Class 13
Infixation

11/19/21

1 Introduction

- **Infixes** are morphemes that (at least some of the time) attach **inside of morphological constituents**, rather than to the edge of a morphological constituent.
  - See Blevins (2014) for a recent overview, focusing on derivational infixes but applicable generally.

- Most of the time, people analyze infixation as being driven by phonological conditions on the position of the morpheme (see especially Yu 2007).

- Just like with allomorphy, mobile affixation, and PCAO, there is a debate in the literature about how these phonological conditions should be implemented:
  - Via **Subcategorization** (Yu 2007, Paster 2009, Kalin 2020)
  - By **P** ≫ **M** (McCarthy & Prince 1993a, Wolf 2008, *a.o.*; cf. Zukoff to appear)

- The arguments for and against are similar to those in the other domains:
  - Many infixal distributions seem to be governed by optimizing phonotactics, so P ≫ M.
  - Some infixal distributions seem to be non-/anti-optimizing, so Subcategorization.

- Kalin's (2020) arguments from the interaction between allomorphy and infixation are nuanced and novel, and may help untangle some of the persistent problems in adjudicating between the theories.
- But Zukoff's (to appear) [esp. Version 2:§5] introduction of alignment-driven (in addition to phonotactically-driven) infixation may re-complicate some of the questions.

2 Tagalog

- The classic case of (alleged) phonologically-driven infixation is *um*-infixation in Tagalog:
  → In Tagalog (Austronesian, Philippines), the actor focus (AF) morpheme /um/ alternates between a prefix and an infix (Schachter & Otanes 1972), seemingly to optimize syllable structure.

⋆ There has been a long debate about the data and the analysis. Here's how it went:

2.1 McCarthy & Prince's (1993a) analysis

- McCarthy & Prince (1993a:101) (following Prince & Smolensky [1993] 2004:§4.1) assume the following data:

(1) **Distribution of Tagalog AF *-um-* morpheme** (according to McCarthy & Prince 1993a)

a.	V-initial root:	/abot/	'reach for'	→	[<**um**>abot]
b.	C-initial root:	/sulat/	'call'	→	[s<**um**>ulat]
c.	CC-initial root:	/gradwet/	'graduate'	→	[gr<**um**>adwet]

- When the root is underlyingly vowel-initial (1a), the AF morpheme surfaces as a prefix.
- However, when the root begins in a consonant (1b,c), the AF morpheme surfaces as an infix.
  - With initial single consonants (6b), the AF morpheme surfaces after the root-initial C.

  - With initial clusters (6c), the AF morpheme surfaces after the cluster.

- McCarthy & Prince (1993a:103–104) argue that this distribution can be explained in full by the ranking:

(2) **M&P's Tagalog Ranking:** NoCoda ≫ Align-AF-L

- When there's a single root-initial consonant (3):
  - Prefixation puts the [m] of /um/ in coda position (3a), violating NoCoda.
  - Infixing past the root-initial /s/ (3b) allows that [m] to surface as an onset, creating no codas beyond the root-final one.
    - Codas can't be gotten rid of (3f) by an unfaithful phonological mapping (Faith-IO ≫ NoCoda), so the root-final coda has to stay.
    - This also means that you can't delete the AF /m/ (3e).
  - Since Align-AF-L is evaluated gradiently, infixing any further into the word (3c,d) will incur unnecessary violations.

(3) **Infixing past the first C to avoid a NoCoda violation:** *s<um>ulat* (1b)

/sulat, um/	Faith-IO	NoCoda	Align-AF-L	Align-Root-L	Contig
a. .<**um**>.su.lat.		**!		**	
b. ☞ .s<**u.m**>u.lat.		*	*		*
c. .su.l<**u.m**>at.		*	**!*		*
d. .su.la.t<**um**>.		*	**!***		
e. .<**u**>.su.lat.	*!	*		*	
f. .s<**u.m**>u.la.	*!		*		*

- This analysis predicts that /um/ will infix past an entire initial cluster (4c), because infixing past just the first consonant (4b) will create a coda.
  ∗ Assume rising-sonority clusters are parsed as complex onsets.

