Lexical category and the distribution of word-final vowels in Hiaki*

Jianrong Yu, Adam King, Maria Florez Leyva, Santos Leyva, and Heidi Harley
University of Arizona

Abstract: We report on an automated extraction and analysis of lexical items from a digitized dictionary of Hiaki (Yaqui), which enabled the discovery of patterns in the distribution of word-final vowels according to lexical category. Headwords were extracted from a Microsoft Word dictionary file with a Perl script, then binned and counted according to lexical category and word-final vowel. There were statistically significant interactions between lexical category and word-final vowel, and a near-categorical prohibition on word-final [i] in verbs and word final [e] in adjectives. The exceptions were few enough in number that we undertook detailed item-by-item investigations of each, discovering that in the case of putative 'verbs' ending in [i], the dictionary-makers had miscategorized, misspelled, or otherwise mistakenly presented the forms. We present a theory that active Voice has an exponent [-e] in Hiaki, which coalesces with stem-final [i] verbs to yield surface [e] word-finally in free verb forms.

Keywords: Voice, dictionary, verbs, adjectives, derivation, categorization, Yaqui

1 Introduction

In this paper we investigate an initial impression that the distribution of word-final vowels in Hiaki verbs and adjectives is more skewed than normal morphological processes would predict. The specific goal is to determine if the distribution of word-final vowels in Hiaki is truly asymmetric across lexical categories, and, if so, to provide an account for the asymmetry. We find that this is indeed the case, especially for the categories of verbs and adjectives. Specifically there is a striking asymmetry in the distribution of word-final [e] and [i] between verbs and adjectives: [e] is much more common for verbs than [i], with the reverse being true for adjectives. Further detailed lexical work on apparent exceptions reveal that the prohibition on word-final [i] in verbs is apparently categorical. We propose a non-lexicalist account for this distribution couched in terms of Distributed Morphology. Through this study we also hope to illustrate that simple digital tools, such as the digitization of dictionaries combined with simple computational methodologies, can help uncover and profile linguistic phenomena of potential interest in fast and efficient ways. These discoveries in turn can help feed documentary goals and raise new theoretical questions and hypotheses to be tested.

The paper is organized as follows. Section 2 provides a brief background about Hiaki and describes linguistic patterns relevant to our investigations. Section 3 describes the methodologies used and discusses the results. Section 4 discusses the results of examining the exceptions from the results in Section 3 with our native speaker consultants, and shows that many of the exceptions are either the result of misanalysis or reveal potentially deeper questions about Hiaki morphology. Section 5 provides a theoretical analysis of our results and Section 6 concludes.

* This work was supported by NSF grant BCS-1528295 to Harley.
Contact info: jianrongyu@email.arizona.edu, adamking@email.arizona.edu, hharley@email.arizona.edu

In Proceedings of the Workshop on the Structure and Constituency of Languages of the Americas 21, University of British Columbia Working Papers in Linguistics 46,
Megan Keough, Natalie Weber, Andrei Anghelescu, Sihwei Chen, Erin Guntly, Khia Johnson, Daniel Reisinger, and Oksana Tkachman (eds.), 2018.

2 Background, motivation, and hypothesis

Hiaki is an Uto-Aztecan language spoken in Arizona and Northern Sonora.[1] Word order is rigidly SOV, with rich inflectional and derivational morphological processes which will be important for our analysis. The vowel inventory is a typical five-vowel system comprising [a], [e], [i], [o], and [u]. In traditional descriptions of Hiaki, vowel length is often described as phonemic. We set aside vowel length here since long vowels seldom occur word-finally.

The numerous inflectional and derivational suffixes of Hiaki often alter and affect the distribution of word-final vowels. A sampling of some relevant to our purposes is provided below.

(1) a. *V* + participial suffix *-ri* → adjectives and nouns
b. *N* + 'do/make' verbalizer *-te* → verbs
c. *V* + agentive nominalizer *-reo* → nouns
d. *Adj* + adverbializer *-si* → adverb

After years of studying the rich morphological system of Hiaki, it was the the fifth author's impression that there might be a significant asymmetry in the distribution of [e] and [i] word-finally between the lexical categories of verbs and adjectives. We thus set out to determine if vowels, in particular [e] and [i], exhibit an asymmetric distribution in word-final positions across different lexical categories and, if so, why this might be the case.

3 The distribution of word-final vowels: method[2]

The primary data for this study was extracted from a digitized version of the Yoeme-English side of the Yoeme-English/English-Yoeme Standard Dictionary by Molina, Valanzuela, & Shaul (1999). Each lexical entry in the dictionary is arranged in the following format, with each entry mandatorily containing the headword, its lexical category, and at least one definition.

(2) headword *lexcat.* definition1. definition2. (grammatical notes); example1.
translation1. example2. translation2. example3. translation3....

Although headwords are both multimorphemic and monomorphemic, and although suffixation could affect the distribution of word-final vowels, we did not attempt to distinguish multimorphemic from monomorphemic words on our first pass at the question; it subsequently turned out that the results were informative even without implementing that refinement. Within the dictionary, the headword in the lexical entry is almost always a *free form*, which in Hiaki is the form that surfaces without any suffixes and also the form that serves as the stem for inflection. Most Hiaki words also have a bound form, which, in contrast to the free form, is used as the stem for derivational and compounding processes.[3]

To sort the relevant headwords into their lexical categories, a Perl script was written to extract regular expressions of the form ¶ * {*n., iv., tv., adj., adv.*}. Thus any string corresponding to a headword was extracted together with its lexical category, since each lexical entry in the dictionary

[1] Yoeme, Yaqui, Hiaki, and Jiaki are all names for the same language; Hiaki is the name preferred by our native speaker consultants, and represents the correct spelling in the Arizona tribe's relatively phonemic orthographic system.
[2] The data and R scripts used can be found at: https://github.com/King-of-Ling/Hiaki
[3] The distinction between free and bound forms will be crucial to the analysis we propose below.

was arranged in the format as in (2). After the headwords were extracted, they were sorted by lexical categories in a simple spreadsheet. There were five lexical categories: nouns, adjectives, adverbs, intransitive verbs, and transitive verbs. Function words were not collected.

