Phonological reanalysis is guided by markedness: the case of Malagasy weak stems

Abstract. A key goal in phonology is to understand the factors that affect phonological learning. This paper addresses the issue by examining how paradigms are reanalyzed over time. Malagasy has a class of stems, called weak stems, where final consonants alternate under suffixation. Comparison of historical and modern Malagasy shows that weak stem paradigms have undergone extensive reanalysis in a way that cannot be predicted by the probabilistic distribution of alternants. This poses a problem for existing quantitative models of reanalysis, where reanalysis is always towards the most probable alternant. I argue instead that reanalysis in Malagasy is driven by both distributional factors and a markedness bias. To capture the Malagasy pattern, I propose a maximum entropy learning model (Goldwater & Johnson, 2003), with a markedness bias implemented via the model's prior probability distribution. This biased model successfully predicts the direction of reanalysis in Malagasy, outperforming purely distributional models.

1 Introduction

Understanding the extent to which different biases affect phonological acquisition is a central question in phonology. This question has been addressed extensively through experimental work (e.g. Wilson, 2006; Moreton & Pater, 2012b,a) and research on child language acquisition (e.g. Singleton & Newport, 2004; Peperkamp et al., 2006). Since Kiparsky's seminal work on phonological change (1965; 1968; 1978, et seq), it has been recognized that studying language change over time can also give us insight into the factors that drive phonological learning. The data may be harder to interpret due to the large time depth, but also potentially offer more contextual validity than experimental work. Insights from language change can therefore complement experimental and acquisition research.

The current study focuses on a specific type of change, reanalysis in paradigms. Morphological paradigms can have neutralizing alternations that cause ambiguity in one or more slots of the paradigm. For example, Middle High German (MHG) had a well-known process of final obstruent devoicing that created ambiguity in non-suffixed forms (Sapir, 1915, p.237; Kiparsky, 1968, p.177, etc.). As demonstrated by the examples in (1a), given a non-suffixed MHG stem with a final voiceless obstruent, the final obstruent could either surface as voiceless (e.g. zak∼zakə), or show a voicing alternation (e.g. vek∼vegə).

(1) Reanalysis of obstruent voicing in Yiddish (nominative sg. vs pl. paradigm)

(a)			(b)			(c)		
MHG		→	Early Yiddish		→	Modern Yiddish		
sg.	pl.		sg.	pl.		sg.	pl.	
ve**k**	ve**g**ə		ve**k**	ve**g**(ə)		ve**g**	ve**g**ən	'way'
za**k**	za**k**ə		za**k**	ze**k**(ə)		za**k**	ze**k**	'sack'

Neutralizing alternations like this can be challenging to the language-learning child, and be prone to reanalysis over time. This was the case for voicing alternations in Yiddish, a direct descendant of MHG. Final obstruent devoicing was present in early Yiddish (1b), but subsequently lost in Modern Yiddish, where the singular forms were reanalyzed to remove neutralization. As shown in (1c), the voicing value of the plural was reintroduced to the singular (Albright, 2010).

Notably, there are relatively few quantitative models that can make strong, language-specific predictions about the output and direction of reanalysis. Existing models predict reanalysis to be solely based on the probabilistic distribution of segments. In these models, reanalysis is always in the direction of the more probable alternant.

In the current study, however, I find that for Malagasy, there has been extensive reanalysis that contradicts the predictions of purely distributional models. Specifically, in a class of stems called ‘weak stems’, there has been extensive reanalysis in a direction that is not predicted by distributional properties in the lexicon. I argue that reanalysis in Malagasy is sensitive to both distributional and markedness effects. Building on these results, I propose a constraint-based model of reanalysis which has a markedness bias.

The rest of the paper is organized as follows: §2 introduces existing models of reanalysis, and presents the descriptive facts of Malagasy weak stems. In §3, I present results of a corpus study comparing historical Malagasy forms with modern Malagasy data, to show that reanalysis has occurred in a direction that cannot be predicted by purely inductive models of reanalysis. Finally, §4 proposes a model of reanalysis which incorporates a markedness learning bias.

2 Background

2.1 Quantitative approaches to modeling reanalysis

Existing quantitative models of reanalysis (or more generally of morphophonological paradigm learning) are inductive, and therefore predict change to be driven purely by statistical distributions. One representative model of this variety is the Minimal Generalization Learner (MGL; Albright & Hayes, 2002; Albright, 2002; Albright & Hayes, 2003, et seq.).

The MGL first compares different members of the paradigm, and learns word-specific rules mapping from one form to another. With regards to the MHG pattern introduced above, the MGL would generate rules like in (2). When forms share the same change, the model finds what features they share in common, and generalizes rules based on these shared features. For example, a rule ∅→ə/[−voiceless, −continuant]_# may be generated from comparison of forms (a) and (b). The result is a system of stochastic rules which predict the inflected form of a paradigm given an input base.

(2) Word-specific rules learned by the MGL for MHG

	sg.	pl.	word-specific rule
(a)	za**k**	za**k**ə	∅→ə/ vek_#
(b)	mu**t**	za**t**ə	∅→ə/ mut_#
(c)	ve**k**	ve**g**ə	k→gə/ ve_#

In the MGL, reanalysis occurs when the grammar derives the incorrect output for certain derived forms, and these errors come to replace the older, exceptional forms. This model has been shown to explain the direction of historical restructuring in various languages, including Lakhota (Albright, 2008b), Yiddish (Albright, 2010), and Korean (Kang, 2006). Details of model implementation can be found in Albright & Hayes (2003). What is important to note is that this model learns rules inductively, and predicts reanalysis to be in the direction of the statistically most probable outcome, given the distribution of sounds in a paradigm.

Albright’s model is rule-based, and generates sets of rules that predict the outcome of paradigm reanalysis. An alternative analogical approach is exemplified by the Generalized Context Model (GCM Nosofsky, 2011). This approach is ‘similarity-based’, meaning that in principle, any words that are similar enough to each other can serve as the basis for reanalysis. Broadly speaking, similarity-based models are less restrictive than rule-based models, and are potentially able to capture a wider range of effects (Albright & Hayes, 2003). However, both approaches predict that reanalysis will match the distributions of the input data.

Inductive learning is also possible in stochastic constraint-based models such as Maximum Entropy Harmonic Grammar (MaxEnt; Goldwater & Johnson, 2003; Smolensky, 1986). As a preview, in §4, an inductive constraint-based model will be used as a baseline, and compared to more models which incorporate learning biases.

2.2 Malagasy phonology and weak stem alternations

Malagasy, the national language of Madagascar, is an Austronesian language belonging to the Malayo-Polynesian subgroup (Rasoloson & Rubino, 2005). The term Malagasy really refers to a macro-language that covers many dialects distributed throughout Madagascar (Lewis et al., 2014). The following study uses data from Official Malagasy (OM), which is the standardized, institutional dialect that is based on the Merina dialect spoken in the capital city Antananarivo. All subsequent descriptions and analysis will assume data from OM.

Malagasy has inflectional and derivational morphology, much of which involves morphophonological alternations. In a subset of so called **weak stem** consonant alternations, the expected alternant (based on historical evidence) often does not match the observed alternant, suggesting that substantial reanalysis has occurred.

Malagasy has been studied extensively. The phonetic system is described by Howe (2021), and basic facts on the morphology and phonology are documented in work such as Keenan & Polinsky (2017) for OM, and O'Neill (2015) for the closely related Betsimisaraka dialect. Formal analyses of Malagasy phonology, including of weak stem alternations, have been done in both generative rule-based frameworks (Dziwirek, 1989) and OT (Albro, 2005). Moreover, the history of Malagasy can be traced in some detail through the work of Austronesianists (e.g. Dahl, 1951; Mahdi, 1988; Adelaar, 2013). Additionally, dictionary data is digitized in the Malagasy Dictionary and Encyclopedia of Madagascar (MDEM; de La Beaujardière 2004), which compiles data from multiple Malagasy dictionaries. Historical comparative data is also available the Austronesian Comparative Dictionary (ACD; Blust et al., 2023).

In this section, I provide a descriptive account of Malagasy phonology and weak stem alternations, based on work by Keenan & Polinsky (2017) and Howe (2021).

2.2.1 Malagasy phonology

Malagasy words have a strict (C)V syllable structure, where codas are not allowed. Word stress is phonemic but generally penultimate, though there are exceptions to be discussed in the following section.

Malagasy has five phonemic monophthongs /i e a o u/. /o/ is considered to be non-phonemic (or marginally phonemic) in many descriptions of Malagasy (e.g. Rasoloson & Rubino, 2005; O'Neill, 2015). However, it has become much more common because /ua/ and /au/ sequences have merged to /o/ in OM (Howe, 2021).

The consonants of Malagasy are given in Table 1. /ŋ/ is given in parentheses because although it is non-phonemic in OM, it is phonemic in many dialects of Malagasy.

All subsequent examples are presented in IPA, with the following caveats. Prenasalized obstruents are written as nasal-obstruent sequences (e.g. *mb* corresponds to [^{m}b]). [ʈʂ] and [ɖʐ] are generally retroflex, but can vary in production between speakers (Howe, 2021), and have been described in prior work as post-alveolar (e.g. Keenan & Polinsky, 2017). In addition, [r] is a short

	bilabial	labiodental	dental	alveolar	retroflex	velar	glottal
plosives	p, b		t, d			k, g	
	ᵐp, ᵐb		ⁿt, ⁿd			ᵑk, ᵑg	
affricates				ts, dz	ʈʂ, ɖʐ		
				ⁿts, ⁿdz	ⁿʈʂ, ⁿɖʐ		
nasals	m		n			(ŋ)	
trills/flaps				r∼ɾ			
fricatives		f, v		s z			h
lat. approximants				l			

Table 1: Malagasy consonant chart

alveolar trill in most dialects including OM, but is often realized as a tap [ɾ] in casual speech (Howe, 2021).[1]

2.2.2 Weak stems

Malagasy has a class of forms that Keenan & Polinsky (2017) refer to as weak stems. These roots have antepenultimate stress (if long enough), and always end in one of the three ‘weak syllables’ [ʈʂa], [ka], or [na].[2]

When weak stems are suffixed, the consonant of the weak syllable ([ʈʂ], [k], or [n]) may alternate with another consonant. Patterns of alternation are summarized in Table 2, using the active and passive forms of verbs. In addition to these alternants, the lexicon also contains four words few minority patterns, including stems where final ʈʂa alternates with [s]. I exclude these here because they are so low in frequency that they do not affect my analysis, but they are given in the Appendix (Table 17) for reference. In the suffixed forms, the final vowel of the weak stem is not present, leaving the alternating consonant at a morpheme boundary. As demonstrated in these examples, suffixation also shifts stress one syllable to the right.

pattern		**active (m + stem)**	**passive (stem + ana)**	
na ∼	n	manˈɖʐavi**n**a	anɖʐaˈvi**n**ana	‘to bear leaves’
	m	maˈnanɖʐa**n**a	aˈnɖʐá**m**ana	‘to try’
ka ∼	h	maˈngata**k**a	angaˈta**h**ana	’to ask for’
	f	maˈnaha**k**a	anaˈha**f**ana	’to scatter’
ʈʂa ∼	r	miána**ʈʂ**a	ianá**r**ana	‘to learn’
	t	maˈnanɖʐa**ʈʂ**a	anaˈnɖʐa**t**ana	‘to promote’
	f	maˈnɖʐaku**ʈʂ**a	anɖʐaˈku**f**ana	‘to cover’

Table 2: Patterns of consonant alternation in Malagasy weak stems

Note that even though the weak stem alternants are neutralized in stem-final prevocalic position, the same phonemes are fully contrastive in other positions (i.e in initial and medial position).

---

[1] My personal observations in work with a consultant matches Howe’s phonetic descriptions.

[2] According to Howe (2021), the final vowel of weak stems is often devoiced or reduced.

Table 3 provides minimal pairs that demonstrate this. For example, /t/ and /r/ are neutralized to [ʈʂ̺a] in unsuffixed weak stems, but are contrastive as demonstrated by minimal pairs like [a**t**u] 'close at hand' and [a**r**u] 'barrier, rampart'.

contrast	**position**	**word 1**		**word 2**	
/t/ vs. /r/	initial	**t**aba	'grasp, grab'	**r**aba	'do w/o thought or order'
	medial	a**t**u	'close at hand'	a**r**u	'barrier, rampart'
/h/ vs. /f/	initial	**h**ana	'lend/borrow money'	**f**ana	'heat'
	medial	a**h**u	'I, myself'	a**f**u	'fire, calamity'
/n/ vs. /m/	initial	**n**ani	'neck of fishing basket'	**m**ani	'stink'
	medial	le**n**i	'wet'	le**m**i	'softness'

Table 3: Minimal pairs showing that weak stem alternants are contrastive

The standard formal analysis for weak stems is that they are underlyingly consonant-final (Albro, 2005). For example, the surface forms [m-iˈanaʈʂ̺a]∼[iaˈnar-ana] would have the stem UR /ianar/, with surface forms derived as in (3). First, all words are assigned penultimate stress, and the stem-final consonant is neutralized to [ʈʂ̺], [k], or [n] (here, /r/ neutralizes to [ʈʂ̺]). In the suffixed form, /r/ is medial and therefore protected from neutralization. Finally, an epenthetic /a/ is added to resolve the violation against codas (counterbleeding the final-C neutralization). Antepenultimate stress falls out naturally from the rule ordering, where stress assignment precedes vowel epenthesis. As I discuss below, the analysis in (3) is in fact a recapitulation of the historical development of weak stem alternations.

(3) Derivation for surface forms of /ianar/ in a formal analysis of weak stems

UR	/m-ianar/	/ianar-an/
Penultimate stress assignment	miˈanar	iaˈnaran
Final C neutralization (/r/→ʈʂ̺/_#)	miˈanaʈʂ̺	iaˈnaran
Vowel epenthesis (∅→a/C_#)	miˈanaʈʂ̺a	iaˈnarana
SR	[miˈanaʈʂ̺a]	[iaˈnarana]

2.2.3 Historical development of weak stem alternations

The linguistic history of Malagasy has been studied in detail. The following description summarizes findings from a large body of scholarship, including Dahl (1951), Hudson (1967), Mahdi (1988), and Adelaar (2012, 2013).

Malagasy weak stem alternations started as a series of relatively common final consonant neutralizations, which were subsequently obscured by a process of final vowel epenthesis. Vowel epenthesis was motivated by a phonotactic restriction against codas which developed around 600AD, when speakers of proto-Malagasy migrated from Kalimantan into the Comoro Islands. Contact with Bantu during this migration significantly influenced Malagasy grammar, and is largely thought to have caused the development of final open syllables in Malagasy. For most final consonants, epenthesis of a final vowel removed final codas, resulting in the weak stems of current Malagasy.