(4) **Infixing past the first CC to avoid a NoCoda violation:** *gr<um>adwet* (1c)

/gradwet, um/	NoCoda	Align-AF-L	Align-Root-L	Contig
a. .<**um**>.gra.dwet.	**!		**	
b. .g<**um**>.ra.dwet.	**!	*		*
c. ☞ .gr<**u.m**>a.dwet.	*	**		*
d. .gra.dwe.t<**um**>.	*	***!***		

- This analysis also predicts that you will not get infixation (5b,c), but rather prefixation (5a) for vowel-initial roots, because prefixation does not create a new coda.
  - Since prefixation and infixation are equivalent with respect to the relevant markedness constraint, the preferred alignment is able to surface.

(5) **Prefixation when it doesn't violate NoCoda:** *<um>abot* (1a)

/abot, um/	NoCoda	Align-AF-L	Align-Root-L	Contig
a. ☞ .<**u.m**>a.bot.	*		**	
b. .a.<**um**>.bot.	**!	*		*
c. .a.b<**u.m**>ot.	*	*!*		*
d. .a.bo.t<**um**>.	*	*!***		

2.2 Revising the data

- However, subsequent work showed that this isn't the whole story about the data:
  - Orgun & Sprouse (1999:204) find that, for CC-initial roots, at least some speakers exhibit variation in the site of infixation, between post-C_1 and post-C_2 (6c).
  - McCarthy (2003:91) clarifies, following the original description by Schachter & Otanes (1972), that all "vowel-initial" words surface with an epenthetic initial glottal stop (6a).

(6) **Distribution of Tagalog AF *-um-* morpheme**

a.	V-initial root:	/abot/	'reach for'	→	[ʔ<**um**>abot]
b.	C-initial root:	/sulat/	'call'	→	[s<**um**>ulat]
c.	CC-initial root:	/gradwet/	'graduate'	→	(i) [g<**um**>radwet] ~ (ii) [gr<**um**>adwet]

> ∗ We could alternatively assume that the initial glottal stops are underlying, which would simply collapse the (a) cases with the (b) cases. This may become useful later...

⋆ These facts transform the analysis from being driven by NoCoda to having to be driven by Onset.

- When the root begins in a single consonant (7):
  - Onset rules out full left-alignment of the affix (7a).
  - If Dep-C is next highest-ranked, it will rule out repairing that Onset via epenthesis (7b) as long as other candidates remain.
  - Since there are candidates (7c–e) that avoid these two problems at the expense just of Align-AF-L, the evaluation selects the infixal order where the affix is closest to the left (7c), i.e. after C_1.

(7) **Infixing past the first C to avoid an Onset violation:** *s<um>ulat* [.su.mu.lat.] (6b)

/sulat, um/	Onset	Dep-C	Align-AF-L	NoCoda	Align-Root-L	Contig
a. .<**um**>.su.lat.	*!			**	**	
b. .ʔ<**um**>.su.lat.		*!	*	**	***	
c. ☞ .s<**u.m**>u.lat.			*	*		*
d. .su.l<**u.m**>at.			**!*	*		*
e. .su.la.t<**um**>.			**!***	*		

- When the root is underlyingly vowel-initial (8):
  - There's no way to avoid an Onset violation without epenthesis, because both morphemes are vowel-initial (8a,d).
  - Since the desire to satisfy these two constraints is what motivates infixation (Align-AF-L violation), prefixation + epenthesis (8b) is optimal here.

(8) **No infixation if it doesn't fix Onset:** *<um>abot* [.ʔu.ma.bot.] (6a)

/abot, um/	Onset	Dep-C	Align-AF-L	NoCoda	Align-Root-L	Contig
a. .<**u.m**>a.bot.	*!			*	**	
b. ☞ .ʔ<**u.m**>a.bot.		*	*	*	***	
c. .ʔa.ʔ<**um**>.bot.		**!	***	**	*	*
d. .a.b<**u.m**>ot.	*!		**	*		*
e. .ʔa.b<**u.m**>ot.		*	**!*	*	*	*
f. .ʔa.bo.t<**um**>.		*	**!***	*	*	

- For roots beginning in two consonants, just like those beginning in one, infixation can avoid violation of both Onset and Dep.
  → The variable outputs can be derived if we have a variable ranking between the two lower-ranked constraints, NoCoda and Align-AF-L.