The resulting data allowed for an analysis of the distribution of all vowels in Hiaki by lexical category as well as of the distribution of syntactic categories for each individual vowel. We calculated the ratio of word-final to non-word-final vowels (final:non-final) across all lexical categories, which we take to be the baseline ratio of comparison. We then calculated the final:non-final ratio for each individual vowel in Hiaki by lexical category for comparison with the baseline ratio. χ^2 tests (X-squared = 360.48, df = 4, p-value < 2.2e-16) were run to ensure that the distribution of counts and ratios for final vowels was statistically significant.

3.1 The distribution of word-final vowels by category

In total 3230 vowel-final headwords were extracted from the dictionary via the aforementioned methods. The breakdown of the number of headwords per lexical category is given below.

Table 1: Total number of entries by lexical category

Category	Word Count
Noun	1421
Adjective/Adverb	682
Verb	1127
Total	3230

The figures below, generated using standard packages in the statistical programming language R, illustrate the results. The x-axis shows the count of total vowel tokens while the y-axis indicates the individual vowel phonemes. Non-final:final proportions are indicated by the shaded and non-shaded parts of each bar in the graphs, while percentages indicate the numerical ratio between final and nonfinal positions.

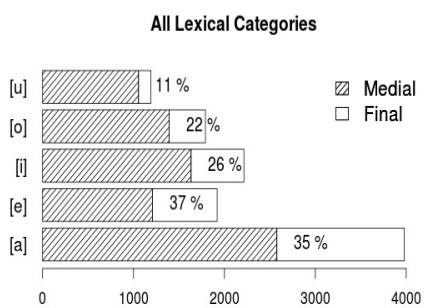


Figure 1: Distribution of vowel tokens by position in word X-squared = 360.48, df = 4, p-value < 2.2e-16, Cramer's V: .1802

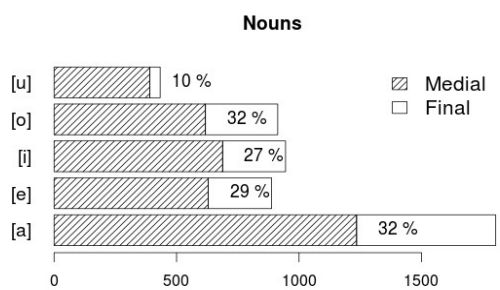


Figure 2: Distribution of final vowels with nouns X-squared = 90.529, df = 4, p-value < 2.2e-16 Cramer's V: .1348

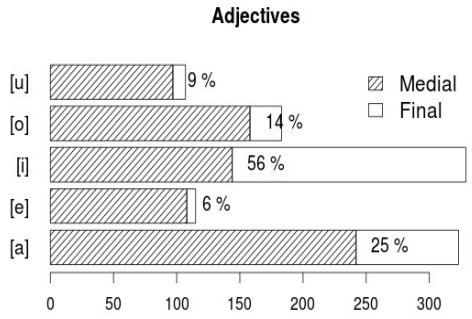


Figure 3: Distribution of final vowels with adjectives X-squared = 190.66, df = 4, p-value < 2.2e-16, Cramer's V: .4247

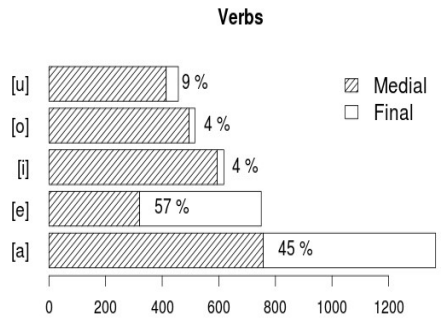


Figure 4: Distribution of final vowels with verbs X-squared = 856.58, df = 4, p-value < 2.2e-16, Cramer's V: .4806

Figure 1 illustrates the distribution of vowel tokens in medial and final positions across all lexical categories. Figures 2, 3, and 4 illustrate the same distribution of vowels in medial and final positions broken down by the lexical categories of nouns, adjectives, and verbs, respectively. The verbal category was further split up into intransitive and transitive verbs as indicated in the dictionary since the categories of intransitive and transitive are morphologically significant. This is shown in Figures 6 and 7.

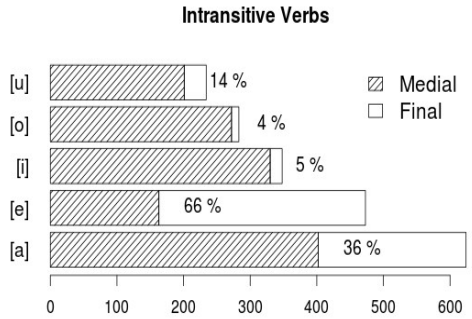


Figure 5: Distribution of final vowels with intransitive verbs X-squared = 513.18, df = 4, p-value < 2.2e-16, Cramer's V: .5114