The development of Malagasy from Proto-Austronesian (PAn) can be broadly be split into three stages: Proto-Malayo-Polynesian (PMP), Proto-Southeast Barito (PSEB), and Proto-Malagasy (PMlg). The examples in (4) trace a subset of weak stems through these stages, to illustrate the historical development of some weak stem alternations.

(4a) illustrates the development of a ʈʂa∼t alternating weak stems, which historically end in voiceless coronal stops, in this case *t. Final **-t* neutralized to **-ʈʂ* in PMlg; this affected the non-suffixed forms, while stem-final [t] was preserved in suffixed forms. Following this, epenthesis of a final vowel resulted in the current ʈʂa∼t alternation.

In (4b), on the other hand, the PMP stem ends in **D* [ɖ]. In the non-suffixed form, this final consonant devoiced to **-t*, and then neutralized to [ʈʂ]. In the suffixed form, **D* lenited to [r] due to regular sound change (*D>r; Adelaar 2012). This was followed by final vowel epenthesis, resulting in the observed ʈʂa∼r alternation. Note that while final devoicing (*-D >-t) and lenition (*D > r) are both thought to have taken place in PSEB, devoicing must have preceded lenition for the observed alternations to be possible.

Examples (4c-4d) provide similar illustrative cases for ka-final alternations. First, in PMlg, historical **k* spirantized to [h] intervocalically (before the epenthesis of final vowels). This affected the stem-final **-k* of suffixed forms, but not the unsuffixed forms, resulting in ka∼h alternations, as shown in (4c). The development of ka∼f alternating follows from a similar process, given in (4d). First, **-p* and **-k* neutralized to [-k] word-finally. This affected the unsuffixed form, but not the suffixed forms, where stem-final *p is intervocalic. This was followed by spirantization in the suffixed forms from **p*> [f].

(4) Examples: historical basis of final consonant alternations; changes relevant to the consonant alternation are given in parentheses.[3]

a. ʈʂa∼t alternation[4]

PMP	*yawut	*piyawutan	
PSEB	*ˈawut	*piaˈwutan	
PMlg	*ˈavuʈʂ	*fiaˈvutan	(Final affrication, *-t > -ʈʂ)
	*ˈavuʈʂa	*fiaˈvutana	(Final V epenthesis)
Mlg	ˈavuʈʂa	fiaˈvutana	‘to uproot’

b. ʈʂa∼r alternation

PMP	*bukiD	*bukiD-ən	
PSEB	*ˈwukit	*wuˈkiDən	(Final devoicing, *-D > *-t)
	*ˈwukit	*wuˈkirən	(Lenition, *D, *d > r)
PMlg	*ˈwukiʈʂ	*wuˈkirən	(Final affrication, *-t > *-ʈʂ)
	*ˈwukiʈʂa	*wuˈkirəna	(Final V epenthesis)
Mlg	ˈvuhiʈʂa	vuˈhirina	‘to make convex’

c. ka∼h alternation

PSEB	*ˈtətək	*təˈtək-ən	
PMlg	*ˈtetek	*teˈtehen	(spirantization, *k > h/V___V)
	*ˈteteka	*teˈtehena	(Final V epenthesis)
Mlg	ˈtetika	teˈtehina	‘to cut into small pieces’

d. ka∼f alternation

PMP	*heyup		
PSEB	*ˈtiup	*pi-tiˈup-an	
PMlg	*ˈtiuk	*pitiˈupan	(Final stop neutralization, *-p >*-k)
	*ˈtiuka	*fitsiˈufana	(Final V epenthesis; spirantization, *p > f/V___V)
Mlg	ˈtsiuka	fitsiˈufana	‘to lick’

---

[3] Stress becomes non-contrastive and uniformly penultimate in PSEB; later on, epenthesis of a final vowel resulted in forms with antepenultimate stress, making stress contrastive.

[4] Protoforms use the orthographic conventions established by Dyen (1951). The phonetic value of *R is thought to be [ʀ], *C to be [cç], *y to be [j], *D to be [ɖ].

stem-final	**alt.**	**example**	**PMP/PAn**
n	n	ˈanki**n**a∼aˈnki**n**-ina	<*n, *ŋ, *l
	m	aˈmpiri**n**a∼ampiˈri**m**-ana	<*m
ʈʂ̺	r	ˈampa**ʈʂ̺**a∼ aˈmpa**r**-ana	< *j [gʲ],*d,*D [ɖ]
	t	ˈhara**ʈʂ̺**a∼ haˈra**t**-ana	< *t, *C [cç]
	f	ˈdidi**ʈʂ̺**a∼ diˈdi**f**-ana	< *p,*b
k	h	baˈlia**k**a∼ibaliˈa**h**-ana	<*k,*g
	f	ˈhiri**k**a∼ hiˈri**f**-ana	<*p,*b

Table 4: Weak stem alternants and corresponding historical consonants

Table 4 summarizes all the expected weak stem alternants in Malagasy, given the historical final consonants in PMP. In general, the historical origin of weak stems are well-understood, and the observed alternants in modern Malagasy are expected to correspond to specific historical final consonants.

As a caveat, most consonant-final PMP forms reflect as weak stems in Malagasy, but there are three exceptions. First, PMP *s, *q, *h were deleted in all environments in PSEB, so do not result in consonant alternations. Additionally, PMP glides *w,*y [j] deleted or coalesced with the preceding vowel in final position, and hardened to *v and *z elsewhere. Stems with a historic final glide therefore have ∅∼C alternations in modern Malagasy (e.g. [ˈlalu∼laˈluv-ana] <*lalaw, 'pass without stopping'). Finally, *s in early Malay loanwords was deleted word-finally, but retained in other positions. These forms have ∅∼s alternation in modern Malagasy (e.g [miˈlefa∼leˈfa**s**-ana] <*ləpas (Malay) 'gone, escaped'). The reflexes of different PMP final consonants are summarized in Table 5.

Coda resolved by...	**PMP cons.**	**Mlg alternation**	**Example**
Vowel epenthesis	*-k,*-g	ka∼h	baˈlia**k**a∼ibaliˈa**h**-ana
	*-p, *-b	ka/ʈʂ̺a∼f	ˈhiri**k**a∼ hiˈri**f**-ana
	-t,-c	ʈʂ̺a∼t	ˈhara**ʈʂ̺**a∼ haˈra**t**-ana
	*-d, *-D,*-j	ʈʂ̺a∼r	ˈampa**ʈʂ̺**a∼ aˈmpa**r**-ana
	-n,-ŋ,*-l	na∼n	ˈanki**n**a∼aˈnki**n**-ina
	*-m	na∼m	aˈmpiri**n**a∼ampiˈri**m**-ana
Deletion/coalescence	*-y [j]	∅∼z	ˈalu∼aˈlu**z**-ina
	*-w	∅∼v	ˈlalu∼laˈlu**v**-ana
Deletion	*-s (loan phoneme)	∅∼s	miˈlefa∼leˈfa**s**-ana

Table 5: Malagasy reflexes of stem-final PMP consonants

Where there is a lot of mismatch between the expected alternant (given the PMP final consonant) and the actual alternants observed in Malagasy, this suggests that reanalysis has occurred. Examples of mismatches are given in (5). In (5a), for example, [ˈlumuʈʂ̺a] is expected to have [t] as the alternant because the stem historically ended in *t. Instead, the alternant that surfaces is [r], indicating reanalysis in the direction of t→r. As will be seen in §3, the ʈʂ̺a-final weak stems in particular seem to have undergone extensive reanalysis, and often do not surface with the expected alternant given the PMP final consonant.

(5) *Examples of mismatches between PMP and Malagasy*

	PMP	**Malagasy**	**gloss**
a.	*lumu**t**	ˈlumu**ʈʂ**a∼luˈmu**r**-ina	‘seaweed’
b.	*qade**p**	ˈatri**k**a∼aˈtre**h**ina	‘face, facade’
c.	*dale**m**	ˈlali**n**a-laˈli**n**-ina	‘deep, profound’

In fact, in modern-day Malagasy, weak stem alternations appear to be partially conditioned by phonological factors, and partially dependent on the historical final consonant. Mahdi (1988), in one of the most comprehensive studies of Malagasy weak stems, notes the following generalizations. First, na-final weak stems usually alternates with [n], but may alternate with [m] if the stem-final consonant was historically *m. Final ka usually alternates with [h], but may alternate with [f] if the historical stem-final consonant was labial, or if the nearest consonant in the stem is [h]. In other words, alternation in ka-final weak stems is partially driven by a dissimilative pattern.

For final [ʈʂa], Mahdi again finds a dissimilative effect. Specifically, in present-day Malagasy, [ʈʂa] alternates with [r] in general, but will alternate with [t] if the stem already contains an [r]. Where there are exceptions to this pattern (i.e. where the alternant is [t] or [f] in a non-dissimilatory environment), it is because the historical final consonant was historically *t, *p, or *b.

Mahdi’s findings (and existing work on Malagasy weak stems) have noted the connection between Malagasy alternants and their historical consonant. However, they have not focused on exactly what direction reanalysis happened in, or why there is so much mismatch between the historical consonant and observed alternant in modern-day Malagasy.

3 Reanalysis in weak stems

Although the historic basis of weak stems is relatively well-understood, there are many mismatches between the observed and expected alternants in Malagasy (given the historic PMP consonant), suggesting that substantial reanalysis has occurred in Malagasy. In the following section, I discuss the predicted outcome of reanalysis under a distributional approach, and show that reanalysis in Malagasy differs from these predictions.

Reanalysis of weak stems in Malagasy always results in the suffixed forms being changed. However, reanalysis may still vary in terms of which alternants are more likely to be reanalyzed, and which alternants are the preferred output of reanalysis.

For example, final [ʈʂa] can alternate with [t], [r], or [f] in the suffixed form. Given these possible alternants, one possible direction of reanalysis is t→r, where a ʈʂa∼t alternating stem is reanalyzed as r-alternating. Conversely, reanalysis could happen in the opposite direction, where a historically ʈʂa∼r alternating stem becomes t-alternating. (6) summarizes the possible outcomes of reanalysis, given the hypothetical ʈʂa-final weak stem [ˈpakuʈʂa].

(6) *Possible directions of reanalysis for ʈʂa-final weak stems (example stem: [ˈpakuʈʂa])*

Direction	passive (stem + ana)
t → r	pakut-ana→pakur-ana
t → f	pakut-ana→pakuf-ana
r → t	pakur-ana→pakut-ana
r → f	pakur-ana→pakuf-ana
f → t	pakuf-ana→pakut-ana
f → r	pakuf-ana→pakur-ana

In this section, I examine the directions of reanalysis in Malagasy weak stems in detail. As a preview of the results, for the ka- and na-final weak stems, reanalysis is generally in the direction predicted by an inductive approach (i.e. in the direction of the historically more frequent alternant). For the ʈ͜ʂa-final weak stems, however, there has been extensive reanalysis in the direction of t→r, which is *not* predicted by distributional information. I will argue that this reanalysis is driven by a markedness bias, specifically a tendency to avoid intervocalic stops.

Results of this section are based off of comparison of historical and modern Malagasy data, where historical data refers to PMP protoforms. Many Malayo-Polynesian languages maintain the final consonant contrasts that were neutralized in Malagasy. This is demonstrated in (7), which shows examples of final consonant contrasts that were neutralized in Malagasy (as weak syllables), but maintained in other related languages. As a result, PMP reconstructions provide a reliable picture of what the Malagasy weak stem pattern may have looked like before reanalysis.

(7) *Final consonant contrasts across Malayo-Polynesian languages*

alt.	**PMP**	**Malagasy**	**Malay**	**Javanese**	**Tagalog**	**Balinese**
n	*bula**n**	vula**na**	bula**n**	bula**n**		
m	*dale**m**	lali**na**	dala**m**	dale**m**		
t	*buha**t**	vua**ʈ͜ʂa**	bua**t**			bua**t**
r	*hate**D**	ati**ʈ͜ʂa**	(h)anta**r**	ate**r**	hati**d**	
h	*ana**k**	ana**ka**	ana**k**	ana**k**	ana**k**	
f	*qade**p**	atri**ka**		hada**p**		hare**p**

Historical data are taken from the Austronesian Comparative Dictionary (ACD; Blust et al., 2023) and Adelaar (2012). Protoforms had to reconstruct back to PMP, and were excluded if they were only reconstructable back to Proto-Western Malayo-Polynesian (PWMP). Additionally, protoforms were excluded if they had less than 6 cognates.

Modern Malagasy words are taken from the Malagasy Dictionary and Encyclopedia of Madagascar (MDEM; de La Beaujardière 2004), which is an online dictionary that compiles data from multiple Malagasy dictionaries.[5]

§3.1 will discuss the distribution of final obstruents in PMP, and what this predicts about the direction of reanalysis in Malagasy. These predictions are compared to the actual observed directions of reanalysis in §3.2. §3.3 provides additional indirect evidence on the directions of reanalysis using data from modern Malagasy.

3.1 Predicted reanalyses under an inductive approach

In a purely inductive model of morphophonological learning, reanalysis would always be in the direction of the more frequent alternant (subject to phonological conditioning). The alternants predicted under this approach can be approximated by looking at the distribution of final consonants in PMP, before extensive reanalysis had taken place. Table 6 shows the distribution of all PMP protoforms with final consonants which would be reflected as weak syllables in Malagasy ($n = 805$). Results are organized by which alternant each PMP final consonant would correspond.

There is one complication when [f] is the alternant. Historically, final *-p and *-b neutralized to either *-k or *-t, with a slight bias towards *k (Dahl, 1951; Adelaar, 2012). Consequently, PMP forms ending in a labial stop tend to reflect as ka-final weak stems, but also often reflect as ʈ͜ʂa-final weak stems. In Table 6, all PMP forms ending in labial stops are assumed to correspond to

---

[5] The primary dictionaries that the MDEM sources from were all published from 1885-1998; more details can be found in `https://en.mondemalgache.org/bins/sources`.

Type	alternant	count	P	Predicted reanalysis
ka	h (<*k)	183	0.81	f→h
	f (<*p,*b)	42	0.19	
na	m (<*m)	35	0.10	m→n
	n (<*n,*ŋ)	302	0.90	
ʈʂa	r (<*j,*r,*d,*ɖ)	52	0.25	r→t
	t (<*t)	162	0.75	

Table 6: Expected distribution of Malagasy weak stem alternants, based on the distribution of PMP final consonants.

ka-final weak stems in Malagasy. This simplification should not impact the analysis, since ʈʂa∼f alternating forms make up a very small proportion of ʈʂa-final weak stems (n = 7,≈2.4%).

From this data, we see that ka-final weak stems have more h-alternating forms, na-final weak stems have more non-alternating forms, and ʈʂa-final weak stems have more t-alternating forms. An inductive approach predicts that reanalysis should generally be in the direction of these more frequent alternants. For example, reanalyses of ʈʂa-final stems should be in the direction of r→t, rather than t→r. Predictions are summarized in the rightmost column of Table 6.