- When ALIGN-AF-L ≫ NOCODA (9):
  - It will be preferable to align the affix closer to the left (9c), even though it creates a coda, than to place it after the cluster (9d), which avoids the coda at the expense of an extra ALIGN violation.

(9) **Variable infix position in CC-initial roots:** ALIGN-AF-L ≫ NOCODA → *g<um>radwet* (6c.i)

/gradwet, um/	ONSET	DEP-C	ALIGN-AF-L	NOCODA	ALIGN-ROOT-L	CONTIG
a. .<**um**>.gra.dwet.	*!			**	**	
b. .ʔ<**um**>.gra.dwet.		*!	*	**	***	
c. ☞ .g<**um**>.ra.dwet.			*	**		*
d. .gr<**u.m**>a.dwet.			**!	*		*
e. .gra.dwe.t<**um**>.			***!***	*		

- On the other hand, when NOCODA ≫ ALIGN-AF-L (10), the reverse will be true:

(10) **Variable infix position in CC-initial roots:** NOCODA ≫ ALIGN-AF-L → *gr<um>adwet* (6c.ii)

/gradwet, um/	ONSET	DEP-C	NOCODA	ALIGN-AF-L	ALIGN-ROOT-L	CONTIG
a. .<**um**>.gra.dwet.	*!		**		**	
b. .ʔ<**um**>.gra.dwet.		*!	**	*	***	
c. .g<**um**>.ra.dwet.			**!	*		*
d. ☞ .gr<**u.m**>a.dwet.			*	**		*
e. .gra.dwe.t<**um**>.			*	***!***		

> ∗ Klein (2005:968–969) accounts for the variation in (6c) by positing a variable ranking between NOCODA and ***COMPLEXONSET**.
> - This predicts covariation between infix placement (post-C_1 vs. post-C_2) and the syllabification of medial clusters ($[...d]_\sigma[w...]_\sigma$ vs. $[...]_\sigma[dw...]_\sigma$): [g**um**.ra<u>d.w</u>et] vs. [gr**u.m**a.<u>dw</u>et].
>
> → There's no evidence for variable syllabification, so we should prefer the analysis with variation involving ALIGN.

⋆ This works, as long as we assume that medial rising sonority clusters are always parsed as complex onsets.
  - The activity of NOCODA means that we generate medial complex onsets.
  - If we needed to generate heterosyllabic parsing ([VC.CV] not [V.CCV]), we'd need *COMPLEXONSET to rank higher than NOCODA.

→ This would categorically result in the post-C_1 outcome (✓(9c)/✗(10c)), contrary to fact.

⋆ Indeed Zuraw (2007:298–299, fn. 27) asserts that medial clusters are always heterosyllabic in Tagalog.

→ This would break the analysis.
  - But other sources (e.g. Schachter & Otanes 1972, French 1988) aren't super clear on Tagalog's syllabification, so maybe it's still viable.

2.3 Zuraw (2007)

- Regardless of the syllabification issues, Zuraw (2007) adduces additional evidence that leads to a slightly different analysis, which sidesteps syllabification entirely.

- Zuraw (2007:esp. 295) finds that different types of initial clusters have different frequency distributions for the two different infix positions:
  - For ST clusters (and /sm/), speakers prefer the post-C_2 position to the post-C_1 position (11a).
  - But for CR clusters (except /sm/), speakers prefer the post-C_1 position to the post-C_2 position (11b).

(11) **Preferred infix site by cluster type**
  a. **ST:** #ST<**um**>V... > #S<**um**>TV...
  b. **CR:** #C<**um**>RV... > #CR<**um**>V...