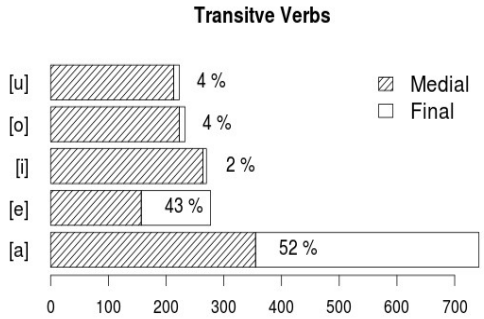


Figure 6: Distribution of final vowels in transitive verbs X-squared = 433.85, df = 4, p-value < 2.2e-16, Cramer's V: .4986

3.2 The association of categories with word-final vowels

Having determined the ratios of non-final:final vowels for each category, we specifically investigated the distribution of final vowel tokens across the syntactic categories. For each final vowel, we charted the number of nouns, verbs (split into the two categories), adjectives, and adverbs. Each of the graphs below shows the raw counts and percentages of total vowel tokens that occurred in final position.

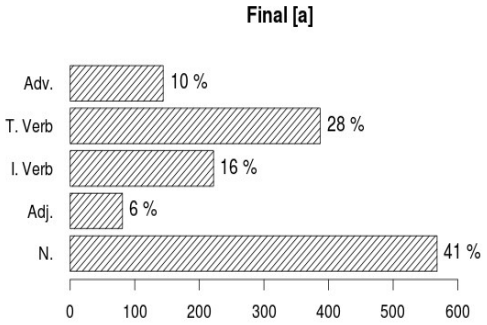


Figure 7: Distribution of final [a] across categories
X-squared = 555.82, df = 4, p-value < 2.2e-16,
Cramer's V: .3148

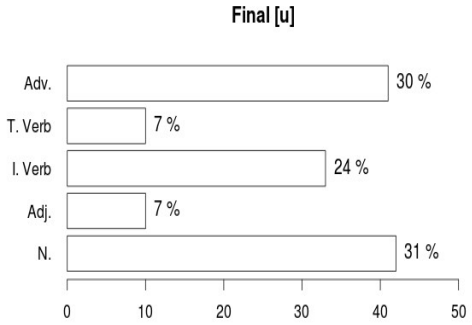


Figure 8: Distribution of final [u] across categories
X-squared = 38.044, df = 4, p-value = 1.097e-07,
Cramer's V: .2644

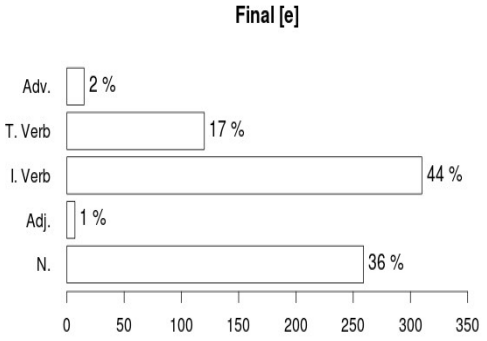


Figure 9: Distribution of final [e] across categories
X-squared = 539.74, df = 4, p-value < 2.2e-16,
Cramer's V: .4356

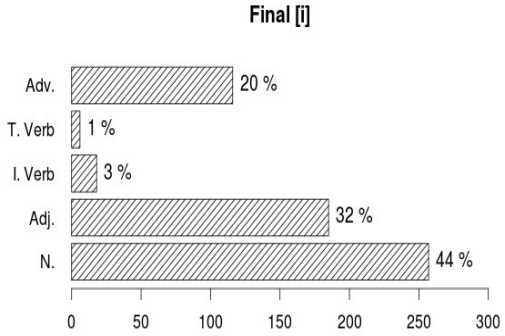


Figure 10: Distribution of final [i] across categories
X-squared = 398.15, df = 4, p-value < 2.2e-16,
Cramer's V: .4135

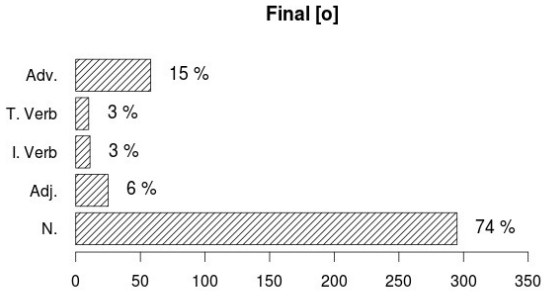


Figure 11: Distribution of final [o] across categories
X-squared = 744.3, df = 4, p-value < 2.2e-16,
Cramer's V: .6828

From Figures 7–11, it is clear that nouns represented a very high percentage of the final vowel tokens for all vowels, an expected result because it is the most frequent lexical category within the dictionary in terms of raw frequency. Looking at each individual vowel, [a] was most frequently word-final in nouns and transitive verbs; [e] was most common in intransitive verbs and nouns and much less so for adjectives; [i] had the highest ratio of final vowels in nouns, adjectives, and adverbs; [u] was the least likely vowel to be found in final position, and most likely to terminate adverbs, intransitives, and nouns; [o] was overwhelmingly dominated by nouns. Importantly, the number of verbs that ended in [i] was very small, as was the number of adjectives that ended in [e].

The results above seemed to suggest a clear discrepancy in the distribution of [e] and [i] across the lexical categories of verbs, adjectives, and adverbs. We thus compared the distribution of [e] and [i] across these lexical categories, collapsing adjectives and adverbs together as they seemed to exhibit parallel behaviors. As seen in table 2, the inverse pattern of [i] in adjectives and adverbs and [e] in verbs is so striking that it suggests that the distinction may perhaps be categorical.

