among allomorphs), there is a preference to learn URs that match the unsuffixed stem allomorph, in line with predictions of KK-B.

One variant of KK-B, well-substantiated in the literature, is Albright’s single surface-base hypothesis, which predicts restructuring towards the overall most informative allomorph. Recall from Section 3.3 that an informativeness-based account would actually predict restructuring towards the negative allomorph, contradicting the results found here. This is because more forms are predictable from the negative allomorph ($n = 230$, 22% of the corpus) than from the stem allomorph ($n = 96$, 9%). In Section X below, I discuss some potential reasons for why results do not line up with the predictions of the single surface-base hypothesis.

5 Paradigm leveling

In this section, I discuss cases where paradigms appear to have leveled towards either the stem or negative allomorph, removing $\emptyset \sim V$ alternations. Rather than direct comparison with Proto-Rukai (as was done in Section 4 above, direction of leveling is inferred from the syncope pattern found in Modern Rukai. For example, given the word /simitu/ ‘lips’, there are three potential syncope patterns (not leveled, leveled to stem, leveled to negative) given in (26); where leveling has occurred, the vowel that should have deleted is underlined, and vowels that were exceptionally deleted are shown as ‘ \emptyset ’. If we observe a paradigm like (26, where the negative allomorph does not show the regular syncope pattern, and instead has the same syncope pattern as the stem allomorph, this implies that leveling in the direction of stem \rightarrow negative has occurred.

(26) *Types of leveling for ‘lips’ /simitu/*

PATTERN	PARADIGM
a. Not leveled (regular syncope)	[smítu] \sim [i-k-símtu:]
b. Leveled to stem	[smitu] \sim [i-k-s \emptyset mítu:]
c. Leveled to negative	[sím \emptyset tu] \sim [i-k-smitu:]

Recall from Section 3.3 that Maga paradigms can be grouped in terms of whether they are predictable from one (or both) surface allomorphs. For example, /tupe/ [tpé:] \sim [i-k-tupé:] ‘gourd dipper’ is predictable from the negative allomorph, which contains both underlying vowels. With this in mind, we can categorize paradigms along two dimensions: the original base (i.e. predictable allomorph prior to leveling), and the direction of leveling. Table (27) provides examples of each type of leveling; the UR listed in each example is the hypothetical UR that we would expect, if the syncope pattern were regular.

For example, /udali/ ‘rain’ represents a case where the base is the stem, and leveling towards the stem has happened. The stem [událi] is a predictable base in that all three vowels surface. The ‘regular’ negative form should be [i-udlí:], but instead, it has leveled towards the stem and is now [i-udalí:]. On the other hand, /idipi/ ‘quench’ is also predictable from the stem allomorph,

but has instead leveled towards the negative allomorph. As a result, the isolation stem surfaces as [ídpi], with the syncope pattern of the negative form.

(27) *Examples of leveling in Maga*

Base	Level to	UR	Expected paradigm	Actual paradigm	Gloss
stem	stem	/ud <u>ali</u> /	událi~i-udlí:	událi~i-udalí:	‘rain’
	negative	/id <u>i</u> pi/	ídípi~i-idpí:	ídpi~i-idpí:	‘quench’
negative	stem	/k <u>u</u> cia/	u-kcyá:~i-kucyá:	u-kcyá:~i-kcyá: ^a	‘scissor’
	negative	/te <u>θ</u> o/	tóó:~ik-teóó:	téθo~i-k-teóó:	‘turnip’
neither	stem	/b <u>u</u> lavani/ ^b	blávni~ik-bulvani	blávni~ik-blavni:	‘silver’
	negative	/θil <u>i</u> bi/	u-θlíbi~i-θilbí:	u-θlíbi~i-θilbí:	‘fly’

^aThe negative of ‘scissor’ does actually have a regular variant [i-kucyá:]

^bThe first vowel of this (hypothetical, pre-leveling) UR is inferred from Proto-Rukai *bolavana.