- She proposes using CONTIGUITY constraints relativized to different cluster types to capture this difference.

- One way to capture frequency-based variation is by using **weighted constraints** in Harmonic Grammar (Legendre, Miyata, & Smolensky 1990, Smolensky & Legendre 2006), where the weights are fitted to the data using a Maximum Entropy (**MaxEnt**) model (Goldwater & Johnson 2003, Hayes & Wilson 2008).
  - Heuristically, the *relative weights* of the constraints determined by MaxEnt for the variable outputs would map onto the *relative rankings* of the constraints in OT if the differences were categorical.

→ So, abstracting away from the variation and assuming categorical outputs, we can derive the distribution by ranking a constraint against splitting ST clusters (CONTIG-ST) *above* ALIGN-AF-L:

- When there's an initial ST cluster (12):
  - The high ranking of ONSET and DEP-C continue to rule out left-aligning /um/ (12a,b).
  - The minimal infixation candidate (12c) is now ruled by relatively high-ranking CONTIG-ST.
  - For these clusters, the least displaced possible infixal candidate is thus (12d), where the /um/ lands after the initial cluster.

(12) **Post-C_2 position for ST-initial roots:** (nonce) *sp<um>in* (11a)

/spin, um/	ONSET	DEP-C	CNTG-ST	ALN-AF-L	ALN-ROOT-L	CNTG-CR
a. <**um**>spin	*!				**	
b. ʔ<**um**>spin		*!		*	***	
c. s<**um**>pin			*!	*		
d. ☞ sp<**um**>in				**		
e. spin<**um**>				***!*		

- When there's an initial CR cluster (13):
  - ONSET and DEP-C still to rule out left-aligning /um/ (13a,b).
  - But now, the fact that the minimal infixation candidate (13c) splits the cluster is not fatal, because it violates only low-ranked CONTIG-CR.
  - ALIGN-AF-L is now able to rule out all but the minimal infixation candidate (13d–g).

(13) **Post-C_1 position for CR-initial roots:** *g<um>radwet* (11b)

/gradwet, um/	ONSET	DEP-C	CNTG-ST	ALN-AF-L	ALN-ROOT-L	CNTG-CR
a. <**um**>gradwet	*!				**	
b. ʔ<**um**>gradwet		*!		*	***	
c. ☞ g<**um**>radwet				*		*
d. gr<**um**>adwet				**!		
e. grad<**um**>wet				**!**		*
f. gradw<**um**>et				**!***		
g. gradwet<**um**>				**!****		

- Since C-initial roots (14) and V-initial roots (15) don't involve clusters, their analysis works exactly the same as before.

(14) **Infixing past the first C to avoid an ONSET violation:** *s<um>ulat*

/sulat, um/	ONSET	DEP-C	CNTG-ST	ALN-AF-L	ALN-ROOT-L	CNTG-CR
a. <**um**>sulat	*!				**	
b. ʔ<**um**>sulat		*!		*	***	
c. ☞ s<**um**>ulat				*		
d. sul<**um**>at				**!*		

(15) **No infixation when it doesn't fix ONSET:** *ʔ<um>abot*

/abot, um/	ONSET	DEP-C	CNTG-ST	ALN-AF-L	ALN-ROOT-L	CNTG-CR
a. <**um**>abot	*!				**	
b. ab<**um**>ot	*!			**		
c. ☞ ʔ<**um**>abot		*		*	***	
d. ʔaʔ<**um**>bot		**!		***	*	
e. ʔab<**um**>ot		*		**!*	*	
f. ʔabot<**um**>		*		**!***	*	

→ One additional upshot of this analysis is that it is not dependent on syllabification.
- ○ Therefore, it is consistent with medial heterosyllabic parsing, unlike the NOCODA-based analysis.

- Zuraw (2007) actually uses high-ranked "ALIGN-STEM" (≈ALIGN-ROOT-L) to generate infixation, rather than high-ranked {ONSET ≫ DEP-C}.
  - ○ This amounts to saying that infixation is the *default* (ALIGN-ROOT-L ≫ ALIGN-AF-L).
  - → This creates a problem for "vowel-initial" roots.