Table 2 Number of adjectives, adverbs, and verbs with final -e and -i

	/i/	/e/
Adjectives or adverbs	301	22
Verbs	24	330

This stimulated us to pursue documentary work with the 46 exceptions to the overwhelming trends for these categories, which we report in the next section.

4 Testing the categories of apparent exceptions

4.1 Tests for lexical categories

The results revealed that there is a significant asymmetry in the distribution of vowels in word-final position across lexical categories, specifically for [e] and [i] across adjectives, adverbs, and verbs. Even so, there were 22 adjectives and adverbs ending in [e] (22/682 = **3%** of all adjectives/adverbs) and 24 verbs ending in [i] (24/1129 = **2%** of all verbs). We thus set out to more precisely determine and investigate the categorial status of these exceptions.

Each exception was tested to determine its categorial status, relying on the native-speaker intuitions of the third and fourth authors. This was done using a variety of diagnostic tests discussed by Jelinek et al. (1998) and Harley, Haugen, & Tubino Blanco (*in prep*). We describe these tests for each lexical category as below.

(3) **Verbs**

- Can be suffixed with tense, aspect, and mood (TAM) morphology directly.
- To be used attributively, must be derived into a participle or relativized; in either case, does not appear pre-nominally.

(4) **Adjectives**

- Can appear pre-nominally in an attributive use.
- Can appear with nominal inflections, such as the plural suffix *-(i)m* and the accusative suffix *-ta*, when used alone in argument position or post-nominally.
- When used predicatively, can be inflected with certain tense, aspect, and mood morphology but needs to be verbalized by suffixation of *-tu* which indicates *being ADJ* or *becoming ADJ* (Jelinek 1998:178).

(5) **Nouns**

- Take nominal inflectional morphology like plural *-(i)m* and accusative *-ta* in argument positions.
- With tense, aspect, and mood morphology directly attached, return possessive/in-use meanings.
- When used predicatively in non-present tenses, require the suffixation of *-tu*, like adjectives.

(6) **Adverbs**

- None of the above, i.e.
- Does not occur as a predicate or inflect with TAM.
- Does not occur as an argument or inflect with case/number.
- Occurs in the left periphery or middle field.

4.2 Results for verbs ending with [i]

It was found the exceptional 'verbs' ending in [i] were for the most part miscategorized or otherwise erroneously recorded in the dictionary in a variety of ways.[4] For example, three forms were simply not attested in the vocabulary of the third and fourth authors:

(7) **Example of an unattested *i*-final verb:**
havoi, given as: tv. 'miss out on food': not attested

Eight forms were found to be mistakenly categorized as verbs; according to the diagnostics for adjectives, these forms behaved like prototypical adjectives rather than verbs. One example of such a form, *nasonti*, is given in (8) below. This word forms part of a typical stative-adjective/inchoative/causative alternation determined by the word-final vowel: *nasonti/nasonte/nasonta*, 'ruined/get.ruined/ruin;' there are dozens of parallel triples in the language and in most of them the *-i* final alternant is correctly categorized in the dictionary as an adjective.

(8) **Example of a miscategorized *i*-final 'verb'**
nasonti, given as: iv. 'ruined, blotched': actually behaves as an adjective

Five other forms showed word-final vowel alternations *-i* ~ *-ia* and *-i* ~ *-e* in reduplication contexts, as shown in the representative case below.

(9) e.g. *bwisi*, given as: tv. 'hold, grasp it'; must be *bwibwis**e*** when reduplicated

The form *bwisi* behaves in certain respects like a verb, e.g. inflecting for the past imperfective *-n* directly. Under contexts of verbal reduplication though, *bwisi* surfaces as *bwibwis**e*,** showing the word-final *-e* in compliance with the pattern discovered in Section 3. The form is also related to the forms *hi'ibwisia/hi'ibwise*, formed from the indefinite detransitivizing prefix *hi'i-* plus *bwisia/bwise* meaning 'lean (on something)'. However **bwisia* is not good as a word on its own.

[4] A full list of the results for both verbs and adjectives is given in Appendix A.

We hypothesize that *bwisi* may be a shortened form of an underlying *bwisia*, the result of a novel morphophonological process to be investigated.

Four other forms remain uninvestigated. The form *valumai*, for example, was indicated by the dictionary to be a verb occurring in deer song;[5] the third and fourth authors note that it is not used in everyday Hiaki and also that it is not appropriate to discuss deer song lyrics, which should be respected.

(10) *valumai*, given as: tv. 'wash' (song language)

Three other yet-to-be investigated forms are all derived by the suffixation of *-machi*, which is itself derived from *maachi(a)*.

(11) *maachi(a)*: usually treated as a verb meaning 'appear, be light'

In its suffixal form, *-machi*, again demonstrating the *-i* ~ *-a* alternation, this morpheme has a variety of readings, including 'seem', 'appear', and evidential modal 'should'. These are highly complex forms and require extensive documentation and description, which we leave for future elicitation work. We note, however, that the form *machitukan*, using the verbalizing suffix *-tu* on the stem before the TAM suffix *-kan*, is attested. This failure to directly inflect for TAM suggests that, like other forms ending in [i], *maachi* may not be verbal.

Taking into consideration the above miscategorizations, the number of genuinely exceptional verbs was reduced to only two, both based on the same stem, *temai* and *nattemai,* and both broadly translated as 'ask'. These forms behaved like prototypical verbs according to all categorial diagnostics. However, although it appears they end in [i], they do not; rather, they end in a diphthong [aj], whose glide [j] is represented as 'i' in Hiaki orthography.

(12) e.g. *nattemai*, given as: tv. 'ask about, ask for'

Following this detailed documentary work, we conclude, then, that it is very likely to be the case that there are actually no verbs in Hiaki ending in [i].