Mahdi's (1988) findings on dissimilatory effects in weak stems are also partially replicated in the PMP data. Consider (8), which summarizes the protoforms corresponding to ʈʂa-final stems by whether or not there is a preceding (non-final) [r]. PMP *r ,*d, and *j (in non-final position) are coded as corresponding to Malagasy [r], but excluded if they occurred as the first consonant in a CC cluster. This is because consonant clusters were historically simplified in PMP by deleting the first consonant (e.g. vavaʈʂa, < *bajbaj).

From this data, there appears to be evidence for r-dissimilation. Out of the 28 protoforms coded as containing a preceding [r], only one would reflect as [t]-alternating in Malagasy. Put another way, when the expected alternant is [r], only one form was coded as containing a preceding [r] (n = 1/52, 2%). In contrast, when the expected alternant is [t], 27 forms have a preceding [r] (n = 27/163, 17%).

(8)

	Does stem have [r] (<*r,*d,*ɖ,*j)?	
alternant	yes	no
t	27	136
r	1	51

For ka-final weak stems, the evidence for a dissimilatory pattern in PMP is weaker. If dissimilation were present, we would expect the proportion of stems with an immediately preceding *k (corresponding to [h] in modern Malagasy) to be smaller when the expected alternant is [h]. When the expected alternant is [h], around 7% (n = 13/183) of protoforms have a preceding *k. When the expected alternant is [f], 22% of forms (n = 9/42,21%) have a preceding *k. In other words, there is a dissimilatory pattern, but it is weaker than the r-dissimilation pattern observed in (8).

(9)

	does stem have h (<*k)?	
alternant	yes	no
h	13	170
f	9	33

3.2 Observed directions of reanalysis

In this section, I discuss form-by-form comparisons of PMP stems to their weak stem reflexes. Where there is a mismatch between PMP and Malagasy, the direction of reanalyses can be inferred. The ACD contains 143 protoforms that reflect as productive suffixed forms in Malagasy. 56 were removed following the exclusionary criteria discussed above, leaving 87 forms to be analyzed. The data is also supplemented with 49 Malay and Javanese loanwords from the World Loanword Database (WOLD; Adelaar, 2009) and Adelaar (1994). These are all early loans, introduced to Malagasy before the development of weak stems (Adelaar, 1989). Tables 7-9 summarize whether the alternant observed in Malagasy matches the expected one given the historical consonant (or in the case of loanwords, the final consonant of the source word).

Table 7 shows the results for na-final weak stems. The column named 'PMP' shows the expected alternant given the PMP protoform, while the column named 'Mlg' shows the actually observed alternant in Malagasy. Mismatches between PMP and Mlg indicate that a reanalysis has occurred. Overall, there are relatively few reanalyses (n = 3), but most are in the direction of m→n (e.g. [ˈlali**na**∼lalˈ**in**-ina] <*dale**m** 'inside, deep'). This is in line with the predictions of an inductive approach.

Of the stems expected to be n-alternating, only one has been reanalyzed in the direction of n→m (n = 1/39, 3%); the reanalyzed stem is [ˈteno**na**∼teˈno**m**-ina] (<*tenu**n**) 'to weave/be woven'. Given the lack of data, it is hard to tell what the cause is.[6] Overall, comparisons for the na-final weak stems are tentatively in line with a statistical learning approach.

PMP	**Mlg**	**Match?**	**Count**
m	m	yes	2
	n	no (m→n)	2
n	n	yes	38
	m	no (n→m)	1

Table 7: Expected (PMP) vs. observed (Malagasy) alternant of na-final stems, based on known protoforms and loanwords

Table 8 shows the reanalyses for ka-final weak stems. Once again, there are relatively few cases of reanalyses (n = 2). However, both case of reanalysis are in the direction of f→h (e.g. [ˈaʈʂi**ka**∼fiaˈʈʂe**h**-ana] <*qade**p** 'face, facade'), in line with the predictions of an inductive approach. In contrast, there are no reanalyses in the direction of f→h.

Note that the data did not contain any stems where the immediately preceding consonant is [h]. As such, it is unclear whether a dissimilatory effect was active in the reanalysis of ka-final weak stems. However, one item, which was excluded because it was only reconstructed to PWMP (Proto-Western Malayo-Polynesian), shows reanalysis in the direction of h→f that could potentially be attributed to h-dissimilation. This word, [ˈlau**ka**∼laˈu**f**ana] (<PWMP *lahu**k**) 'meat/relish eaten with rice', historically had a preceding [h] which was subsequently elided in PSEB.

Table 9 shows results for ʈʂa-final weak stems. The rightmost column, 'has r?', indicates, for each row, the number of forms which had an [r] in the stem. For ʈʂa-final stems, extensive reanalysis has occurred towards [r]. Of the stems that were historically expected to have [t] as the alternant, over half (23/40, 57%) have been reanalyzed in the direction of t→r (e.g.

---

[6] This change of n→m does not seem to be from a dissimilatory effect, since there was no nasal dissimilation found in either PMP or modern Malagasy. However, nasal dissimilation is documented the Betsimisaraka dialect of Malagasy (O'Neill, 2015)

PMP	Mlg	Match?	Count
f	f	yes	3
	h	no (f→h)	2
h	h	yes	36
	f	no (h→f)	0

Table 8: Expected vs. observed alternant of ka-final stems, based on known protoforms and loanwords

[ˈhudi**ʈ͜ʂ**a∼huˈdi**r**-ina] <*kuli**t**, ‘skin, hide’). In contrast, when the expected alternant is [r], there is only one case of reanalysis (n = 1). Moreover, the one case of reanalysis in the ʈ͜ʂa∼f alternating forms is in the direction of f→r ([ˈhala**ʈ͜ʂ**a∼aŋaˈla**r**-ina] < *ala**p**, ‘theft, robbery’).

PMP	Mlg	Match?	Count	has r?
t	t	yes	17	7 (41%)
	r	no (t→r)	23	0
	f	no (r→f)	0	0
r	r	yes	11	0
	t	no (r→t)	1	1
	f	no (f→t)	0	0
f	f	yes	3	1 (33%)
	t	no (f→t)	0	0
	r	no (f→r)	1	0

Table 9: Expected vs. observed alternant of ʈ͜ʂa-final stems, based on known protoforms and loanwords

Additionally, r-dissimilation appears to be active in the reanalysis of ʈ͜ʂa-final weak stems, in that reanalysis to [r] is blocked if the stem has a preceding [r]. As seen in Table 9, when the alternant was reanalyzed to be [r], the stem never contained a preceding [r]. In addition, out of the t-alternating stems that were not reanalyzed, a relatively larger proportion (n = 7/17, 41%) had a preceding [r] (e.g.[ˈur̠i**ʈ͜ʂ**a∼uˈr̠i**t**ana] <*quri**t**, ‘stroke, line’).

The only example of reanalysis in the direction of r→t is likely also motivated by r-dissimilation. The reanalyzed form [ˈsanɖ͜ʐa**ʈ͜ʂ**a∼anaˈnɖ͜ʐa**t**-ana] (<sanda**r**, Malay loan) does not have a preceding [r] in modern Malagasy, but [nɖ͜ʐ] sequences are historically [nr], and only affricated to [nɖ͜ʐ] in a later stage of PSEB (Proto Southeast-Barito).

The direction of reanalysis in ʈ͜ʂa-final weak stems goes against predictions of an inductive approach. Based on the PMP distribution, there should more [t]-alternating forms than [r]-alternating forms. However, reanalyses are overwhelmingly towards the less frequent alternant, in the direction of t→r.

3.3 The result of reanalysis: weak stem alternations in modern Malagasy

This section describes the distribution of weak stem alternants in modern Malagasy, using 1628 stems taken from the MDEM. This data supplements the above results, by providing indirect evidence for the direction of reanalysis that has taken place.

Table 10 summarizes the distribution of weak stem alternants in modern Malagasy; the rightmost column shows the expected directions of reanalysis for each weak stem type, given the histor-

ending	alternant	Freq		Expected reanalyses
na	n	580	(97.3%)	m→n
	m	13	(2.2%)	
	other	3	(0.5%)	
ka	h	668	(95.0%)	f→h
	f	35	(5.0%)	
ʈʂa	r	231	(70.2%)	t,f→r
	t	89	(27.1%)	
	f	7	(2.1%)	
	s	2	(0.6%)	

Table 10: Proportion of alternants for modern Malagasy weak stems

ical distributions discussed so far. The na-final weak stems are overwhelmingly non-alternating, where 97.7% of the sampled forms are non-alternating. This distribution is consistent with the finding that reanalyses have been in the direction of m→n, increasing the relative frequency of non-alternating na-final weak stems.

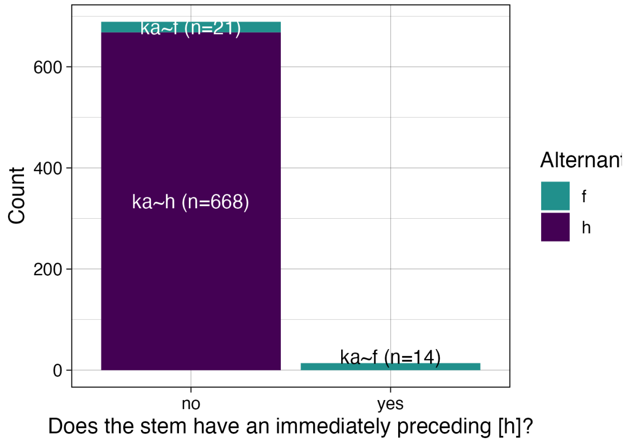

Figure 1: Distribution of alternants in ka-final weak stems

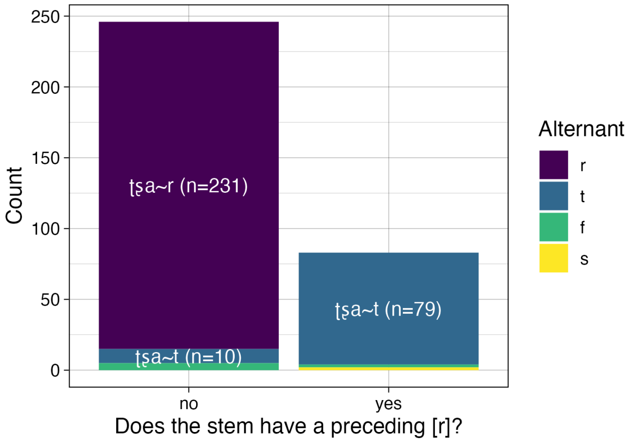

Figure 2: Distribution of alternants in ʈʂa-final weak stems

For ka-final weak stems, [h] is overwhelmingly the preferred alternant, accounting for 94.8% of the sampled forms. Again, this distribution is consistent with the finding that reanalyses have been in the direction of f→h.

In addition, recall that Mahdi (1988) finds evidence for h-dissimilation in ka-final weak stems. Although no such effect was found in PMP (or in the attested reanalyses), h-dissimilation does seem to be present in modern Malagasy. This is illustrated in Fig. 1, which shows the distribution of alternants for ka-final stems, by whether or not the consonant nearest to the alternant is [h]. When there is an immediately preceding [h], the observed alternant is always [f]. In contrast, when the stem does not have a preceding [h], only 3% (n = 21/689) stems have [f] as the alternant. Based on these results, dissimilation could have affected reanalyses of ka-final stems.

The data in Table 10 shows that for ʈʂa-final stems, there is a general preference for alternation with [r] (relative to [t] or [f]), such that around 70.2% (231/329) of relevant stems are r-alternating. Fig. 2 shows the proportion of alternants, organized by whether or not there is a preceding [r] somewhere in the stem. From here, it is evident that in modern Malagasy, there is a strong dissimilatory pattern. Specifically, final ʈʂa *never* alternates with [r] if there is already an [r] in the stem. In contrast, when there is no preceding [r], there is a strong preference for alternation with [r]. Overall, the distribution of alternants in modern Malagasy supports the finding that reanalysis in ʈʂa-final weak stems is in the direction of t→r, except when blocked by r-dissimilation.

3.4 Markedness effects in the reanalysis of ʈʂa stems

For the ʈʂa-final weak stems, reanalysis in the direction of t→r cannot be explained by an inductive approach. Additional factors are needed to explain this direction of reanalysis.

I propose that reanalysis towards [r] is the result of a markedness bias in Malagasy against intervocalic oral stops. There is support for the presence of this constraint internal to the Malagasy lexicon. Historically, Malagasy underwent intervocalic lenition which affected all oral stops except for *t (*b > v, *p > f, *d,*ɖ > r, *k,*g > h)(Adelaar, 1989, 2012). As such, it's likely that there were very few intervocalic stops at some point in historical Malagasy.

A constraint against intervocalic stops is also independently motivated cross-linguistically. Studies have found phonetic support for intervocalic lenition, from both an articulatory (Kirchner, 1998) and perceptual (Kaplan, 2010; Katz, 2016) point of view. There is also sizeable typological support for intervocalic lenition at morpheme boundaries, including (among many other examples) Sanskrit stop voicing (Selkirk, 1980), English phrasal tapping (Hayes, 2011, p. 143-144), Korean lenis stop voicing (Jun, 1994), and Catalan fricative weakening (Wheeler, 2005, p. 163). Malagasy ʈʂa∼r alternation fits into this typology, and can be explained as the result of stop lenition at morpheme boundaries.

The fact that only ʈʂa-final stems, and not other weak stems, have undergone reanalysis in a direction not predicted by distributional information, follows naturally from this markedness-based account. For ka-final stems, the possible alternants are [f] and [h]; both are fricatives and would not violate a constraint against medial stops. For na-final stems, the attested alternants are [m] and [n]. Again, neither violate a constraint against medial oral stops, so are equally unmarked if all else is held equal.

One alternative possibility is that speakers are driven by a perceptual bias, rather than a markedness bias (Steriade, 2009 [2001]; Wilson, 2006; White, 2013). That is, if the retroflex affricate [ʈʂ] has a smaller perceptual distance to [r] than to [t], reanalysis towards [r] could be explained as the result of a bias towards perceptually similar alternations.