- ONSET must dominate ALIGN-ROOT-L, because the reverse ranking would block epenthesis as a means of repairing an ONSET violation (because it introduces a pre-root segment). This means we'd need the ranking in (16).
  - ○ But the fact that ALIGN-ROOT-L ≫ ALIGN-AF-L means that we now predict infixation past the first C in these roots, **because infixation is the default** given the alignment ranking.

(16) **Incorrect prediction of ALIGN-ROOT-L ≫ ALIGN-AF-L for V-initial roots**

/abot, um/	ONSET	DEP-C	ALN-ROOT-L	CNTG-ST	ALN-AF-L	CNTG-CR
a. <**um**>abot	*!		**			
b. ab<**um**>ot	*!				**	
c. ☹ ʔ<**um**>abot		*	**!*		*	
d. ʔaʔ<**um**>bot		**!	*		***	
e. 💣 ʔab<**um**>ot		*	*		***	
f. ʔabot<**um**>		*	*		****!*	

- A way to circumvent the problem is to say that these roots are actually underlyingly /ʔ/-initial.
  - ○ If so, they will behave exactly like other C-initial roots, e.g. /sulat/.

⋆ However, we still have a Richness of the Base (Prince & Smolensky [1993] 2004) problem here:
- → If there were vowel-initial roots, they would be predicted to behave differently (as in (16)).
- ○ One could tell a story about lexicon optimization (Prince & Smolensky [1993] 2004, McCarthy 1998) based on the isolation forms, but it would be pretty tenuous.

2.4 P ≫ M vs. Subcategorization in Tagalog

- The analysis outlined above is a **P** ≫ **M** approach to infixation:

(17) **P ≫ M analysis**

a. Assuming the MAP (Zukoff to appear), the morphosyntax wants the AF morpheme to be a prefix (ALIGN-AF-L ≫ ALIGN-ROOT-L).
b. This succeeds in vowel-initial roots, because infixation would not improve on any phonological problems.
c. This fails in consonant-initial roots, because infixation can avoid more important phonological problem (ONSET and DEP-C).
d. The infixation site is regulated (gradiently) by (morpho)phonological alignment, subject to purely phonology CONTIGUITY constraints.

⋆ **What would a Subcategorization approach look like?**

- If we assumed that CC-initial roots uniformly infixed after C_2 (following McCarthy & Prince 1993a), then we could say that /um/ wants to attach *to the left of the first vowel/mora*:

(18) ACTOR.FOCUS ⇔ *um* / [(C)(C)_V...]$_{\text{STEM}}$ (cf. Paster 2009:19)

(19) ALIGN(/um/$_{\text{ACTOR.FOCUS}}$, R; μ_1, L) (cf. Yu 2007:91 on Leti)

- In order to account for the difference in behavior for different cluster types (per Zuraw 2007) with Paster-style subcat frames, we could consider specifying a distinct frame for #CR-roots (20a).
  → Since (20a) specifies *two* necessary segments whereas (20b) specifies only *one*, the Subset Principle / Elsewhere Condition should preferentially select (20a) when both are compatible.

(20) a. ACTOR.FOCUS ⇔ *um* / [C_R...]$_{\text{STEM}}$
b. ACTOR.FOCUS ⇔ *um* / [(C)(C)_V...]$_{\text{STEM}}$

  - If we assume that the elsewhere condition is *gradient* rather than *categorical* (not something usually assumed), we could assign some frequency distribution to both exponents in the case of #CR-root.
  - ∗ But this won't generate any frequency for post-C_1 infixation in the case of an #ST-root.

- Alternatively, we could consider the following:

(21) a. ACTOR.FOCUS ⇔ *um* / [(C)(C)_V...]$_{\text{STEM}}$
b. ACTOR.FOCUS ⇔ *um* / [C_...]$_{\text{STEM}}$

  - For V-initial stems, only (21a) would apply, so we'd still have categorical prefixation.
  - For C-initial stems, both frames would have the same result (because *after the first C* is the same place as *before the first V*), so we'd always get infixation in the right place.
  - For CC-initial stems, (21a) generates post-C_2 infixation while (21b) generates post-C_1 infixation.