4.3 Results for 22 adjectival and adverbial forms ending in [e]

The results of applying diagnostic tests to the 22 adjectival and adverbial forms ending in [e] showed similar results as the verb exceptions. Firstly, one form was unattested.

(13) *amene*, given as: adv. '(so) much, this much'

This might be a misspelling of *amuine*, the future form of the verb *amue* with the future suffix *-ne* meaning 'suffice'. In addition, two other forms turned out to be misspelled in the dictionary, and should have ended with 'i' rather than 'e':

(14) e.g. *kutvene*, given as: adj. 'dark (color)' should in fact be spelled *kutvene**i***

Two forms turned out to be miscategorized and behaved as verbs according to the verb diagnostics.

[5] Deer song is a type of ceremonial music and dance of the Hiaki, traditionally performed during important occasions such as the celebration of Easter.

(15) e.g. *kutwatwatte*, given as: adv. ‘in the dark’ is in fact a verb, ‘be.dark’, iv.

Four other forms were misanalyzed. Specifically, some adverbs were postpositional phrases, derived from nominal stems combined with the postposition *-e*:

(16) e.g. *amae*, given as: adv. ‘over there’ is in fact, *ama* ‘over there’ + =*e* ‘from’, used in

(17) contexts of movement away

Three other forms showed puzzling misspelling or alternations. A representative case involves the adverb *vahiwe*, given as below in the dictionary.

(18) e.g. *vahiwe*, given as: adv. ‘in three ways’

The third and fourth authors suggest this should actually be spelled *vahiweeye*, which looks to be a compound verb, *vahi-weeye* ‘three-wander’, but behaves like a nominal. They further pointed out that it can also occur as *vahiweyame*, which looks like a subject relative of ‘*vahi-weeya*’, also ‘three-wander’. Both *vahiweeye* and *vahiweyame* behave syntactically like nouns meaning ‘three things’/‘three topics’, and are not adverbial at all. Interestingly, when in object position and hence suffixed with accusative *-ta*, the subject relative form *vahiweyame* — and not *vahiweeye* — must be used. This suggests *vahiweeye* may be a clipping of *vahiweyame*. Despite the dictionary’s adverbial categorization, *vahiwe* may be best characterized as a singular noun meaning something like ‘a triplet (of topics)’.

Taking into account the above results, the number of genuine exceptions — adverbs and adjectives ending in [e] — is then reduced to eight. They all behave like prototypical adjectives according to the diagnostics for adjectives.

(19) e.g. *teeve*, given as: adj. ‘tall’

Although the pattern for adjectives is still striking, it did not prove to be as categorical as the prohibition on [i]-final verbs.

5 Hypothesis

5.1 Active voice has exponent [e]

The key findings thus far are that there is an asymmetry in the distribution of the vowels [e] and [i] in word-final positions between verbs and adjectives/adverbs. To begin to understand why this might be the case, it is necessary to first understand the difference between free and bound verbal forms in Hiaki.

Hiaki verb forms alternate between a free form and a bound form. The free form serves as the stem for the attachment of inflectional affixes, while the bound form serves as the stem for derivational affixes.

(20) **Free form inflects**

a. <u>poona</u>-**k**
pound-PRF

b. vicha-**kai**
see-PPL

(21) **Bound form derives**

a. pon-**vae**
pound-PROSP

b. vit-**ria**-k
see-APPL-PRF

The alternation between free and bound stems takes many morphophonolgical patterns, but there are three major classes (Harley and Tubino-Blanco 2013): truncation, insertion of an echo-vowel or remain invariant, as the examples below show.

(22) a. Truncation: *poona~pon-*
b. Echo-vowel: *kivacha~kivacha'a-*
c. Invariant: *kotta~kotta-*

Apart from these three main classes of free~bound alternations, there exists another group of irregular correspondences that involve various vowel changes. A sample list of such irregular correspondences is illustrated below; importantly, many of these irregular forms involve vowel changes at the right edge, highlighting the importance of the right edge of a verb as a major site for morphophonological alternations.

(23) **Irregular stem form correspondences**

yaha → ya(h)i-	'arrive.pl'	*yepsa → yevih-*	'arrive.SG'
he'e → hi'i-	'drink'	*suulu → suluu-*	'slide, slip'
kiimu → kimoe-	'bring.SG'	*sevea → seve-*	'get cold'
hi'ivoa → hi'ivoo-	'cook.intr'	*hia → hiu-*	'vocalize'
eo'otea → eo'ote'e-	'be.nauseated'	*kepe → kup-*	'close one's eyes'
me'a → me'e-	'kill.PL'	*vicha → vit-*	'see'
nu'e → nu'u-	'get'	*ke'e → ki'i-*	'bite'
yu'a → yu'u-	'push'		

As can be seen from the list in (22), there is no prohibition on a verb's *bound* stem ending in [i]; it's only free stems which never end in [i]. In particular, the right edge serves as the site for the morphophonological change of a major class of transitivity alternating verbs that exhibit a pattern with terminal [i] in the intransitive bound stem and terminal [e] in the intransitive free stem (Jelinek 1998). Importantly, as seen in the intransitive verb stems below, every verb whose bound form ends in [i] has a free form counterpart ending in [e], and, as noted above, there are no verbal free forms that end in [i].