Although there have been no studies on perceptual distance of Malagasy phonemes, there is indirect evidence from English that [ʈʂ] is perceptually closer to [t] than to [r]. If this is true, than a perceptual distance account predicts that [ʈʂ]∼[t] alternation is preferred over [ʈʂ]∼[r] alternation. English does not phonemically have [ʈʂ] and [r], but Warner et al. (2014) have found that for English, [tʃ] is perceptually closer to [t] than to [ɾ]. If we use [tʃ] and [ɾ] respectively as proxies for Malagasy [ʈʂ] and [r], this would suggest that [ʈʂ] is perceptually more similar to [t] than to [r]. This assumption is not unreasonable because Malagasy [ʈʂ] is variably realized as postalveolar, and [r] is realized as a tap in fast speech (Howe, 2021).[7]

Finally, it is worth noting that the pattern of r-dissimilation, though already present in the distributional information, also has typological support. Suzuki (1998), in a typological study of dissimilation, finds multiple examples of tap dissimilation. More generally, liquid dissimilation

---

[7] There is also evidence of low discriminability between retroflex and coronal affricates ([ʈʂ] vs. [ts]; [ʈʂʰ] vs. [tsʰ]) in Mandarin Chinese, where the two places of articulation are phonemically contrastive (Cheung, 2000; Tsao et al., 2009).

is also crosslinguistically attested, both as a phonotactic tendency and in active phonological processes (e.g. French and Spanish; Colantoni & Steele, 2005).

3.5 Interim summary

Comparison of PMP protoforms with Malagasy suggests that reanalysis of weak stems is driven not just by distributional probabilities of the lexicon, but also by additional markedness effects. Findings of this section are summarized in (10). On one hand, reanalysis of na- and ka-final weak stems is largely predictable from distributional probabilities.

(10) Summary: directions of reanalysis in Malagasy

Type	Pattern	Distributional?
na	m→n	yes
ka	f→h	yes
	h-dissimilation	yes
ʈʂa	t→r	**no**
	r-dissimilation	yes

However, the ʈʂa-final stems underwent reanalysis towards r-alternation, which is opposite of what is predicted by lexical statistics. In other words, a purely inductive model of reanalysis would fail to predict the direction of reanalysis found in Malagasy.

Instead, reanalysis of ʈʂa-final stems is argued to be driven by a markedness constraint against intervocalic stops. In the following section, I propose a model of reanalysis that incorporates a markedness bias, and show that it better captures the Malagasy data than an unbiased model.

Note that for the ka-final weak stems, there is some evidence for h-dissimilation both in the historical distribution and in Malagasy. However, the pattern is hard to confirm due to the lack of evidence; as such, the rest of this paper will not consider effects of h-dissimilation. Additionally, the rest of the paper will focus on the ʈʂa-final weak stems, where the effects of markedness are most pronounced.

4 Modeling reanalysis with a markedness bias

In this section, I test the predictions of the previous section (that reanalysis in Malagasy is driven by both distributional and markedness effects) using a quantitative model of reanalysis. In particular, a constraint-based model of reanalysis which incorporates a markedness bias is compared to baseline control models.

As a preview, results in this section explicitly demonstrate that both distributional and markedness effects are needed to explain the direction of reanalysis found in Malagasy. The model will also make strong, empirically testable predictions about how markedness can influence reanalysis, which can then be applied to other case studies.

The model has three main components. First, it uses Maximum Entropy Harmonic Grammar (MaxEnt; Goldwater & Johnson, 2003; Smolensky, 1986), a probabilistic variant of Optimality Theory. Additionally, to mirror the effect of reanalyses over time, the model will have an iterative (generational) component, in which the output of one iteration of the model becomes the input for the next. Finally, to incorporate markedness effects, a bias is implemented as a Gaussian prior, following the methodology of Wilson (2006) and White (2013, 2017). This biased model will be compared to control models that do not have a markedness bias.

The rest of this section is organized as follows. §4.1 outlines the different components of the grammar, including the inputs and constraint set (§4.1.1-4.1.3), a procedure for implementing markedness bias (§4.2-4.3), and the iterative component of the model (§4.4). Finally, §4.5 compares the markedness-biased model against several control models, to show that a markedness bias significantly improves model performance.

4.1 Components of a MaxEnt model of reanalysis

Because rates of Malagasy weak stem alternation are probabilistic (as opposed to categorical), I adopt MaxEnt, which uses weighted (instead of ranked) constraints and generates a probability distribution over the set of candidate outputs. In principle, other stochastic inductive models of morphophonological learning, such as the MGL (§2.1), would work equally well in matching the Malagasy input data. MaxEnt is adopted because there is existing work on incorporating learning biases in MaxEnt (Wilson, 2006; White, 2013).

Note that unlike classic OT, where strict ranking ensures that losing candidates never surface, all candidates in MaxEnt grammars receive some probability. However, if constraint weights are sufficiently different, MaxEnt produces results that are functionally very similar to classic OT, where the winning candidate gets near-perfect probability.

In all subsequent models, constraint weights were learned in R (R Core Team, 2021), using the *maxent.ot* package (Mayer et al., 2022). Constraint optimization is done using the *optim* function from the R-core statistics library. Constraint weights are restricted to finite, non-negative values.[8]

For explanatory ease, tableaux used to demonstrate the effect of different constraints will be shown in classic strictly ranked OT. However, for the actual model, the output is a set of candidates, each with a predicted probability.

4.1.1 Inputs

The input to the model is a set of 1616 nonce weak stem, designed to represent historical Malagasy, presumably before extensive reanalysis had occurred. The value 1616 was chosen to match the number of weak stems found in the MDEDM (i.e. modern Malagasy) corpus (after removing irregularly alternating forms). Relative frequencies of ka, ʈʂa, and na stems match that of the MDEM corpus. The relative frequency of each alternant was based on the distribution of final consonants in the historical PMP data. Nonce stems are used in place of actual PMP stems because of number PMP forms available is too few.

For simplicity, only candidates with observed alternants are included in the model. A potential alternate like [p], which is in the Malagasy inventory, but not observed as a weak stem alternant, is assumed to be ruled out by highly weighted faithfulness constraints. In addition, ʈʂa ~ f alternating forms and irregular alternants (e.g. na∼f alternating forms) are excluded, because they are extremely low-frequency and do not influence model outcomes. The input data is summarized in Table 11.

The input matches the *surface* stem allomorphs, and the output candidates are suffixed allomorphs. This is because all reanalyses in Malagasy weak stems are from the non-suffixed to suffixed allomorphs. Reanalysis happens in this direction if speakers have access to the surface stem (or another non-suffixed allomorph), but not the suffixed allomorph. The inputs therefore match the conditions under which speakers would reanalyze weak stems.

---

[8] Nearly identical results were found using the Excel Solver (Fylstra et al., 1998), which uses the Conjugate Gradient Descent method.

Input	Candidate	Freq	P
ˈvukiʈ͜ʂa	vuˈkiʈ͜ʂana	0	0
	vuˈkirana	65	0.27
	vuˈkitana	176	0.73
ˈvuritra	vuriʈ͜ʂana	0	0
	vuˈrirana	3	0.04
	vuˈritana	76	0.96
ˈvukika	vuˈkikana	0	0
	vuˈkihana	567	0.81
	vuˈkifana	136	0.19
ˈvukina	vuˈkinana	534	0.9
	vuˈkimana	59	0.1

Table 11: Sample inputs to the Malagasy model of reanalysis

This choice of inputs relies on the assumption that the base of reanalysis is *always* a non-suffixed allomorph. A similar approach is taken by Albright (2008b; 2010, etc.), who argues that the base of reanalysis is fixed, and is always a single slot of a morphological paradigm.

Albright also argues that the base should be the most **informative** allomorph, which has the most contrastive information. The Malagasy base appears to contradict this hypothesis, since it is the suffixed forms that are more informative, and retain contrastive information about weak stem consonant alternations. The Malagasy data may lead us to slightly rethink Albright's hypothesis that informativeness always determines the base of reanalysis. In particular, the base of reanalysis is generally the most informative one (per Albright's hypothesis). However, if learners only have access to limited paradigm slots, reanalyses may still occur from these paradigm slots even if they are not the most informative.

Other factors such as token frequency may also affect how learners select the base of reanalysis. Albright (2008a) suggests that when one slot of the paradigm is used with much higher frequency than others, it may be preferred as the base of reanalysis. However, Keenan & Manorohanta (2001) find, based on written corpora, that actives (unsuffixed) and passives (mostly suffixed) occur at roughly equal rates, making this explanation less likely. Another possible factor is the tendency for bases to be isolation stems or other shorter, 'unmarked' forms (Vennemann, 1972; Kuryłowicz, 1945).

4.1.2 Faithfulness constraints

The model uses the *MAP family of faithfulness constraints, instead of classical feature-based faithfulness constraints (McCarthy & Prince, 1995). *MAP constraints, proposed by Zuraw (2010, 2013), assess violations between pairs of surface forms. A constraint *MAP(a, b) assesses a violation to a candidate if *a* is mapped to a corresponding *b*. The corresponding segments *a* and *b* can differ more than one feature. For example, a constraint like *MAP(k,f), where segments [k] and [f] differ in multiple features ([continuant], [LABIAL], [DORSAL]), is allowed.[9]

The tableau in (11) demonstrates how *MAP violations are assessed for the candidate [ˈvuliʈ͜ʂa]. Candidate (a), where [ʈ͜ʂ] alternates with [t], incurs a violation of *MAP(ʈ͜ʂ, t). Meanwhile, candidate (b), where the alternant is [r], incurs a violation of *MAP(ʈ͜ʂ,r).

---

[9] Zuraw also permits *MAP constraints to include contexts. For the present paper, context-free *MAP constraints suffice.

(11)

ˈvuliʈʂa	*Map(ʈʂ,t)	*Map(ʈʂ,r)
a. vuˈlit-ana	*	
b. vuˈlir-ana		*
c. vuˈliʈʂ-ana		

*Map constraints are more powerful than traditional faithfulness constraints, but are also constrained in substantive terms. Specifically, Zuraw assigns *Map constraints a default weighting (or ranking) based on the **p-map**. The p-map, proposed by Steriade (2009 [2001]), is a language-specific perceptual map which encodes the perceptual distance between all segment pairs in all contexts. *Map constraints which ban changes that cover a larger perceptual distance are assigned a default ranking higher (or weighted more) than constraints banning smaller changes.

In an inductive model of Malagasy, traditional output-output identity constraints actually do just as well as *Map constraints in frequency-matching the input data. However, the current study adopts *Map constraints because they more straightforwardly allow different types of learning bias to be incorporated, and have been more successful at modeling phonetic bias in prior work (Wilson, 2006; Hayes & White, 2015).

4.1.3 Markedness constraints

The inductive model has four markedness constraints. All four constraints are included because they can be learned simply from local distributional information, and would be learned in comparable inductive models of morphophonological learning.

First, the three markedness constraints *ʈʂ]V, *k]V, and *n]V assess violations for every C]V, where C is at a morpheme boundary. These constraints motivate alternation of the final consonant in weak stems. Reference to morpheme boundaries is necessary because within stems, prevocalic ʈʂ, k, and n are allowed.[10] This approach is similar to the one taken by Pater (2007) and Chong (2019) to explain morphologically-derived environment effects (MDEEs), where static phonotactic patterns mismatch the alternations allowed at morphological boundaries.

The effect of *ʈʂ]V is demonstrated in tableau (12); *k]V and *n]V work in parallel ways. ʈʂa-final weak stems always alternate in the suffixed form. This can be achieved by ranking *ʈʂ]V above competing faithfulness constraints (or by giving *ʈʂ]V a much higher weight). As a result, the faithful candidate (c) is eliminated.

(12)

ˈvuliʈʂa	*ʈʂ]V	*Map(ʈʂ,t)	*Map(ʈʂ,r)
a. vuˈlit-ana		*	
b. vuˈlir-ana			*
c. vuˈliʈʂ-ana	*!		

A fourth constraint, *r...r], is used to enforce dissimilation of [r] at the right edge of morpheme boundaries. Again, reference to morpheme boundaries is necessary because within stems, r...r sequences are permitted (e.g. [ˈraraka] ‘spilled’, [buˈrera] ‘weak, limp’, [ˈrirana] ‘edge’). The effect of *r...r] is demonstrated in tableau (13), where the input stem has a preceding [r], *r...r]. In this tableau, highly ranked *r...r] rules out the r-alternating candidate (b).

(13)

ˈvuriʈʂa	*r...r]	*Map(ʈʂ,t)	*Map(ʈʂ,r)
☞ a. vuˈrit-ana		*	
b. vuˈrir-ana	*!		*

---

[10] Examples: be**ʈʂ**oka ‘to swell up’, **ʈʂ**ano ‘box’, fo**k**a ‘smoke, suck in’, a**k**a ‘familiar with’, a**n**ika ‘to climb’

The model laid out so far is inductive, and able to match the input data perfectly ($R^2=1$). However, the goal of the model is not to fit the input data. Instead, given input data that represents Malagasy before reanalysis, it should predict predict the correct direction of reanalysis, and match the distribution of alternants in modern Malagasy. The current inductive model will not be able to do this, as it predicts reanalysis to be in the direction of high frequency alternants (r→t, f→h, m→n). This makes the wrong prediction for ʈʂa-final stems, where reanalysis is in the direction of t→r.

4.2 Learning additional markedness constraints

The central argument of the current study is that reanalysis in Malagasy is partially driven by markedness effects that *cannot* be learned inductively. In this section and the subsequent section, I outline a process for incorporating this markedness component to the model.

First, when we consider markedness bias in reanalysis, it is also important to consider how such effects are constrained—in other words, what is the range of markedness effects that are able to influence reanalysis? I propose that markedness constraints can only affect reanalysis if they are already active in the lexicon, in the form of stem phonotactics.

This "active markedness" proposal is attractive because it ties into existing theories of acquisition and empirical findings about the relationship between phonotactics and morphophonology. First, this approach predicts a strong relationship between phonotactics and alternations. Crosslinguistically, similar phonological generalizations tend to hold within morphemes and across morpheme boundaries; in other words, alternations are consistent with stem phonotactics (Chomsky & Halle, 1968; Kenstowicz, 1996). This is especially true once we consider gradient effects; Chong (2019) shows that even in cases of apparent mismatch between phonotactics and alternations, there is often some gradient phonotactic support for an alternation pattern. Additionally, alternations that are not supported by phonotactics tend to be under-attested.

In work on compound formation, Martin (2011) also finds similar effects of active markedness. In particular, Martin presents evidence from Navajo and English that the same phonotactic constraints present within morphemes are also active in compound formation, albeit as a weaker, gradient effect. In other words, there is evidence that speakers generalize phonotactic constraints across morpheme boundaries. Given Martin's findings, it is conceivable that stem-internal phonotactics could also constrain cross-morpheme alternation patterns.

An active markedness restriction is also consistent with the view that phonotactics guide alternation learning (Tesar & Prince, 2003; Hayes, 2004; Jarosz, 2006), which is supported by experimental evidence (see for example: Pater & Tessier, 2005; Chong, 2021). This restriction also makes empirically testable, language-specific predictions that should be tested in follow-up work, about which markedness effects can affect reanalysis.

For these reasons, I propose that markedness bias is restricted to active markedness effects. As a preview, the Malagasy results are consistent with this active markedness principle. In §5.1, other alternatives are discussed.