For a subset of forms that would otherwise require composite URs (i.e. neither allomorph is a fully informative base), one or both allomorphs fail to undergo syncope at all. As a result, the UR is predictable from at least one allomorph. Examples of this are given in (28), and the vowel(s) that should have deleted are underlined.

(28) *Forms where syncope exceptionally fails to apply*

Form w/o syncope	Paradigm	Gloss
Stem	ma-rimúru~i-k-rimru:	‘forget’
Negative	plíŋi~ik-pilíŋí:	‘ghost’
Both	l u bili~i-k-lub i li:	‘mouth harp’

Table (29) summarizes the proportion of forms that have been leveled; the row titled ‘no syncope’ shows cases where syncope exceptionally failed to apply in either one or both allomorphs. Examples from (27) are given again; where leveling has occurred, the vowel that should have deleted is underlined, and vowels that were exceptionally deleted are shown as ‘∅’.

First, as expected, non-alternating forms (i.e. where the original base is both allomorphs) have not undergone any change. For forms that were predictable from one allomorph, there has generally been leveling towards the predictable allomorph. For example, paradigms that were predictable from the negative allomorph almost always leveled towards the negative allomorph, if they were leveled (29f; n=69). In contrast, there is only one example of leveling towards the stem allomorph (29e).

For the composite UR forms (i.e. ones predictable from neither allomorph), there has generally been less leveling (N=61, 13%).⁴ Where leveling has occurred, it has happened towards both stem and negative allomorph, with a bias towards the negative allomorphs.

⁴Note that this is also the subset of forms that underwent vowel-matching restructuring.

(29) *Patterns of leveling*

	Orig. base	Leveled...	N	P	Example	
a.	both	not leveled	62	1	ábu~i-abú:	‘ash’
b.	stem	towards stem	32	0.37	událi~i-udalí:	‘rain’
c.		towards negative	2	0.02	ídØpi~i-idpí:	‘quench’
d.		not leveled	51	0.61	amíci~i-k-amcí:	‘tree root’
e.	negative	towards stem	1	0.01	u-kcyá:~i-kØcyá:	‘scissor’
f.		towards negative	69	0.42	téθo~i-k-teθó:	‘turnip’
g.		not leveled	94	0.57	tpe:~i-k-tupé:	‘gourd dipper’
h.	neither	towards stem	34	0.07	blávni~i-k-bØlavni:	‘silver’
i.		towards negative	27	0.06	u-θflØbi~i-θilbí:	‘fly’
j.		no syncope	97	0.20	lubili~i-k-lubíli:	‘mouth harp’
k.		not leveled	322	0.67	qmari~i-k-qamri:	‘moon’

In sum, forms show leveling towards both the stem and negative allomorphs, supporting KK-C levels of representation (i.e. choosing among allomorphs). Additionally, leveling appears to be driven in part by informativeness; forms that are predictable from one of the surface allomorphs have almost exceptionlessly leveled towards that allomorph. This is in line with the existing literature (see review in Section 3) on base-driven leveling. For the composite UR cases, there is a slight preference for leveling towards the stem allomorph. Notably, the composite UR forms have also undergone less leveling ($P = 0.13$, or 0.33 if non-syncope are counted) than other forms ($P = 0.39$ for stem-base paradigms; $P = 0.33$ for negative-base paradigms).

5.1 Alternative accounts to leveling

5.1.1 Syncope as the innovative variant

Chen (2008) puts forth the idea that syncope is a relatively innovative and recent development, and that paradigms which fail to undergo syncope actually reflect an earlier stage of Maga. Under this account, a case like (29b) could be treated as failure of the negative allomorph to undergo syncope, rather than leveling of the paradigm towards the stem allomorph. Similarly, cases like (29f) can be treated as failure of the stem allomorph to undergo syncope. More generally, if we adopt this idea, then around 75% of the leveling cases in (29) can be re-interpreted as failure to undergo syncope, reflecting a more conservative stage of Maga.