(24) **Transitivity alternations**

Transitive	Tr. Stem	Intransitive	Intr. Stem
chep-ta: 'step on x'	*chepta-*	*chep-te*: 'jump'	*chepti-*
om-ta: 'scold x'	*omta-*	*om-te*: 'be angry'	*omti-*
noi-ta: 'take/bring x'	*noita-*	*noi-te*: 'go/come'	*noiti-*
vehuk-ta: 'duck under x'	*vehukta-*	*vehuk-te*: 'bow'	*vehukti-*
vee-ta: 'burn x'	*veta-*	*vee-te:* 'burn'	*veti-*
kot-ta: 'break x'	*kotta-*	*kot-te:* 'break'	*kotti-*
yook-a: 'color x'	*yoka-*	*yook-e:* 'change color'	*yoki-*
moh-ta: 'grind x finely'	*mohta-*	*moh-te:* 'disintegrate'	*mohti-*

Taking the patterns discussed above into consideration, we propose that in verbal free forms, word-final [e] is a morphological exponent of active Voice and that Voice is a mandatory category in the verbal extended projection. Hiaki verb forms are spelled out cyclically, and inflectional and derivational suffixes are spelled out in different cycles (Harley and Tubino Blanco 2013). Inflectional suffixes are *word-attaching* and are attached outside the first cycle while derivational suffixes are *stem-attaching* and attach within the first cycle. Voice morphology (*PASS/IRR*) occurs at a cyclic boundary between where the first cycle is spelled out and where the second cycle begins, as shown in the Hiaki verbal template.

(25) ...V - [(Derivation*)] - **[(PASS/IRR)]** |- [(Tense/Asp)] - [(Relativizer)]
√ - v° - Voice°- T/Asp° - C°
BOUND STEMS REQUIRED | FREE STEMS REQUIRED

Harley and Tubino Blanco (2013) note further that the phase boundary indicated by voice morphology is significant for the bound and free form distinction. As shown in the alternating pair below, the bound form of the verb without word-final [e] is required as a stem for the attachment of passive morphology.

(26) a. Aapo aman kom chept**e** (Active)
3SG there down step$_{FREE}$
'He stepped down there.'

b. Aman kom chept**i**-wa (Passive)
there down step$_{BOUND}$-IMP.PASS
'There was stepping down.' (lit. It was stepped down.)

Cyclic spell-out thus applies at VoiceP, at which point the complex head in Voice° must be interpretable at PF. The verb stem to which T/Asp and C suffixes attach must be free. In free forms then, cyclic spell-out ensures that active Voice is spelled out as [e], and vowel coalescence overwrites verb-stem-final [i]. Alternatively, stem-final [i]s are epenthetic; not inserted if Voice supplies an [e].

Support for this view of word-final [e] comes from the possession/use construction in Hiaki (Jelinek 2003).

(27) a. Inepo chu'u-ne.
I dog-IRR
'I will have a dog.'

b. Maria siali-k karo-k
Maria green-ACC car-PRF
'Maria has a green car.'

The possession construction as shown in (25) involves taking a noun N and inflecting it directly with verbal TAM morphology, yielding a 'have N' interpretation. Importantly, nominals that end in [i] must surface with stem-final [e] when inflected in the possessive constructions.

(28) a. kava'i
horse

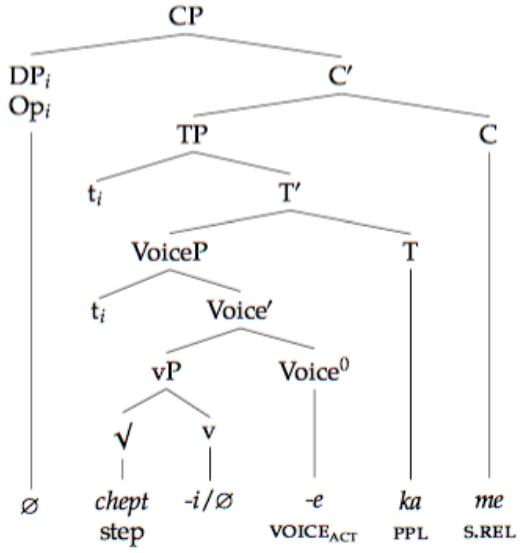
b. kava'e-k
horse-PRF
'have a horse'

c. ka-kava'e
RED-horse
'habitually have horses'

Haugen (2004, 2008) hypothesized that the *i~e* alternation at the end of nouns in the possessive construction reflected the morphophonological exponence of an otherwise null possessive predicate. However, if the possessive construction involves embedding a N under an active Voice head whose exponent is [e], the [i] to [e] change here could instead be subsumed as part of the broader phenomenon of the cyclic spelling out of active Voice whose exponent is [e].

Assuming that verb-final *-e* is an exponent of Voice when it occurs, the morphosyntactic structure of the subject relative *cheptekame* meaning 'one who stepped' is illustrated below.

(29) *cheptekame*, 'one who stepped'



5.2 Possible reasons for association of final [i] with adjectives

Section 4 showed that there were genuine exceptions to the claim that adjectives and adverbs do not generally end with a final [e] (eight such exceptions). This casts doubt on the search for a reason to completely exclude [e] for adjectives and adverbs. Nevertheless there is a clear morphological reason why many adjectives end in [i]. In Hiaki, deverbal adjectives can be derived by the suffix /-ri/, as the examples below illustrate.

(30) a. Inepo uka ki'i-**ri**-k atteak
1.SG DET.ACC bite-**RI**-ACC own
'I have that bitten (one).'

b. Uka voa hutta-**ri**-k ne-u bwise
DET.ACC feather pluck-**RI**-ACC 1.SG-to hand
'Hand me the plucked one.'