To test whether a constraint against intervocalic stops is present in Malagasy phonotactics, I constructed a phonotactic model of Malagasy stems using the UCLA Phonotactic Learner (Hayes & Wilson, 2008), which learns a grammar of n-gram constraints that fits the distribution of natural classes in a set of learning data. The grammar was restricted to learning maximally trigram-length constraints. The UCLA Phonotactic Learner also allows the user to specify different projections, in order to test for long-distance dependencies. The Malagasy phonotactic grammar included two projections, a vowel tier ([+syllabic]) and consonant tier ([-syllabic]). The consonant tier is included to test whether r-dissimilation (and potentially other dissimilative effects) are present in

Malagasy stem phonotactics. The vowel tier is included because, although it is not directly relevant to the current study, there is evidence for vowel dissimilation in Malagasy (Zymet, 2020).

The input to the grammar was 3800 Malagasy stems sampled from the MDEM. Completely reduplicated forms were automatically removed (e.g. pakapaka), but partially reduplicated forms still remain. Only non-suffixed stems were used; suffixed allomorphs were not included because the alternants of weak stems reflect the distribution of the lexicon *after* reanalysis, while the phonotactic grammar is supposed to approximate patterns already present in Malagasy pre-reanalysis.

The resulting grammar learned four constraints, given in (14), that penalize intervocalic stops and specifically favor [r] over [t] as the alternant for ʈʂa-final weak stems. The constraints listed here all motivate reanalysis of t→r. Crucially, they also do not affect the relative preference for different alternants in ka- or na-final weak stems.

(14) *Phonotactic constraints penalizing intervocalic stops*

Constraint	Violations
*[+syll][-cont,-vc][+syll]	V{p,t,ts,dz,ʈʂ,k}V
*[+syll][-son,-cont][+syll]	V{p,b,t,d,ts,dz,ʈʂ,ɖʐ,k,g}V
*[+syll][-tap,-nasal,+coronal][+syll]	V{t,d,ts,dz,ʈʂ,ɖʐ,s,z,l}V
*[+syll][-son,-cont,-labial][+syll]	V{t,d,ts,dz,ʈʂ,ɖʐ,k,g}V

In general, the phonotactic grammar also learned constraint weights in a way that favored r-alternating candidates over t-alternating candidates. This is demonstrated in (15), which shows the Harmony scores assigned by the phonotactic grammar to suffixed forms of hypothetical weak stems. The higher the Harmony, the more a form is penalized by the grammar and phonotactically dispreferred.

For the ʈʂa-final weak stems, the grammar assigns the lowest harmony to the r-alternating candidate ([vuki**r**-ana]). Notably, for the na- and ka- final weak stems, the phonotactic grammar also assigns harmony scors that are either neutral or favor the statistically preferred alternant. Specifically, for ka-final weak stems, the grammar assigns very similar Harmony scores to all three candidates. For the na-final weak stems, the grammar assigns a higher Harmony to the m-alternating candidate, which is statistically dispreferred.

(15) *Harmony scores assigned by phonotactic grammar to suffixed form candidates*

stem	**suffixed**	**H**
vuki**ʈʂ**a	vuki**ʈʂ**ana	13.8
	vuki**t**ana	13.3
	vuki**r**ana	12.3
vuki**k**a	vuki**k**ana	13.1
	vuki**h**ana	13.3
	vuki**f**ana	13.3
vuki**n**a	vuki**n**ana	13.2
	vuki**m**ana	14.3

For simplicity, I added only the constraint *V[-cont,-voice]V to the model of reanalysis. Although the phonotactic grammar found multiple constraints which penalize intervocalic stops, I included only one because all four constraints have the same violation profile with respect to the candidates in weak stem reanalysis.

Alternation in ʈʂa-final weak stems is also driven by a strong r-dissimilation constraint. The phonotactic grammar did not learn this constraint in the consonant tier; other projections that were tested, such a Coronal tier, also did not learn a constraint for r-dissimilation. Constraints on dissmilation of larger classes of segments (e.g. approximants) were also found to be non-significant. As such, r-dissimilation differs from lenition in that it is a markedness constraint learned from the local distribution of weak stem alternants, and does not receive additional phonotactic support.

In other words, although *r...r] and *V[-cont,-voice]V look similar on the surface, they have different underlying mechanisms. Reanalysis driven by *V[-cont,-voice]V is supported by stem phonotactics. In contrast, reanalysis driven by *r...r] is better characterized as frequency-matching of patterns within the weak stem paradigm.

4.3 Incorporating a soft markedness bias

The constraint *V[-cont,-voice]V is added to the model, and assigned a bias towards higher weight. Following Wilson (2006) and White (2017), a bias term, or 'prior', is implemented as a Gaussian distribution over each constraint weight. The bias term, calculated as in (16), is defined in terms of a mean (μ) and standard deviation (σ). For each constraint, ***w*** is its learned weight, and μ can be thought of as the 'preferred' weight. As such, the numerator of the bias term reflects how much the actual weight deviates from the preferred weight of each constraint, and the penalty resulting from the bias term increases as constraint weights diverge from μ.

(16) $$\sum_{i=1}^{m} \frac{(w_i - \mu_i)^2}{2\sigma^2}$$

The value of σ^2 determines how much effect the preferred weight (μ) has; lower values of σ^2 result in a smaller denominator, and therefore greater penalty for weights that deviate from their μ. In unbiased models, the goal of learning is to maximize log probability. With the inclusion of the prior, the goal becomes to maximize a different OBJECTIVE FUNCTION, which is the prior term subtracted from the log probability of the observed data.

In principle, both μ and σ^2 can be varied to give constraints a preference towards a certain weight. In the current models, σ^2 is set to fixed values. The markedness constraints *ʈʂ]V, *k]V, *n]V, and *r...r] have no phonotactic support from the lexicon, but are supported by distributional information within the paradigm. For these constraints, I assume that the weight is learned from the input data, and that the effect of a bias is negligible. This is done by setting σ^2 to an arbitrarily high value (1000).

For the rest of the constraints, σ^2 is set to 0.5, and μ is varied to implement different learning biases. For example, a markedness bias is implemented by assigning *V[-cont,-voice]V a higher μ than competing faithfulness constraint(s). As a result, *V[-cont,-voice]V will be biased to have a higher weight than the relevant faithfulness constraints. In §4.5, I provide the specific μ values used for the markedness-biased model, as well as the μ values of baseline control models.

4.4 Iterated modeling

To simulate reanalysis over time, I use a generational model, in which the output of one iteration of the model becomes the input to the next iteration. Similar models of language change are by no means new, and there are various approaches to doing so. For example, Weinreich et al. (1968) use phonological rules that apply variably to predict change in progress. Other approaches that have been explored include modeling change in dynamical systems (Niyogi, 2006), connectionist frameworks (Tabor, 1994), as the result of competing grammars (Yang, 1976), in exemplar-based frameworks (Pierrehumbert, 2002), and more recently in variants of OT (e.g. Boersma, 1998; Zuraw, 2000, 2003; Staubs, 2014).

To simulate the cumulative effects of reanalysis over time, I assume an agent-based iterated learning model. Under this approach, small changes to an alternation pattern can accumulate over iterations (each corresponding roughly to a generation of speakers), resulting in large-scale reanalyses of a pattern.

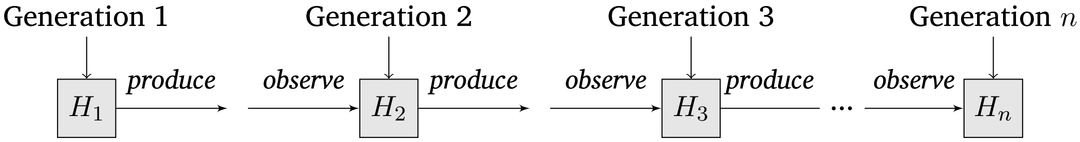

Figure 3: Structure of an iterated learning model, adapted from Ito & Feldman (2022, p. 3). H_i indicates hypotheses of each generation.

In an agent-based iterated model, the output of one model iteration becomes the input to the next iteration. The current study adopts a simplified model in which each generation (or iteration) has just one agent and one learner, as illustrated in Fig. 3. In the first generation, the agent A1 produces the output language based on their grammatical knowledge (i.e. Hypothesis 1; H_1). More concretely, a hypothesis is the speaker's grammar, represented in this case using MaxEnt, as the probabilistic weighting of Optimality-Theoretic constraints. The learner observes these data, induces the relevant generalizations and forms another hypothesis (H_2), which then becomes the basis of the output data presented to the next generation. This process is repeated for many iterations.

When providing input for a learner in the next generation, not all of the information of the language is presented, resulting in a learning "bottleneck" (Brighton, 2002; Kirby, 2001; Griffiths & Kalish, 2007). As a result of this bottleneck, input patterns that are easier to learn should be more likely to pass through this bottleneck, and become more prominent over generations of learning. In the current study, the bottleneck is implemented by having the Agent "forget" some proportion of forms at each iteration. The remembered forms are retained to the next generation, while the forgotten forms are generated from the Agent's grammar (Hypothesis 1, 2, 3, etc.). A similar approach is taken up by Ito & Feldman (2022), who use iterated learning to model accent change in Sino-Korean.

Note that the iterated learning paradigm I adopt makes several simplifying assumptions. In particular, I assume just one agent and one learner, when in fact language change takes place at the level of the population. Future work should therefore consider more complex models which incorporate multiple interacting Agents in a way that models the speech community. Baker (2008) finds that such multi-agent models produce more empirically accurate results.

The iterated learning model has two parameters: forgetting rate and number of iterations. The forgetting rate is the proportion of forms forgotten and relearned in each iteration. I test 5 forgetting rates (0.05, 0.1, 0.15, 0.2, 0.25). In the interest of clarity, and because the model trended in the same direction across all five forgetting rates, the rest of this paper will only present models with a forgetting rate of 0.2.

The number of iterations is set to 50 for several reasons. First, I follow Ito & Feldman (2022) in equating each iteration to roughly one generation of speakers, where a generation lasts 25 years. The number of iterations was then chosen to reflect the maximal span of time in which reanalyses of weak stems could have occurred. The sound changes that resulted in weak stem alternations took place around 600 AD, while the modern Malagasy data starts around the 1800s. Therefore, reanalysis must have occurred within the span of around 1200 years. This corresponds to roughly 50 generations, assuming that each generation is 25 years. 50 generations is meant as a conservative estimate, since in reality, reanalysis of the ʈʂa-final weak stems may have happened in a much shorter span of time.

Forgetting rate and the number of iterations are closely related; in general, when the forgetting rate is low, rate of change over time is slower, but this can be offset by increasing the number of iterations. However, increasing the forgetting rate has the additional effect of increasing variation between different runs of the model. This is because as forgetting rate increases, the input data for each model iteration becomes more variable.

Because random sampling causes each iteration of the model to vary slightly, all subsequent models were run 30 times, and predicted probability values are the mean of these 30 trials.

4.5 Model comparison

This section compares markedness biased models against controls, to evaluate the effect of markedness in improving model predictions. Although it is not the focus of the current paper, models with a p-map bias are also tested. These models are included to confirm that markedness effects improve model predictions after controlling for perceptual similarity effects, which have been substantiated by prior research (White, 2013, 2017).

A total of four models are compared: the first two conditions, FLAT-PRIOR and P-MAP, are controls. They are compared to two conditions with a markedness bias, labeled MARKEDNESS and FULL (which includes both a markedness and p-map bias). The priors assigned to each condition are explained below, and summarized in Table 12.

If reanalysis is in fact driven by a markedness bias in Malagasy, then the MARKEDNESS and FULL models should outperform their respective control conditions, FLAT-PRIOR and P-MAP. If, instead, reanalysis is rooted in a p-map bias, adding a markedness bias should not improve model fit. Instead, the P-MAP condition (and potentially the FULL condition) should perform better than the FLAT-PRIOR model, and the FULL condition should not perform better than the P-MAP condition.

FLAT-PRIOR condition (control). The FLAT-PRIOR model (labeled FLAT in Table 12) is a control condition. In this condition, every constraint with a bias term has the same μ of 3.3, which is the mean of all μ values assigned to the *MAP constraints in the P-map condition below. This condition will be compared against the MARKEDNESS condition. It is included because as discussed in White (2013), a model with uniform (but non-zero) μ values is a better control than a model with no bias terms at all.[11]

P-MAP condition (control). The p-map condition (labeled P in Table 12) has a bias towards higher-weighted faithfulness constraints, scaled by perceptual similarity. The μ of *MAP constraints is higher for mappings between perceptually dissimilar sounds, and lower for mappings between perceptually similar sounds. In addition, all markedness constraints are assigned $\mu = 0$.

To approximate perceptual similarity, I adopt White's (2013; 2017) method of using confusability as a measure of perceptual similarity, where the confusability of two speech sounds is determined according to the results of standard identification experiments.[12] As there are no confusability experiments for Malagasy, I use results from Warner et al. (2014), a study of consonant

---

[11] Note that this model essentially has a smoothing term which serves only to prevent model overfitting. The smoothing term penalizes models with a few closely-fitted constraints, and instead prefers for weight to be more evenly distributed across constraints.

[12] Specifically, confusability values are used to train a separate MaxEnt model, whose weights become the priors for the main model. Details of implementation are given in (White, 2013, 2017).

confusability in English, as a proxy.[13] English [ɾ] is used in place of Malagasy <r> [r∼ɾ]. Additionally, English does not have a retroflex affricate (except allophonically when [t] precedes [ɹ]), so [tʃ] is used as a substitute for [ʈʂ̻].

Markedness condition. The Markedness condition (labeled M in Table 12) assigns a uniform prior, $\mu = 3.3$, to all faithfulness constraints. The markedness constraint *V[-cont,-voice]V is assigned a high prior ($\mu = 7$). This value is higher than the μ assigned to the competing faithfulness constrain *Map(ʈʂ̻,r), but is otherwise arbitrary. This condition differs from the flat-prior condition *only* in the μ value assigned to *V[-cont,-voice]V; the two models are otherwise identical.

Full condition. The Full condition has both a markedness bias and a p-map bias. Like the markedness condition, *V[-cont,-voice]V is assigned a μ value of 7. The P-map and Full conditions are identical except for the μ values assigned to *V[-cont,-voice]V.