This account is less satisfactory for several reasons. First, it does not explain all cases of leveling, as around 25% of cases discussed in (29) cannot be explained as the failure of stems to undergo syncope. In particular, in cases like (29h,i), though the paradigm has been leveled, both allomorphs still undergo syncope.

Additionally, under this account, there is no straightforward explanation for why syncope fails to apply at a much higher rate specifically where there is a predictable base in the paradigm. That

is, why should syncope fail to apply for negative allomorphs specifically when the stem happens to be a predictable base (29a), and for the stem allomorph specifically when the negative allomorph is a predictable base (29f)? Under a leveling analysis, these patterns straightforwardly follow if leveling is towards the most informative allomorph.

Finally, comparison of more recent data and Hsin's (2000) corpus suggests that leveling (or failure to undergo syncope) is the innovative variant. The *Austronesian indigenous words and narrations*, offered by the Indigenous Languages Research and Development Center (ILRDC), is an online corpus made for educational purposes (2022). Henceforth I will refer to this source as the ILRDC corpus. The data were recorded from 2020-2022, almost 20 years after Hsin's corpus, in consultation with a younger generation of speakers. In other words, the ILRDC corpus should reflect a relatively more recent stage of Maga. If we compare the two data sources, there are many examples where syncope applied in the Hsin corpus, but fails to apply to the ILRDC dictionary.

Examples of such mismatches are given in (30). The ILRDC corpus does not present words in morphological paradigms, and does not contain the negative allomorph of most words. As such, I focus on comparing the stem forms, but the negative allomorph from Hsin (2000) is also given in parentheses for reference. In examples (30a-c), the ILRDC stems don't undergo the expected vowel syncope, and also have the same vowels as the negative allomorph given in Hsin (2000). In (30d-e), the ILRDC stems actually show the syncope pattern of the negative allomorph, suggesting that there has been leveling towards the negative allomorph.

(30) *Comparisons between Hsin (2000) and the ILRDC web corpus*

	ILRDC CORPUS	HSIN (2000)	GLOSS
a.	<u>s</u> uro:	sró: (cf. i-k-suro:)	'grass'
b.	u-l <u>u</u> é:	u-lhé: (cf. i-lué:)	'siwm'
c.	u-t <u>u</u> ta	u-ttá: (cf. i-tuta:)	'vomit'
d.	c <u>i</u> ŋØli	u-cŋili (cf. i-ciŋli:)	'look'
e.	d <u>i</u> kØsi	u-dkísi (cf. i-diksi:)	'squeeze'

5.1.2 Historical explanations for failure to undergo syncope

Proto-Rukai has long vowels which resisted syncope. As a result, apparent exceptions to the syncope pattern, as described in (29), may result from historic long vowels, rather than from base-driven leveling. In this section, I show that historic long vowels only account for a small proportion of exceptions to syncope, and cannot fully explain my results.

There are two sources of long vowels that resisted syncope in Maga. First, Proto-Rukai had diphthongs which monophthongized in Maga (e.g. *likolaw → rkúlo); more examples are given in Section 4.1. In addition, Proto-Rukai *ʔ was uniformly deleted in Maga, and as a result *VʔV sequences became either V or bimoraic V: (e.g. *abonoʔo → abúŋu 'ant'); presumably, deletion of *ʔ resulted in an intermediate stage of Maga with VV vowels. Representative examples are given

in (31); words like [dani] ‘house’ don’t undergo the expected syncope, but this falls out from the fact that [a] is derived from *aʔa.

Of the Maga Rukai words where we have available protoforms, historic long vowels block syncope in just 39/109 stems, and 4/6 negative forms.⁵ In other words, most cases of exceptional non-syncope *cannot* be traced back to historic long vowels; examples of forms like this, where a non-syncope vowel corresponds to a historic short vowel, are given in (32). In total, only 39% (N = 43/115, P = 0.37) of non-syncope corresponds to Proto-Rukai long vowels.