- If all CC-initial stems had the same free-variation distribution between the two infixal positions, then this analysis would work.
  - If we assume that optional segments don't count for the determination of specificity of subcat frames, then the two are equally specific, and we might reasonably assume a 50/50 distribution.
  - ∗ If we follow Kalin & Rolle (2021), then indeed the optional segments shouldn't even be included.
  - → But this gives us no mechanism for generating the distinction between ST and CR stems.

⋆ It seems that subcategorization is going to have trouble accounting for the cluster-type differences.

3 Alignment-driven infixation and "anti-optimization"

- As always, one of the arguments against **P** ≫ **M** for infixation is that there are some cases which appear to be non-/anti-optimizing (Paster 2006, 2009, Yu 2007, Kalin 2020, Kalin & Rolle 2021, *a.o.*).
  - i.e., the structures resulting from infixation look like they're equally/more phonologically marked than what would have resulted from prefixation/suffixation.

- One of the cases frequently mentioned in this context is actor focus infixation in Atayal (Austronesian, Taiwan; Egerod 1965, Rau 1992; cf. Huang 2018) exemplified in (22).[1]
  - In this pattern, the morpheme *m* is infixed after the first consonant of the root.
  - This happens even when this position sits inside a long consonant cluster (22d–f).

[1] This discussion is based on §5.2 of Version 2 of Zukoff (to appear), available on Lingbuzz.

(22) **Atayal animate actor focus** (Yu 2007:35, ex. (45); data from Egerod 1965:263–266)

	Root	Root + AF	Gloss
a.	qul	q*m*ul	'snatch'
b.	kat	k*m*at	'bite'
c.	kuu	k*m*uu	'too tired, not in the mood'
d.	hŋuʔ	h*m*ŋuʔ	'soak'
e.	skziap	k*m*ziap	'catch'
f.	sbil	s*m*bil	'leave behind'

- Yu (2007) argues that this pattern cannot be described in terms of phonological optimization, and thus serves as counter-evidence to the **P** ≫ **M** model.
  - i.e., nothing phonological is being gained by infixation relative to prefixation — both (can) result in long consonant sequences.

→ However, this argument does not consider *alignment itself* as a trigger for output optimization.
  - If Align-Root-L outranks Align-AF-L, alignment on its own will generate infixation (23).

(23) **Atayal AF infixation**

/kuu, m/	Align-Root-L	Align-AF-L	Contiguity
a. **m**-kuu	*!		
b. ☞ k<**m**>uu		*	*
c. kuu-**m**		**!(*)	

∗ Despite the typical representation in (22), the *m* infix is usually/always preceded by a schwa / reduced vowel on the surface (Yu 2007:35, n. 12): i.e., *kmuu* might be more accurately transcribed [kəmuu].

→ If this schwa were underlying (see Huang 2018, contrary to most accounts), then this case might be analyzable as prosodic optimization, just like Tagalog. Indeed, Atayal does not allow (word-initial) onsetless syllables (Rau 1992:21).

- There's additional morphological evidence that may speak in favor of the alignment-driven infixation analysis:
  - Some active/agent stems built with the *m* morpheme display infixal ordering, but many show prefixal ordering instead or in addition (see the forms in Egerod 1965:263–267).
  - For some roots, both an infixal and prefixal form is attested, but with differences in meaning.

→ Per Rau (1992:37–38): infixal forms are *transitive* (24a) while prefixal forms are *intransitive/stative* (24b):

(24) **Infix/prefix alternations in Atayal** (see Egerod 1965:263–267, Blevins 2014:12)

a. *h<**m**>utaw* [həmutaw] 'drop'
b. ***m**-hutaw* [məhutaw] 'fall'

∗ Blevins (2014:11–12) asserts that there are two different /m/ morphemes, such that this is not a prefix/infix alternation of the same morpheme.