Constraint	σ^2	μ			
		flat	P	M	full
*ʈʂ̻]V	1000	0	0	0	0
*k]V	1000	0	0	0	0
*n]V	1000	0	0	0	0
*r...r]	1000	0	0	0	0
*Map(tr,r)	0.5	3.3	5.13	3.3	5.13
*Map(tr,t)	0.5	3.3	2.82	3.3	2.82
*Map(n,m)	0.5	3.3	1.83	3.3	1.83
*Map(k,f)	0.5	3.3	2.76	3.3	2.76
*Map(k,h)	0.5	3.3	3.3	3.3	0
*V[-cont,-vc]V	0.5	3.3	0	7	7

Table 12: Constraints and bias terms by condition (P = p-map condition, M = markedness condition)

Note that the Flat-prior condition does bias learners slightly in favor of ʈʂ̻a∼r alternation.[14] To see why, we can consider the constraints that respectively enforce ʈʂ̻a∼t and ʈʂ̻a∼r alternation. *Map(ʈʂ̻,r) enforces ʈʂ̻a∼t alternation, while both *V[-cont,-vc]V and *Map(ʈʂ̻,t) enforce ʈʂ̻a∼r alternation. The Flat-prior condition gives all three constraints the same prior weight, and will therefore prefer an outcome where the two constraints that enforce ʈʂ̻a∼r alternation have a higher combined weight than *Map(ʈʂ̻,r). As will be seen in the rest of the section, however, the magnitude of t→r reanalysis predicted by the Flat-prior condition is too small to match the amount of reanalysis that has occurred between PMP and Malagasy.

4.5.1 Model results after one iteration

Table 13 shows results after one model iteration. The column titled 'Obs (PMP)' shows the observed probability of the input candidates, and reflects the distribution of alternants before reanalysis. The column 'Obs (Mal)' reflects the distribution of alternants in modern Malagasy, *after*

---

[13] I use Warner et al. (2014) because unlike other studies of English consonant confusability (e.g. Wang & Bilger, 1973; Cutler et al., 2004), it tests for confusability of phonemes with [ɾ].

[14] Thank you to an anonymous reviewer for pointing out this important detail.

reanalysis. Due to reanalysis of ʈʂa-final forms in the direction of t→r (see §3), modern Malagasy shows a much higher rate of ʈʂa∼r alternation than PMP.

Results in the control conditions (FLAT-PRIOR and P-MAP) are comparable. Both match the frequencies of the input data closely. The two conditions with a markedness bias perform essentially the same. Both predict an increase in the probability of [vuˈkirana] (by 4%), and therefore reanalysis to be in the direction of t→r. In other words, adding a markedness bias does appear to improve model predictions. The magnitude of change is relatively small after one iteration of the model. However, as seen in the following section, the model will approach the distribution seen in modern Malagasy after multiple iterations.

Input	Cand	Obs (PMP)	Obs (Mal)	Predicted			
				FLAT	P-MAP	MARK	FULL
vukiʈʂa	vukirana	0.27	0.95	0.27	0.27	**0.31**	**0.31**
	vukitana	0.73	0.05	0.73	0.73	**0.69**	**0.69**
	vukiʈʂana	0	0	0	0	0	0
vuriʈʂa	vurirana	0.04	0	0.03	0.03	0.03	0.03
	vuritana	0.96	1	0.97	0.97	0.97	0.97
	vuriʈʂana	0	0	0	0	0	0
vukika	vukikana	0	0	0	0	0	0
	vukihana	0.81	0.95	0.81	0.81	0.80	0.81
	vukifana	0.19	0.05	0.19	0.19	0.20	0.19
vukina	vukinana	0.90	0.98	0.90	0.90	0.90	0.90
	vukimana	0.10	0.02	0.10	0.10	0.10	0.10

Table 13: Predicted probability of models after one iteration (mean of 30 trials)

4.5.2 Model results after 50 iterations

Table 14 shows the constraint weights learned by each model after 50 iterations. Because each model was run 30 times, these weights are averaged over 40 runs. Additionally, Table 15 shows the proportion of variance explained (adjusted R^2) and log likelihood ($\hat{L}$) for each model after 50 iterations, fit to the modern Malagasy distribution. Log likelihood was calculated by fitting model predictions to the frequency counts of different weak stem alternants in Malagasy (given in §3.3).

	FLAT	P-MAP	MARK	FULL
*ʈʂ]V	7.53	5.55	7.08	8.53
*k]V	4.24	5.59	6.51	8.21
*n]V	1.08	0.79	0.76	0.77
*MAP(ʈʂ,r)	2.37	3.23	3.40	4.35
*MAP(tr,t)	0.17	0.16	0.05	0.00
*MAP(n,m)	3.33	2.70	2.92	2.72
*MAP(k,f)	1.41	1.63	1.26	2.20
*MAP(k,h)	0.13	0.03	0.02	0.01
*V[-cont,-vc]V	1.87	2.36	6.16	7.10
*r...r]	4.26	2.94	5.61	5.81

Table 14: Model predicted weights after 50 iterations (mean of 30 trials)

Condition	**R^2**	**($\hat{L}$)**
Flat	0.70	-516.7
P-map	0.60	-570.6
Markedness	0.97	-361.1
Full	0.99	-303.9

Table 15: Results after 50 iterations: Proportion of variance explained (adjusted R^2) and log likelihood ($\hat{L}$), of model predictions fit to modern Malagasy

Based on Table 15, the models with a markedness bias clearly outperform corresponding control conditions. Between the two control models, the Flat-prior does slightly better than the P-map model. Interestingly, when comparing between the two markedness-biased models, the Full model does slightly better than the Markedness model (in terms of both R^2 and log-likelihood). The models differ primarily in the weights they learn for *V[-cont,-vc]V. In particular, the markedness-biased models (Markedness and Full) both learn a higher weight for *V[-cont,-vc]V than for *Map(ʈ͜ʂ,r), and will therefore prefer ʈ͜ʂa∼r alternation over ʈ͜ʂa∼t alternation.

Fig. 4 compares the proportion of variance explained (adjusted R^2) in the four conditions over 50 iterations. As seen in this figure, the model fit of the flat-prior control model improves only slightly over the 50 iterations (R^2 = 0.70). In the p-map control model, model fit does not really improve over iterations (R^2≈0.6). In contrast, both the markedness and full are able to account for over 97% of the variation in the observed Malagasy data, and achieve this high model fit by around 30 iterations.

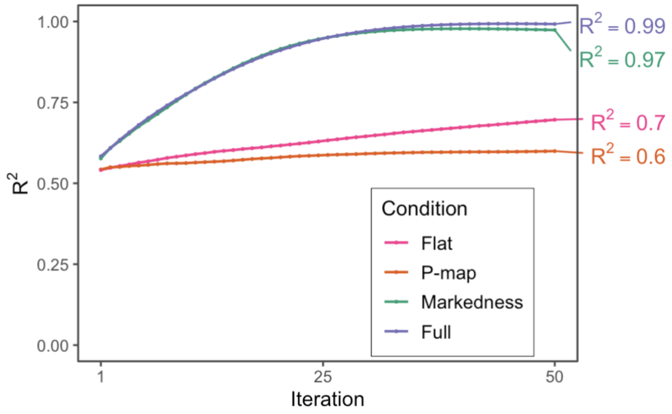

Figure 4: Model fit (adjusted R^2) by condition over 50 iterations (mean of 30 trials)

Overall, adding a p-map bias does not strongly affect model fit, as the P-map condition actually performs worse than the Flat-prior condition. However, the full model ($\hat{L}$=–303.9) actually performs slightly better than the markedness model ($\hat{L}$=-261.1). In other words, adding a p-map bias on top of a markedness bias does slightly improve model fit. This is because, as will be discussed below, adding a p-map bias improves predictions for the ka-final weak stems.

A more detailed examination of model predictions shows that the bulk of improvement in model fit is driven by changes to ʈ͜ʂa-final weak stems. Consider Fig. 5, which plots the change in predicted probabilities over 50 trials for ʈ͜ʂa-final weak stems. Rates of alternation in the input data (PMP) and modern Malagasy (Mlg) are given at the endpoints of the x-axis for reference. The candidates labeled with “(r...)” have a preceding [r] in the stem; for example, “(r...)ʈ͜ʂ∼t” refers to input-output pairs like [ˈvuriʈʂa]∼[vuˈritana]. Non-alternating candidates (e.g. [ˈvukiʈʂa∼vuˈk-iʈʂana]) are not shown, since they are never observed in either PMP or Malagasy, and are consistently assigned zero or near-zero probabilities by all models.

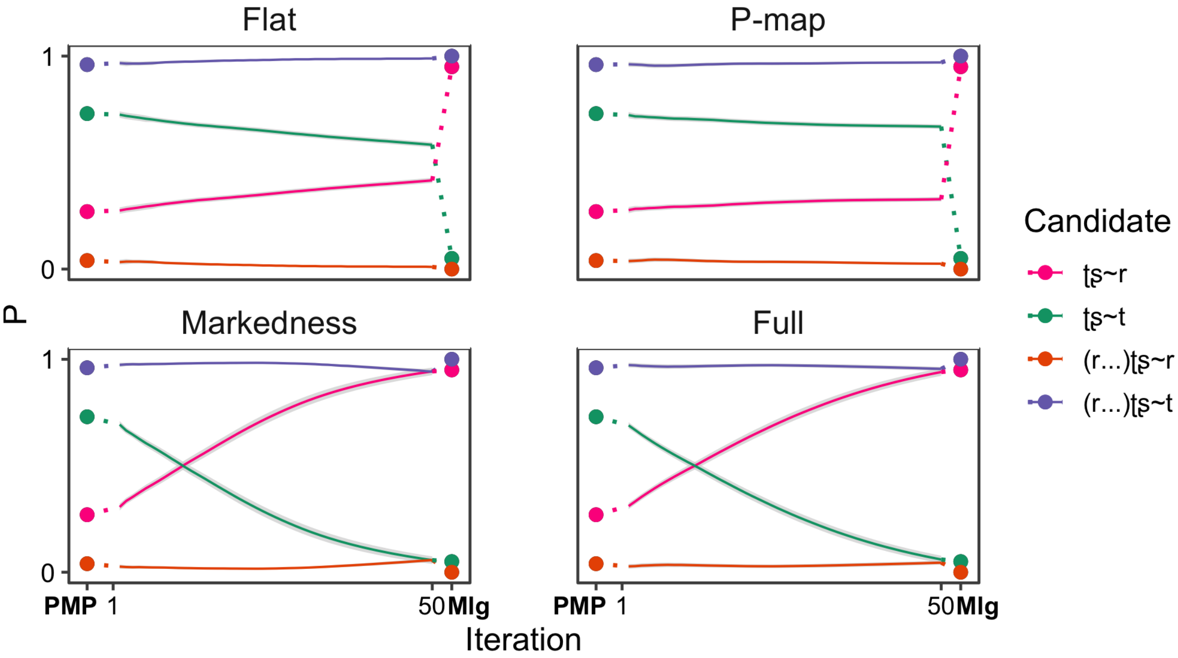

Figure 5: Predicted probabilities of candidates over 50 iterations for ʈʂa-final weak stems (mean of 30 trials). Grey intervals indicate standard error, and observed rates of alternation in PMP and Malagasy are given for reference.

In the two conditions with a markedness bias, the model successfully predicts an increase in the ʈʂ∼r alternating candidate, and therefore closely matches the Malagasy data. At the same time, for inputs with a preceding [r], where r-dissimilation should block the r-alternating candidate, all four models do similarly well and predict the t-alternating candidate at near-exceptionless rates.

The FLAT-PRIOR model also predicts some reanalysis in the direction of t→r. This is because, as discussed above, this model assigned the same μ to *MAP(ʈʂ,r), *MAP(ʈʂ,t), and *V[-cont,-vc]V. This means that the combined μ values of *MAP(ʈʂ,t), and *V[-cont,-vc]V, which both enforce r-alternation, will be greater than the μ of *MAP(ʈʂ,r) (which enforces t-alternation). However, the magnitude of reanalysis predicted by the FLAT-PRIOR model is too small; after 50 iterations, it still predicts a higher rate of ʈʂa∼t alternating forms than ʈʂa∼r alternating forms.

For the na-final weak stems, all four models perform similarly well. This is demonstrated in Fig. 6, which plots the change in predicted probabilities over 50 trials na-final weak stems. In both the historical and modern distributions, there is a strong preference for n-alternation; all four models can capture this pattern. These results show that the markedness-biased models are able to predict frequency-matching in environments where markedness is neutral (i.e. where all alternants are equally marked).

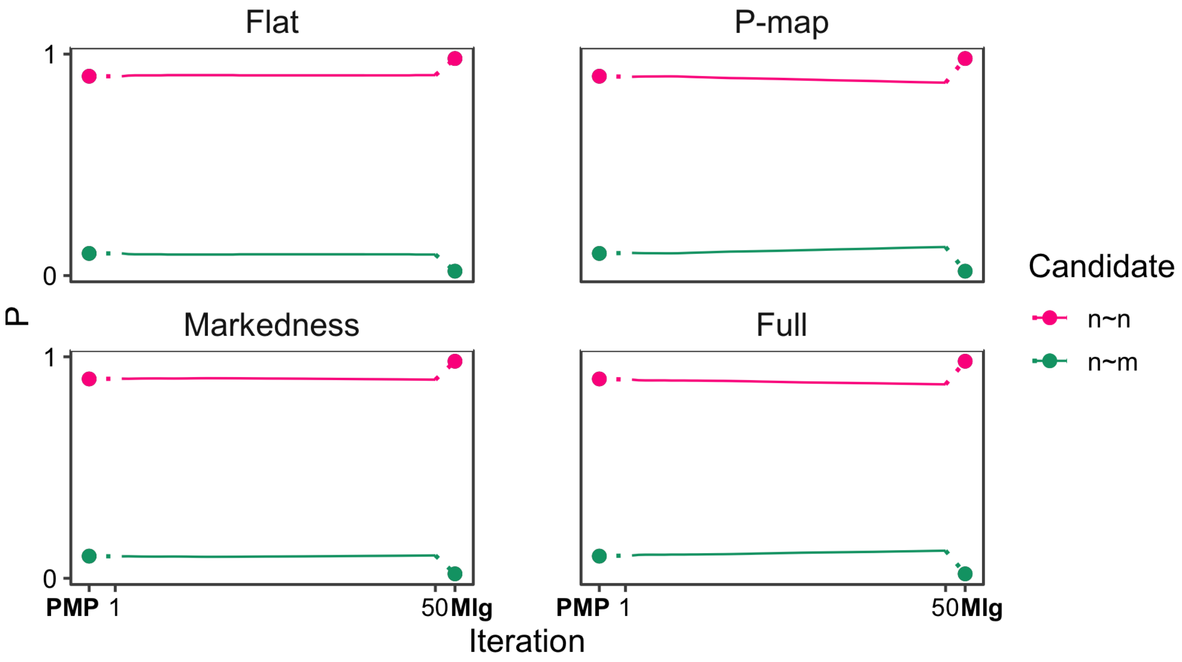

Figure 6: Predicted probabilities of candidates over 50 iterations for ka-final weak stems.