(31) *Maga reflexes of Proto-Rukai diphthongs and *VʔV*

Proto-Rukai	Maga	gloss
*daʔanə	d <u>a</u> ni (dni:)	‘house’
*payso	pe <u>s</u> u (psu:)	‘money’
*bəʔəkə	b <u>i</u> ki (bki:)	‘pig’

(32) *Examples of Proto-Rukai short vowels that fail to undergo regular syncope*

Proto-Rukai	Maga	Expected Maga	Gloss
*d <u>o</u> koʔo	du <u>k</u> u	dØku:	‘grow (plant)’
*p <u>a</u> kə	pi <u>k</u> i	pØki:	‘house lizard’
*ma-rimoro	ma-r <u>i</u> múru	ma-rØmuru	‘forget’

5.2 Picture so far

Overall, examination of Maga paradigms (and comparison of these paradigms with Proto-Rukai) suggests that two types of restructuring have occurred, both of which have reduced the need for composite URs. First, leveling has removed Ø~V alternations in many paradigms, making them amenable to KK-A levels of analysis. Table (33) compares the degree of abstractness required if we assume regular application of syncope (as discussed in Section 3.3) against the degree of abstractness in Hsin’s corpus. If we take the regular syncope pattern to represent Maga pre-restructuring, then there has been an overall shift towards non-alternation, with a 24% increase in forms that are predictable from both allomorphs. Composite UR forms have seen the largest decrease (by 20%), supporting a view where learners are biased against learning composite URs (i.e. KK-D representations).

(33) *Comparison of UR abstractness before vs. after restructuring*

⁵This count does not include Proto-Rukai forms where the long vowel is not in a non-syncope position, e.g. *capəʔə→cpí:.

Predictable				
allomorph	KK-level	Before	After	Change
both	A	0.08	0.30	+0.24
stem	B/C	0.09	0.07	−0.02
negative	B/C	0.22	0.22	0
neither	D	0.61	0.41	−0.20

Aside from leveling, forms have also been restructured in a way that increased the rate of vowel-matching alternations. If we take vowel-matching forms to be predictable from just one allomorph, then an even smaller degree of abstractness is required in the lexicon. Table (34) shows the degree of abstraction required in the Maga lexicon if we account for vowel-matching; the row labeled 'either (vow matching)' represents paradigms that should need composite URs, but have alternating vowels that match. Now, only 16% of forms still require a composite UR.

(34) *Degree of abstractness required in Maga lexicon (assuming that vowel-matching forms do not require composite URs*

Predictable			
allomorph	KK-level	N	P
both	A	238	0.30
stem	B/C	57	0.07
negative	B/C	173	0.22
either (vow matching)	B/C	196	0.25
neither (no matching)	D	126	0.16

Overall, Hsin's corpus suggests that over time, paradigms have been restructured to result in less abstract URs. Additionally, the direction of restructuring suggests that when the two allomorphs are equally informative, there is a strong tendency to level towards the stem allomorph. When one allomorph is more informative than the other, then leveling tends to occur towards the more informative allomorph. This direction of leveling suggests that while URs are amenable to KK-C levels of representation (choosing from allomorphs), there is a bias towards learning URs that match the unaffixed allomorph (KK-B level of representation).

One issue that has not yet been addressed is the fact that the lexicon reflects two restructuring strategies: vowel-matching and leveling. Why should both strategies exist, and did they occur in parallel? It is possible that they reflect different stages of diachronic change in Maga, or that they reflect different pressures on paradigm structure. In the following section, I present data from a newer generation of speakers, whose strategies for paradigm restructuring can provide insight into these questions.

6 Restructuring in a newer generation of speakers

This section reports results from fieldwork with three Maga Rukai speakers. To preview the results, I find that these speakers have leveled paradigms to a greater degree than the speakers represented in Hsin’s corpus. This suggests that leveling is the more innovative strategy; in the following section, I tentatively propose a frequency-based account for the two diverging restructuring strategies.