The adjectival status of these /-ri/ forms is confirmed because they can be inflected with the adjectival accusative suffix /-k/, characteristic of the category of adjectives in Hiaki (Harley, Haugen, and Tubino Blanco in prep). Importantly, intervocalic [r] drop frequently applies to /-ri/ derived forms, yielding the appearance of an [i]-final adjectival form with no obvious signature of deverbal derivation.

6 Conclusion

In this paper we demonstrated how the use of simple digital tools allowed us to quickly and easily investigate the robustness of a morphophonological pattern that would have been challenging and potentially unfruitful to look at by hand. This allowed us to target 46 lexical items as likely needing further descriptive work and we improved on existing documentation about them considerably through traditional grammaticality judgement testing. The morphophonological patterns discovered through these methods in turn have theoretical implications, leading us to suggest that active Voice might have a previously overlooked exponent in Hiaki, also shedding light on the complex morphophonological and morphosyntactic processes that relate bound stem forms to free forms in Hiaki.

References

Harley, H., J. Haugen, & M. Tubino Blanco. (In Prep.). *Lexical Categories and Derivation in Yaqui (Hiaki)*.

Harley, H. & M. Tubino Blanco. (2013). Cycles, Vocabulary Items and Stem Forms in Hiaki. In: O. Matushansky & A. Marantz (eds.), *Distributed Morphology Today: Morphemes for Morris Halle*, pp. 117–134. Cambridge: MIT Press.

Haugen, J. D. (2004). *Issues in Comparative Uto-Aztecan Morphosyntax*. Doctoral dissertation, University of Arizona.

Haugen, J. D. (2008). *Morphology at the Interfaces: Reduplication and Noun Incorporation in Uto-Aztecan* (Vol. 117). Amsterdam: John Benjamins Publishing.

Jelinek, E. (1998). Voice and Transitivity as Functional Projections in Yaqui. In: M. Butt and W. Greuder (eds.), *The Projection of Arguments: Lexical and Compositional Factors*, pp. 195–224. Stanford: CSLI Publications.

Jelinek, E. (2003). Quantification in Yaqui Possessive Sentences. *MIT Working Papers in Endangered and Less Familiar Languages* 5: 201–214.

Jelinek, E., M. Amarillas, F. Delgado, R. Estrella, R. Garcia, M. Molina, & S. Morales. (1998). *Yaqui Language Workbook.* Ms. University of Arizona.

Molina, F., H. Valanzuela, & D. Shaul. (1999). *Yoeme-English English-Yoeme: With a Comprehensive Grammar of Yoeme Language*. Hippocrene Books.

Appendix A: Counterexamples and Test Results

24 verb forms ending in word-final -i

Not attested according to consultants

havoi, given as: tv. 'miss out on food'
Not a word.

yoki, given as: iv. 'be stained'
Instead there is a passive nominal participle, *yoka'i*, derived from *yoka*, 'to paint, mark;' *yoka* + *-ri* → *yoka'ari* → *yoka'i* after intervocalic r-drop and echo-vowel elimination. Without [r]-drop, *yoka'ari* is also an acceptable form. Perhaps dialect variation?

yumvaekai, given as: iv. 'tire one's self'
Instead there is a verb *yumia*, which can in principle be suffixed with propsective *-vae* and adverbial ppl. *-kai*, like any other verb. Suffixed form is *yumiavae*, though, not *yumvae*. Perhaps dialect variation?

Miscategorized

koptiachi, given as: iv. 'be desirable, be unable to keep one's hands off'
In fact an adjective, from *kopte* 'desire, v.' + *-iachi*, 'intensifying adjectivizer'.

kuviari, given as: iv. 'turned, screwed'
In fact nominal participle from *kuvia*, 'twist, turn, v.' + *-ri* 'passive nominalizer'.

sialapti, given as: iv. 'verdant, covered with green'
In fact syntactically adverbial, 'verdantly'.

nasonti, given as: iv. 'ruined, blotched'
in fact an adjective. Part of typical stative-adjective/inchoative/causative alternation triple *nasonti/nasonte/nasonta*, 'ruined/get.ruined/ruin'.

sikisi, given as: iv. 'become red'
In fact adverbial, from *sikii* 'red, adj.' + *-si* 'adverbializer'; should be spelled *sikiisi*.

sunwachi, given as: iv. 'horrible, fierce, frightening'
In fact an adjective.

tevai, given as: iv. 'starving'
In fact seems to mainly be a stem form *tevai-,* related to iv. *tevaure*, 'be.starving, be hungry,' probably formed with *-le/-re*. Still need to further verify it doesn't have independent uses. Occurs with *-tu* when inflected directly with *-kan*.

uhyoi, given as: iv. 1. 'be beautiful' 2. 'interesting'
In fact an adjective, variant of *uhyoli* with [r]-drop.

veakti, given as: iv. 'turned over'

In fact a misspelling of *viakte*, 'turn.over, roll, iv.' which has stem form *viakti-* under derivational affixation. Has causative variant *viakta*, 'roll, tv.'