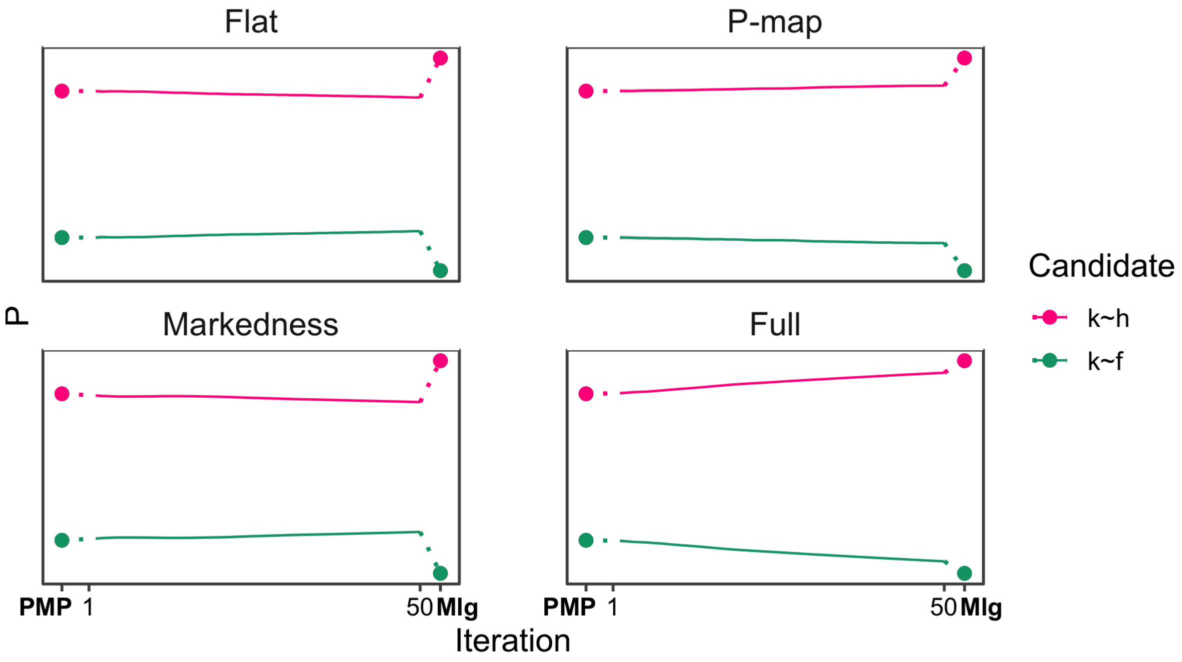

Figure 7: Predicted probabilities of candidates over 50 iterations for ka-final weak stems.

Fig. 7 shows results for ka-final weak stems. Between PMP and Malagasy, there is a slight increase in the rate of k∼h alternation (from P = 0.81 in PMP to P = 0.95 in Malagasy). Notably, the models with a p-map bias (P-MAP and FULL) are able to match this pattern, while the other two models predict roughly stable rates of alternation that match the PMP distribution. The FULL model, in particular, predicts the most increase in k∼h alternation. The P-MAP and FULL models do well because the p-map bias assigns a higher μ to *MAP(k,f) than *MAP(k,h), motivating higher rates of ka∼h alternation.

Table 16 shows the detailed predictions of each condition on the 50th iteration. The two control models (FLAT-PRIOR and P-MAP) generally match the historical PMP distribution, and therefore under-predict rates of ʈʂa∼r alternation. Although the FLAT-PRIOR model does predict a slight increase in ʈʂa∼r, it still does not come close to matching the Malagasy pattern. In contrast, both the MARKEDNESS and FULL conditions predict a large magnitude of reanalysis in the direction of t→r, and assign the r-alternating candidate (vuki**r**ana) a high probability ($P = 0.94$ in both models).

As previewed above, the FULL model actually does better than the MARKEDNESS model for ka-final weak stems. In particular, it predicts higher rates of k∼h alternation ($P_{FULL} = 0.90$ vs. $P_{MARK} = 0.78$). This explains why the FULL model does slightly better than the MARKEDNESS model in terms of overall model fit (as measured by R^2 and log-likelihood).

Input	Cand	Obs (PMP)	Obs (Mal)	Predicted			
				FLAT-PRIOR	PMAP	MARK	FULL
vukitra	vukirana	0.27	0.95	0.42	0.33	**0.94**	**0.94**
	vukitana	0.73	0.05	0.58	0.67	**0.06**	**0.06**
	vukitrana	0	0	0	0	0	0
vuritra	vurirana	0.04	0	0.01	0.03	0.06	0.04
	vuritana	0.96	1	0.99	0.97	0.94	0.96
	vuritrana	0	0	0	0	0	0
vukika	vukikana	0	0	0	0	0	0
	vukihana	0.81	0.95	0.78	0.83	**0.78**	**0.90**
	vukifana	0.19	0.05	0.22	0.17	**0.22**	**0.10**
vukina	vukinana	0.9	0.98	0.91	0.87	0.90	0.90
	vukimana	0.1	0.02	0.10	0.13	0.10	0.10

Table 16: Predicted probability of models after 50 iterations (mean of 30 trials)

Overall, model results support the hypothesis that reanalysis in Malagasy weak stems is largely driven by a markedness bias which penalizes intervocalic stops. Additionally, comparison of the MARKEDNESS and FULL models shows that a perceptual bias (when combined with a markedness bias) improves model fit. Where alternants are equally marked, such as with the na-final weak stems, both of the markedness-biased models are also able to match the frequencies of the input data.

4.6 Iterated learning and the choice of σ^2

In the current model, σ^2 is set to 0.5, which allows for the bias to have a small magnitude of effect that adds up over multiple iterations. By the 10th iteration, the model closely matches the rates of alternation observed in modern Malagasy.

A superficially similar outcome can be achieved by removing the generational component of the model, and simply setting σ^2 to a lower value. A lower σ^2 allows the bias to have a stronger effect, so that the model predicts a greater magnitude of change in just one iteration. Fig. 8 shows the model fit over 50 iterations for the FULL model when σ^2 is varied, and μ values are held constant. Both the high-sigma model ($\sigma^2 = 0.5$) and low-sigma model ($\sigma^2 = 0.1$) converge on the same outcome, but the low-sigma model does so much faster, after just 1-2 iterations.

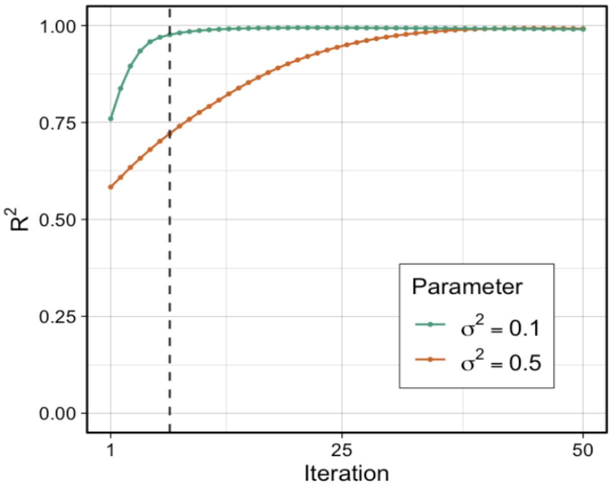

Figure 8: R^2 over 50 iterations of the FULL model, when σ^2 is varied

Although a low-sigma model achieves the same outcome as an iterative high-sigma model, I argue that the multi-generational model is preferable for the following reasons. First, it is conceptually more plausible that reanalysis happens gradually. This is especially true for a case like Malagasy, where the reanalysis of t→r cannot be attributed to sound change, and both alternants are phonemic.

A generational model also predicts randomness and variation; in the current paper, this comes from randomly sampling the winning candidate that becomes the input to the next model iteration. This matches how language change happens in reality, where markedness bias may affect different languages to a different degree, and the same language will undergo dialect divergence.

5 Discussion

5.1 Sources of markedness effects in reanalysis

Throughout this study, I have proposed that active markedness effects in reanalysis are restricted to so-called 'active markedness' effects already active in the stem phonotactics. In other words, learners notice a phonotactic tendency and use it to guide reanalysis. This approach is attractive for the reasons discussed in §4.2; namely, it ties into work showing that people tend to acquire phonotactics before alternations, and use phonotactics to aid in the learning of alternations.

Within work on language change, findings from Garrett (2008) also support the idea that markedness-motivated paradigm reanalyses are a product of language-specific factors rather than a direct manifestation of UG. While Garrett's focus is on semantic (rather than phonological) markedness patterns, his findings still provide support for the idea that reanalysis is driven by markedness effects already present in the language.

Notably, although the Malagasy results are consistent with the active markedness principle, but also amenable to other analyses. One alternative is that external factors have resulted in the bias against intervocalic stops. For example, this bias could be rooted in principles of phonetic naturalness; that is, speakers are biased against intervocalic [t] because it is harder to produce or perceive. Alternatively, sound changes specific to Malagasy made have made intervocalic [t] harder to produce or perceive at some point in the history of the language. This is likely because (as discussed in §3.4), Malagasy underwent multiple intervocalic lenition processes, which affected all oral stops except for [t]. Future work should expand on the typology of markedness effects in reanalysis, to confirm whether the active markedness principle holds true crosslinguistically.

5.2 When can markedness-driven reanalysis occur?

My proposal, broadly speaking, is that reanalysis should be phonologically optimizing. The active markedness principle (the idea that speakers draw on stem phonotactics when reanalyzing paradigms), in particular, predicts that reanalysis will result in a close correspondence between stem-internal phonotactics and cross-morpheme alternations. Importantly, this type of markedness-driven reanalysis only comes into play when there is *uncertainty* in an alternation pattern. In other words, markedness effects in reanalysis are only observed when there is conflicting evidence for which alternant should surface, and one alternant is less marked than the competing alternants.

This distinction is important because it allows mismatches between phonotactics and alternations to persist if an alternation pattern is predictable. In Malagasy, given a ʈʂa-final weak stem, there is generally ambiguity in whether the alternant will be [t] or [r]. This uncertainty allowed a constraint against intervocalic stops (specifically intervocalic [t]) to affect reanalysis. In contrast, for the subset of ʈʂa-final weak stems with a preceding [r], there was near-exceptionless distributional evidence that the alternant should be [t]. In these cases, where the alternation pattern had less uncertainty, the r-dissimilation pattern was able to persist even even in the absence of phonotactic support.

More generally, there is crosslinguistic evidence that phonotactics-alternation mismatches can persist in a language. For example, Turkish vowel harmony operates within stems but not across compounds or phonological words (Kabak & Vogel, 2001); see also Gouskova (2018) for an overview of similar mismatches. Experimental evidence from Gallagher et al. (2019) also supports the idea that speakers are able to learn different cross-morpheme and morpheme-internal phonotactic generalizations.

Relatedly, morphophonological patterns which are not phonologically optimizing can also persist if the relevant pattern is predictable. In particular, there is crosslinguistic evidence for *phonologically conditioned suppletive allomorphy*, or cases where allomorphy has clear phonological conditioning but is not output-optimizing (Paster, 2005, 2009). For example, in Tzeltal, the perfective allomorph that surfaces (*-eh* vs. *-oh*) depends on how many syllables the stem has, in a way that is not output-optimizing.

In summary, although my proposal of markedness-driven reanalysis predicts a strong connection between within-morpheme and cross-morpheme phonotactics, it is also consistent with cases of mismatch because reanalysis occurs only when there is uncertainty in the morphophonology.

6 Conclusion

The current paper looked at reanalysis in Malagasy weak stems, and found that for the ʈʂa-final stems, the direction of reanalysis cannot be predicted by local distributional information. Instead, I argue that reanalysis of t→r is motivated by a markedness constraint against intervocalic (voiceless) stops. This markedness constraint is typologically well-motivated, and also present in the Malagasy lexicon as a phonotactic tendency. Based on these results, I outline a model of reanalysis with a markedness learning bias. This model outperformed control models and was able to closely match the Malagasy data.

From a modeling perspective, the results of this study show that in iterated models where a markedness constraint is biased above a faithfulness constraint, the structure that violates that constraint is likely to be lost over iterations. In the case of Malagasy, suffixed forms which violated a constraint against intervocalic stops were more likely to be reanalyizedd. This ties into other work on iterated modeling, where a learning bottleneck makes the learner more likely to forget structures that are difficult to learn (e.g. Brighton, 2002; Kirby, 2001; Griffiths & Kalish, 2007). In particular, iterated implementations of MaxEnt have similarly found that biased learning, combined with iterated modeling, can be used to model the emergence of unmarked phonological structures (Staubs, 2014; Hughto, 2018, 2020; O'Hara, 2022).

In the current study, I focus on the Official Malagasy dialect. In future work, a comparative analysis of different dialects may also give us insight into the development and reanalysis of weak stems. In particular, different dialects may show different degrees of reanalysis, giving us insight into intermediate levels of change. Where dialects diverge, this could also tell us about how much markedness effects may vary, and how this variation is restricted; a model of reanalysis should ideally be able to capture the range of possible variation.

The approach to incorporating markedness laid out in this study makes empirical predictions about which markedness effects can affect reanalysis. Specifically, I argue that the markedness effects affecting reanalysis are restricted, and must already present in a language's phonological grammar. In the case of Malagasy, the relevant constraint *V[-cont,-voice]V was found to have significant weight in a phonotactic grammar.

To model reanalysis, I adopt a batch learner with a learning bias. However, reanalysis could potentially also be modeled in online variants of MaxEnt (e.g. Perceptron; Rosenblatt, 1958; Boersma et al., 2016). Online implementations of MaxEnt capture learning biases using initial weighting conditions (i.e. by changing the starting weights of each constraint), in a way that can approximate the prior in batch learners. Work such as O'Hara (2020) shows that batch and online learners can differ in subtle ways. As such, future work should consider where the predictions of the two approaches diverge, and which one is a better predictor of reanalysis.

Finally, a model which fully captures reanalysis would be more complex than the one developed here, and should be explored in future work. For one, the current model ignores factors such as usage frequency (Bybee, 2003), and assumes that bias factors remain the same over iterations of the model.

In addition, the current model assumes surface-base representations, where surface stem allomorphs are the inputs. However, reanalysis in Malagasy is also potentially compatible with a model of base competition, in which outputs are faithful to multiple listed allomorphs, but also sensitive to markedness effects (Breiss, 2021). Future work will consider how different parameters can be varied in modeling reanalysis, as well as how input forms should be represented.

References

Adelaar, Alexander (1989). Malay influence on Malagasy: Linguistic and culture-historical implications. *Oceanic Linguistics, A Special Issue on Western Austronesian Languages (Summer, 1989)* **28:1**. 1–46.

Adelaar, Alexander (2009). Malagasy. In Martin Haspelmath & Uri Tadmor (eds.), *World loanword database (WOLD)*, Max Planck digital library.