6.1 Speaker background

Data collection was carried out by the author in Maolin District, Kaohsiung, Taiwan over the course of one week in June 2024. Three native speaker consultants, aged 54-61, were consulted. They represent a younger generation of speakers compared to the two consultants that Hsin (2000) worked with, who at the time of data collection (1999-2000) were both 74 years old.

Consultant details are given in (35). S1 reports speaking Maga regularly with elders (because of his involvement with the District Office, and with church duties). S2 and S3 speak Maga less regularly, and S3 is married to a Tona speaker; while Tona and Maga are mostly mutually intelligible, the two dialects have some vocabulary differences, and Tona has not undergone the vowel reduction/deletion processes that Maga has.

(35) *Consultant details*

SPEAKER	GENDER	AGE	NOTES
S1	M	61	speaks Maga regularly with elders
S2	M	56	
S3	F	54	is married to a Tona speaker

These speakers have one important difference from Hsin’s speakers in terms of their morphology: the negative prefix is no longer productively used with nouns. Instead, speakers express negation paraphrastically with the negation words [teɬa], [eni], and [ikmani]. For example, the negative of ‘mud’ is [ikmai dkaci], rather than [i-k-dukci].

6.2 Data collection

Consultants were asked to provide the citation form (unaffixed stem), negative form (i- prefix), and imperative form of (V- -a) of items from a wordlist. The imperative form has the same syncope pattern as the negative, and was included to check if both forms show a consistent syncope pattern. Items were presented in Chinese orthography using Microsoft PowerPoint; each word was given in its own slide, and accompanied by an illustrative photo.

The wordlist consisted of 60 verbs reported to have a regular syncope pattern in Hsin’s corpus. Words were restricted to verbs because, as stated above, To the extent possible, items were se-

lected to represent types of paradigms: ones predictable from the stem (n = 15), from the negative allomorph (n = 19), and from neither allomorph (n = 26).

6.3 Results: patterns of restructuring

(36) XX

SPEAKER	BASE	N	LEVELED.	TOTAL	P	LEVELED TO STEM
S1	stem	7		12	0.58	6 (0.86)
	negative	6		18	0.33	2 (0.33)
	neither	14		26	0.54	13 (0.93)
S2	stem	11		13	0.85	10 (0.91)
	negative	6		16	0.38	3 (0.50)
	neither	13		22	0.59	10 (0.77)
S3	stem	11		12	0.92	11 (0.92)
	negative	6		18	0.33	3 (0.50)
	neither	15		23	0.65	14 (0.93)

References

- Albright, Adam. A restricted model of ur discovery: Evidence from lakota. Paper given at the GLOW Phonology Workshop, Utrecht. <http://web.mit.edu/albright/www/papers/Albright-Lakhota.pdf>, (2002)a.
- Albright, Adam (2010). Base-driven leveling in Yiddish verb paradigms. *NLLT* 28(3). 475–537. doi: 10.1007/s11049-010-9107-z.
- Albright, Adam C. (2002)b. *The identification of bases in morphological paradigms*. PhD dissertation, University of California, Los Angeles.
- Beckman, Jill N (1998). *Positional faithfulness*. PhD Thesis, University of Massachusetts Amherst.
- Bowers, Dustin (2012). *Phonological restructuring in Odawa*. MA thesis, University of California, Los Angeles.
- Chen, Yin-Ling Christina (2008). On the Rhythmic Vowel Deletion in Maga Rukai. *Concentric: Studies in Linguistics* 34. 47–84.
- Chomsky, Noam and Morris Halle (1968). *The sound pattern of English*. Harper & Row.
- Dresher, B Elan and Aditi Lahiri (1991). The germanic foot: metrical coherence in old english. *Linguistic inquiry* 22(2). 251–286.