<u>Puzzles with *-i ~ -ia,* also *-i ~ -e* in reduplication</u>

bwisi, given as: tv. 'hold, grasp it'

Apparently works as transitive verb as is; can also inflect with past imperfective *-n* directly. But becomes *bwibwis**e*** when reduplicated. Related to form *hi'ibwisia/hi'ibwise*, formed from indefinite detransitivizing prefix *hi'i-* plus *bwisia/bwise* meaning 'lean (on something)'. However *bwisia* is not good as a verb by itself.

bwiichi, given as: iv. 'smoky, filled with smoke'

Related to noun *bwichia*, 'smoke'. When reduplicated in predicate position, *bwichia* form surfaces: *bwibwichia*. In argument position, can be *bwiichi* or *bwichia*; must be the latter when marked accusative. However *bwichi* ok as predicate on its own, doesn't need possessive *-k* as for other nouns, with locative PP.

piiki, given as: iv. 'dive'

Sort of accepted as is, but consultants also said it could be pronounced *piike*, and would be 'better' that way. Dictionary gives reduplicated form *pipik**e***, consultants agreed.

tatai, given as: iv. 1. 'get hot, overheat' 2. 'be in heat'

Variably appears as *tattai*, *tatai(-), tata-* Seems to inflect with *-kan* without requiring *-tu* first, but *-tu* is **possible**. But also seems to have lots of stative uses, also stem-form uses. BUT: Seems to have a nominal meaning with ending *-ia*; got one sentence in argument position: *tatari**a**-ta*, 'heat.ACC'. Could *tattai* be derived from *tataria* + intervocalic [r]-drop + [a]-truncation?

weri, given as: iv. 'be related to'

Better translation might be 'be.siblings'. Seems to inflect like a verb mostly: doesn't need *-tu,* but *-tu* is **possible**, lacks attributive uses, requires subject relativization to behave modificationally. But when *-kan* is attached, surfaces as *weri**a**kan.* In the future can be *werine* or *weritune*; never surfaces as *weri**a**tune/weri**a**ne.* And can also act as a nominal argument, directly inflecting for plural marking: *Nau werim*, 'the related ones/siblings.'

<u>Uninvestigatable</u>

valumai, given as: tv. 'wash' (song language)

The example sentence in dictionary, from deer song, suggests an experiencer-subject predicate. Maria and Santos said it wasn't used in everyday Hiaki, and that deer song lyrics shouldn't be discussed.

<u>Genuine exceptions</u>

nattemai, given as: tv. 'ask about, ask for'

Genuine verb according to all tests, seems to be derived from *temai*. Probable explanation: final segment is not [i] but [j].

temai, given as: tv. ‘ask’

Still uninvestigated: Forms in (-)maachi(a)

mamachi, given as: iv. ‘be invisible, dark’

wattimachi, given as: iv. ‘able to remember’

maachi(a) is usually treated as a verb meaning ‘appear, be light’. In its suffixal form *-machi*, it has a couple of related readings, ‘seem, appear’, and evidential modal ‘should’.

22 adjectival and adverbial forms ending in word-final -e

Not attested according to consultants

amene, given as: adv. ‘(so) much, this much’
Consultants suggested that this might be a misspelling of *amuine*, the future form of the verb *amue* with the future suffix *-ne* meaning ‘suffice’.

Turns out to have been misspelled

kutvene, given as: adj. ‘dark (color)’
In fact should be spelled *kutvene**i**,* adj.

vae, given as: adv. ‘by means of water’
In fact probably adj. or maybe n. *va’ai* ‘dewy, moist’, like *tevai* and *tattai* above occurs mostly as stem form in compounds also has related iv. ‘vahe’, ‘be.foggy/dewy’.

Turns out to really be a verb

kutwatwatte, given as: adv. ‘in the dark’
In fact a verb, ‘be.dark’, iv.

vaamse, given as: adv. ‘in a hurry’
In fact a verb *vamse*, ‘be.in.a.hurry’, iv.

Misanalyzed

amae, given as: adv. ‘over there’
In fact, *ama* ‘over there’ + =*e* ‘from’, used in contexts of movement away.

iae, given as: adv. ‘with this’
In fact, *ia-* ‘this’ + *-e* ‘from’

keche, given as: adv. ‘how’
In fact *ket* ‘still’ + clitic =*ee* ‘2SG.SUBJ’, ‘You still...’, commonly used in greetings that get translated as ‘How are you?’, when what is really being asked is ‘Are you still well?’

kette, given as: adv. ‘still’
In fact *ket* ‘still’ + clitic =*te* ‘1PL.SUBJ’, ‘We still...’, commonly used in greetings.

Puzzling

ute, given as: adv. ‘fast’

Should be *u’ute*, ‘be strong’ iv., somehow related to *uttea*, n. ‘strength’. But *u’ute* also has some adverbial-looking uses.

vahiwe, given as: adv. ‘in three ways’

In fact should be something like *vahiweeye*, which looks something like ‘three-wander’, but doesn’t behave as a verb. Maria and Santos say *vahiweeye* also occurs as *vahiweyame*, which looks like what a subject relative of ‘*vahi-weeya*’, ‘three-wander’ should look like. Both behave like nouns meaning ‘three things’/‘three topics’, serving as arguments of predicates like ‘be.there’: “Q: What’s (discussed) in that book? Three topics are in that book.” When in object position and suffixed with accusative *-ta*, subject relative form must be used, *vahiweeye* cannot be used; suggests *vahiweeye* may be a clipping of *vahiweyame*, maybe. Kind of like *yoeme/yoemta*, **yoemeta.* Note that despite the translation, it’s inflecting as a singular noun — maybe best translation would be ‘a triplet (of topics)’.

woiwe, given as: adv. ‘in two parts’

Same situation as *vahiwe*, above; ‘a pair (of topics)’.

Genuine exceptions

ave, given as: adv. ‘almost’, though should be *have*, sometimes with [h]-drop

bweere, given as: adj. ‘big.PL’

kee, given as: adv. ‘not yet’

poloove, poove, given as: adj. ‘poor’ (borrowed from Spanish, though)

santene, given as: adv. ‘upright’

teeve, given as: adj. ‘tall’

vatte, given as: adv. ‘almost’

vette, given as: adj. ‘heavy’