Adelaar, Alexander (2012). Malagasy phonological history and Bantu influence. *Oceanic Linguistics* **51:1**. 123–159.

Adelaar, Alexander (2013). Malagasy dialect divisions: Genetic versus emblematic criteria. *Oceanic Linguistics* **52:2**. 457–480.

Adelaar, Alexander K (1994). Malay and Javanese loanwords in Malagasy, Tagalog and Siraya (Formosa). *Bijdragen tot de taal-, land-en volkenkunde* **150:1**. 50–66.

Albright, Adam (2008). Explaining universal tendencies and language particulars in analogical change. *Linguistic universals and language change* 144–181.

Albright, Adam (2008). A restricted model of UR discovery: Evidence from Lakhota. *Ms, MIT* .

Albright, Adam (2010). Base-driven leveling in Yiddish verb paradigms. *NLLT* **28:3**. 475–537.

Albright, Adam & Bruce Hayes (2002). Modeling English past tense intuitions with minimal generalization. In *Proceedings of the ACL-02 workshop on morphological and phonological learning*, 58–69.

Albright, Adam & Bruce Hayes (2003). Rules vs. analogy in past tenses: A computational/experimental study. *Cognition* **90:2**. 119–161.

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

Albro, Daniel Matthew (2005). *Studies in computational Optimality Theory, with special reference to the phonological system of Malagasy*. PhD dissertation, University of California, Los Angeles.

Baker, Adam (2008). Computational approaches to the study of language change. *Language and Linguistics Compass* **2:2**. 289–307.

Blust, Robert, Stephen Trussel & Alexander D Smith (2023). CLDF dataset derived from Blust's "Austronesian Comparative Dictionary" (v1.2) [data set]. Zenodo. `https://doi.org/10.5281/zenodo.7741197`.

Boersma, Paul (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Boersma, Paul, Joe Pater et al. (2016). Convergence properties of a gradual learning algorithm for harmonic grammar. In John J McCarthy & Joe Pater (eds.), *Harmonic grammar and harmonic serialism*, Equinox.

Breiss, Canaan (2021). *Lexical conservatism in phonology: theory, experiments, and computational modeling*. PhD dissertation, University of California, Los Angeles.

Brighton, Henry (2002). Compositional syntax from cultural transmission. *Artificial life* **8:1**. 25–54.

Bybee, Joan (2003). *Phonology and language use (Cambridge studies in linguistics)*. Cambridge: Cambridge University Press.

Cheung, Hintat (2000). Three to four-years old children's perception and production of Mandarin consonants. *Language and Linguistics* **1:2**. 19–38.

Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. ERIC.

Chong, Adam J (2019). Exceptionality and derived environment effects: a comparison of Korean and Turkish. *Phonology* **36:4**. 543–572.

Chong, Adam J. (2021). The effect of phonotactics on alternation learning. *Lg* **97:2**. 213–244.

Colantoni, Laura & Jeffrey Steele (2005). Liquid asymmetries in French and Spanish. *Toronto Working Papers in Linguistics* **24**.

Cutler, Anne, Andrea Weber, Roel Smits & Nicole Cooper (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America* **116:6**. 3668–3678.

Dahl, Otto Christian (1951). *Malgache et Maanjan: une comparaison linguistique*. Egede Instituttet.

Dyen, Isidore (1951). Proto-Malayo-Polynesian *Z. *Language* **27:4**. 534–540.

Dziwirek, Katarzyna (1989). Malagasy phonology and morphology. *Linguistic Notes from La Jolla* **15**. 1–30.

Fylstra, Daniel, Leon Lasdon, John Watson & Allan Waren (1998). Design and use of the Microsoft Excel Solver. *Interfaces* **28:5**. 29–55. doi:10.1287/inte.28.5.29.

Gallagher, Gillian, Maria Gouskova & Gladys Camacho Rios (2019). Phonotactic restrictions and morphology in Aymara. *Glossa* **4:1**.

Garrett, Andrew (2008). Paradigmatic uniformity and markedness. In Andrew Garrett (ed.), *Linguistic universals and language change*, 125–143. Oxford University Press.

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*, vol. 111120, .

Gouskova, Maria (2018). Morphology and phonotactics. In *Oxford research encyclopedia of linguistics*, Oxford University Press. doi:10.1093/acrefore/9780199384655.013.613.

Griffiths, Thomas L & Michael L Kalish (2007). A Bayesian view of language evolution by iterated learning. *Cognitive Science* **31**. 441–480.

Hayes, Bruce (2004). Phonological acquisition in optimality theory: the early stages. In Rene Kager, Joe Pater & Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 158–203. Cambridge University Press.

Hayes, Bruce (2011). *Introductory phonology*. John Wiley & Sons.

Hayes, Bruce & James White (2015). Saltation and the P-map. *Phonology* **32:2**. 267–302.

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

Howe, Penelope (2021). Central Malagasy. *Journal of the International Phonetic Association* **51:1**. 103–136.

Hudson, Alfred B (1967). The Barito isolects of Borneo; a classification based on comparative reconstruction and lexicostatistics. *Southeast Asia Program (Dept. of Far Eastern Studies), Data Paper no. 68* .

Hughto, Coral (2018). Investigating the consequences of iterated learning in phonological typology. *Proceedings of the Society for Computation in Linguistics* **1:21**. 182–185. doi:10.7275/R5WH2N63.

Hughto, Coral (2020). *Emergent typological effects of agent-based learning models in maximum entropy grammar*. PhD dissertation, University of Massachusetts Amherst.

Ito, Chiyuki & Naomi H. Feldman (2022). Iterated learning models of language change: A case study of sino-korean accent. *Cognitive Science* **46:4**. e13115.

Jarosz, Gaja (2006). *Rich lexicons and restrictive grammars: maximum likelihood learning in Optimality Theory*. PhD dissertation, Johns Hopkins University.

Jun, Sun-Ah (1994). The status of the lenis stop voicing rule in Korean. In YoungKey Kim-Renaud (ed.), *Theoretical issues in Korean linguistics*, 101–114. Stanford:CSLI.

Kabak, Barış & Irene Vogel (2001). The phonological word and stress assignment in Turkish. *Phonology* **18:3**. 315–360.

Kang, Yoonjung (2006). Neutralizations and variations in Korean verbal paradigms. *Harvard Studies in Korean Linguistics* **11**. 183–196.

Kaplan, Abby (2010). *Phonology shaped by phonetics: The case of intervocalic lenition*. PhD dissertation, University of California, Santa Cruz.

Katz, Jonah (2016). Lenition, perception and neutralisation. *Phonology* **33:1**. 43–85.

Keenan, Edward L & Cecile Manorohanta (2001). A quantitative study of voice in Malagasy. *Oceanic linguistics* **40:1**. 67–84.

Keenan, Edward L & Maria Polinsky (2017). Malagasy (austronesian). *The handbook of morphology* 563–623.

Kenstowicz, Michael (1996). Base-identity and uniform exponence: alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, 363–394. Salford: University of Salford.

Kiparsky, Paul (1965). *Phonological change*. PhD dissertation, MIT.

Kiparsky, Paul (1968). Linguistic universals and linguistic change. In Emmon W. Bach & Robert T. Harms (eds.), *Universals in linguistic theory*, 170–202. New York: Hold, Rinehart & Winston.

Kiparsky, Paul (1978). Analogical change as a problem for linguistic theory. *Studies in the Linguistic Sciences Urbana, Ill* **8:2**. 77–96.

Kirby, Simon (2001). Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation* **5:2**. 102–110.

Kirchner, Robert M. (1998). *An effort-based approach to consonant lenition*. PhD dissertation, University of California, Los Angeles.

Kuryłowicz, Jerzy (1945). La nature des procès dits «analogiques». *Acta linguistica* **5:1**. 15–37.

de La Beaujardière, Jean-Marie (2004). Malagasy dictionary and encyclopedia of Madagascar.

Lewis, Paul M., Gary F. Simons & Charles D. Fennig (2014). *Ethnologue: Languages of Asia*. SIL international.

Mahdi, Waruno (1988). *Morphophonologische Besonderheiten und historische phonologie des Malagasy*, vol. 20. D. Reimer.

Martin, Andrew (2011). Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language* **87:4**. 751–770. doi:10.1353/lan.2011.0096.

Mayer, Connor, Adeline Tan & Kie Zuraw (2022). *maxent.ot: Perform maxent optimality theory analyses in r (version 0.1.0)*. `https://doi.org/10.5281/zenodo.7246367`. Computer software.

McCarthy, John J. & Alan S. Prince (1995). Faithfulness and Reduplicative Identity. In Laura Walsh Dickey Jill N. Beckman & Suzanne Urbanczyk (eds.), *Papers in optimality theory*, 249–384. Amherst: GLSA.

Moreton, Elliott & Joe Pater (2012). Structure and substance in artificial-phonology learning, part I: Structure. *Language and linguistics compass* **6:11**. 686–701.

Moreton, Elliott & Joe Pater (2012). Structure and substance in artificial-phonology learning, part II: Substance. *Language and linguistics compass* **6:11**. 702–718.

Niyogi, Partha (2006). *The computational nature of language learning and evolution*. MIT press.

Nosofsky, Robert M (2011). The generalized context model: An exemplar model of classification. In Emmanuel M. Pothos & Andy J. Wills (eds.), *Formal approaches in categorization*, 18–39. Cambridge University Press.

O'Hara, Charlie (2020). Frequency matching behavior in on-line MaxEnt learners. In *Proceedings of the society for computation in linguistics 2020*, 461–463. New York, New York: Association for Computational Linguistics. `https://aclanthology.org/2020.scil-1.54`.

O'Hara, Charlie (2022). MaxEnt learners are biased against giving probability to harmonically bounded candidates. In *Proceedings of the society for computation in linguistics 2022*, 229–234. Association for Computational Linguistics.

O'Neill, Timothy (2015). *The phonology of Betsimisaraka Malagasy*. PhD dissertation, University of Delaware.

Paster, Mary (2005). Subcategorization vs. output optimization in syllable-counting allomorphy. In *Proceedings of the 24th west coast conference on formal linguistics*, vol. 24, 326. Cascadilla Proceedings.

Paster, Mary (2009). Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering. *Word structure* **2:1**. 18–37.

Pater, Joe (2007). The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In *University of Massachusetts occasional papers 32: Papers in optimality theory III*, 259–296. Amherst: GLSA.

Pater, Joe & Anne-Michelle Tessier (2005). Phonotactics and alternations: Testing the connection with artificial language learning. *University of Massachusetts Occasional Papers in Linguisitcs* **31**. 1–16.

Peperkamp, Sharon, Rozenn Le Calvez, Jean-Pierre Nadal & Emmanuel Dupoux (2006). The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* **101:3**. B31–B41.

Pierrehumbert, Janet (2002). Word-specific phonetics. In C Gussenhoven & N Warner (eds.), *Laboratory phonology VII*, 101—140. Berlin: Mouton de Gruyter.

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Rasoloson, Janie & Carl Rubino (2005). Malagasy. In Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 456–488. Routledge Abingdon.

Rosenblatt, Frank (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review* **65:6**. 386.

Sapir, Edward (1915). Notes on Judeo-German phonology. *The Jewish quarterly review* **6:2**. 231–266.

Selkirk, Elisabeth (1980). Prosodic domains in phonology: Sanskrit revisited. In Mark Arnoff & Mary-Louise Kean (eds.), *Juncture*, 107–129. Anma Libri.

Singleton, Jenny L & Elissa L Newport (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive psychology* **49:4**. 370–407.

Smolensky, Paul (1986). Information processing in dynamical systems: Foundations of harmony theory. Tech. rep. Colorado Univ at Boulder Dept of Computer Science.

Staubs, Robert D (2014). *Computational modeling of learning biases in stress typology*. PhD dissertation, University of Massachusetts Amherst.

Steriade, Donca (2009 [2001]). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. In Kristin Hanson & Sharon Inkelas (eds.), *The nature of the word: Studies in honor of Paul Kiparsky*, 151–180. Cambridge, MA: MIT Press.

Suzuki, Keiichiro (1998). *A typological investigation of dissimilation*. PhD dissertation, The University of Arizona.

Tabor, Whitney (1994). *Syntactic innovation: A connectionist model*. PhD dissertation, Stanford University.

Tesar, Bruce & Alan Prince (2003). Using phonotactics to learn phonological alternations. *CLS* **39:2**. 241–269.

Tsao, Feng-Ming, Ching-Yun Lee, Yi-Hsin Hsieh & Chin-Yeh Chiu (2009). Assessing stop and lexical tone perception in preschool children and relationship with word development. *Journal of the Speech-Language-Hearing Association of Taiwan* **24**. 39–57. doi:doi:10.6143/JSLHAT.2009.12.03.

Vennemann, Theo (1972). Rule inversion. *Lingua* **29:3-4**. 209–42.

Wang, Marilyn D. & Robert C. Bilger (1973). Consonant confusions in noise: A study of perceptual features. *JASA* **54:5**. 1248–1266.

Warner, Natasha, James M McQueen & Anne Cutler (2014). Tracking perception of the sounds of English. *JASA* **135:5**. 2995–3006.

Weinreich, Uriel, William Labov & Marvin Herzog (1968). *Empirical foundations for a theory of language change*. University of Texas Press.

Wheeler, Max W. (2005). *The phonology of Catalan*. Oxford University Press.

White, James (2013). *Bias in phonological learning: Evidence from saltation*. PhD dissertation, University of California, Los Angeles.

White, James (2017). Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Lg* **93:1**. 1–36.

Wilson, Colin (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive science* **30:5**. 945–982.

Yang, Hsiu-fang (1976). The phonological structure of the Paran dialect of Sediq. *Bulletin of the Institute of History and Philology Academia Sinica* **47:4**. 611–706.

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

Zuraw, Kie (2003). Probability in language change. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 139–176. MIT Press.

Zuraw, Kie (2010). A model of lexical variation and the grammar with application to Tagalog nasal substitution. *NLLT* **28:2**. 417–472.

Zuraw, Kie (2013). *map constraints. Ms, University of California, Los Angeles.

Zymet, Jesse (2020). Malagasy OCP targets a single affix: Implications for morphosyntactic generalization in learning. *Linguistic Inquiry* **51:3**. 624–634.

A Irregular alternation patterns in Malagasy weak stems

pattern		**stem**	**passive (stem + ana)**	
ʈs̺a ~	s	ˈburi**ʈs̺**a	buˈri**s**-ana	‘saw off’
ʈs̺a ~	s	ˈrumpu**ʈs̺**a	ruˈmpu**s**-ana	‘to snatch’
n ~	s	ˈreni**n**a	hareˈne**s**-ina	’to be deaf’
n	f	ˈbia**n**a	biˈna**f**-ina	’to open’

Table 17: Irregular alternation patterns in Malagasy weak stems