- Hawkins, W Neil (1950). Patterns of vowel loss in macushi (carib). *International Journal of American Linguistics* 16(2). 87–90.
- Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. University of Chicago Press.
- Hsin, Tien-Hsin (2000). *Aspects of Maga Rukai Phonology*. PhD Thesis, University of Connecticut.
- Hsin, Tien-Hsin (2003). The mid vowels of Maga Rukai and their implications. *Journal of East Asian Linguistics* 12(1). 59–82. Publisher: Springer.
- Indigenous Languages Research and Development Center. Austronesian indigenous words and narrations. <https://web.kloka.h.tw>, (2022). [Online; accessed 9-July-2024].
- Kager, René. Rhythmic vowel deletion in optimality theory. In Roca, Iggy, editor, *Derivations and constraints in phonology*, 463–500. Oxford University Press, (1997).
- Kaohsiung City Government Civil Affairs Bureau. Population census. <https://cabu.kcg.gov.tw/web/HRStatReports.aspx?appname=HRStatReports>, (2023).
- Kaplan, Max J. (2022). Stratal overgeneration is necessary: metrically incoherent syncope in southern pomo. *Phonology* 39(4). 597–639. doi: 10.1017/S0952675723000234.
- Kenstowicz, Michael J. and Charles W. Kisseberth (1977). *Topics in phonological theory*. Academic Press, New York. ISBN 978-0-12-405150-8.
- Kiparsky, Paul. Linguistic universals and linguistic change. In Bach, Emmon W. and Robert Thomas Harms, editors, *Universals in Linguistic Theory*, 170–202. Holt, Rinehart, and Winston, (1968).
- Kiparsky, Paul. Abstractness, opacity, and global rules. In Fujimura, Osamu, editor, *Three dimensions in linguistic theory*, Publications, 57–86. TEC, Tokyo, (1973). URL <https://books.google.com/books?id=YMaS0wAACAAJ>.
- Kiparsky, Paul. Lexical morphology and phonology. In Yang, I.-S., editor, *Linguistics in the morning calm*, 3–91. Hansin, (1982).
- Li, Paul Jen-Kuei (1975). Maga phonology: A preliminary report. *Bulletin of the Department of Archaeology and Anthropology, National Taiwan University* 37-38. 16–28.
- Li, Paul Jen-kuei (1977). The internal relationships of rukai. *Bulletin of the Institute of History and Philology* 48(1). 1–92.
- McCarthy, John J (2008). The serial interaction of stress and syncope. *Natural Language & Linguistic Theory* 26. 499–546.

Saillard, Claire. Is Maga accusative or ergative? evidence from case marking. In Tsao, Feng-fu and Mei-hui Tsai, editors, *Proceedings of the First International Symposium on Languages in Taiwan*. The Crane, (1995).

Valentine, Randy (2001). *Nishnaabemwin reference grammar*. University of Toronto Press.

Zeitoun, Elizabeth (1995). *Issues on Formosan linguistics (Problèmes de linguistique dans les langues aborigènes de Taiwan)*. PhD Thesis, University Paris 7.

A x

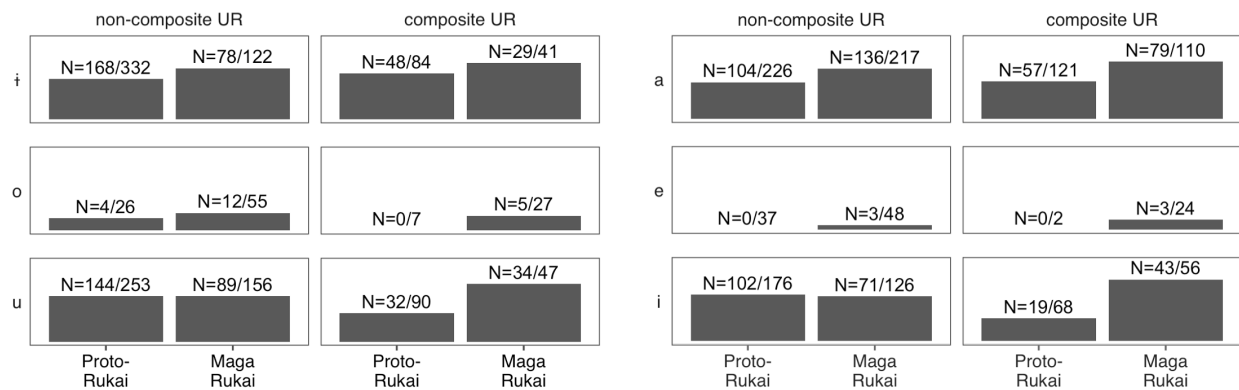


Figure 4: Proportion of vowel-matching in Proto-Rukai vs. Maga, by /V2/

B Monte Carlo test of significance: vowel-matching

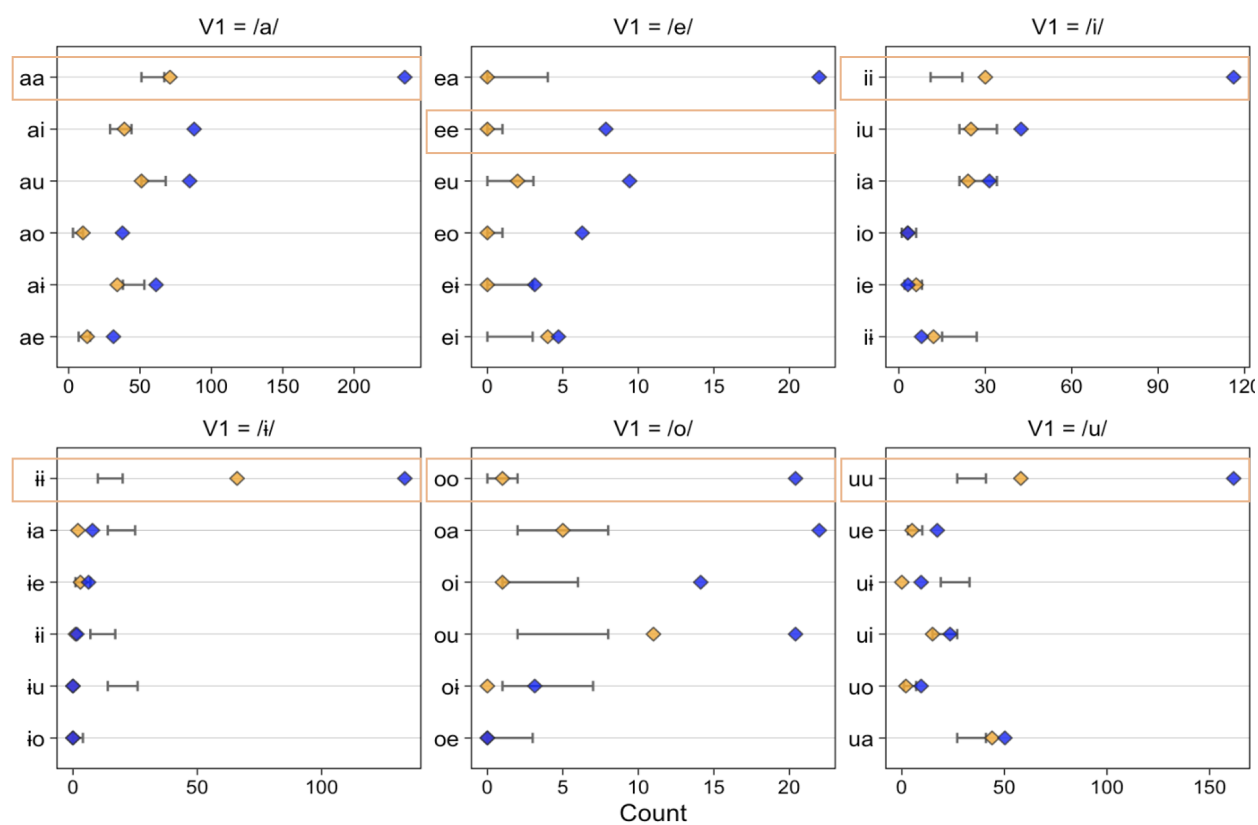


Figure 5: Monte Carlo test of significance: vowel-matching