- This suggests that syntactic differences correlate with ordering differences (à la the MAP; Zukoff to appear).
  - Prefixal ordering is generated when the MAP (plus any attendant relevant default rankings) transmits the ranking Align-AF-L ≫ Align-Root-L.
  - Infixal ordering is generated when it transmits the reverse ranking Align-Root-L ≫ Align-AF-L (as shown in (23)).

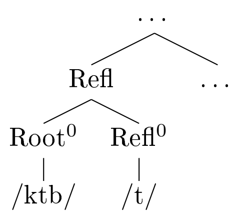
- This would be exactly equivalent to what we find with the Reflexive in Arabic (Zukoff to appear:§4):
  - Infixation occurs when the morpheme *is* the first head to combine with Root (25/26a).
  - Prefixation occurs when the morpheme *is not* the first head to combine with Root (25/26b).

(25) **Arabic Forms with Reflexive /t/** (for example root $\sqrt{ktb}$ 'write'; data from McCarthy 1981:384)

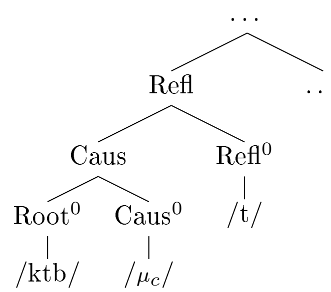
Position		Form	Proposed morphosyntax	Example form	Translation
a.	*Infixal*	VIII	Reflexive	***k*<*t*>*ataba***	'write, be registered'
b.	*Prefixal*	V	Reflexive of the Causative	***ṯak**attaba*	(*constructed form*)
		VI	Reflexive of the Applicative	***ṯak**aataba*	'write to each other'
		X	Causative of the Reflexive	*s**ṯak**taba*	'write, make write'

(26) **Syntactic structures with Reflexive**

a. Form VIII ***k*<*t*>*ataba***



b. Form V ***t*-*akat***$_c$*taba*



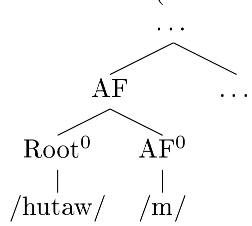
- Projecting this analysis onto the Atayal case (repeated in (27)), we get predict structures like (28).

(27) **Infix/prefix alternations in Atayal**

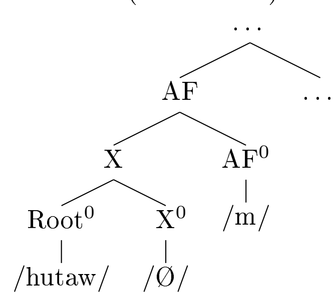
a. *h*<***m***>*utaw* [həmutaw] 'drop'
b. ***m**-hutaw* [məhutaw] 'fall'

(28) **Syntactic structures with Atayal AF**

a. *h*<***m***>*utaw* (transitive)



b. ***m**-hutaw* (intransitive)



- A structure like this would seem to make sense if X is something like Stative or some other valence-reducing head and AF is some sort of active Voice head or *v*.

⋆ Therefore, assuming that infixation in Atayal is driven by alignment constraints themselves, rather than prosodic optimization, we capture not only the surface phonological behavior, but also the morphosyntactically-correlated prefix/infix alternations.

→ This approach gives a roadmap for addressing other "non-/anti-optimizing" cases of infixation presented in Yu (2007) and elsewhere, in a way that is actually consistent with optimization.

4 Subcategorization and pivots

- Of course, subcategorization could get the Atayal case easily (minus accounting for the morphosyntax of the prefix/infix alternation):

(29) ACTOR.FOCUS ⇔ *m* / [C_...]$_{\text{STEM}}$

(30) ALIGN(/m/$_{\text{ACTOR.FOCUS}}$, L; C_1, R)

- In Yu's (2007) typological survey of infixation, he finds that this subcat frame (after the first consonant) is one of a fairly small number of locations where infixes can end up.
  → He calls these positions (or rather, the units delimiting the positions) **pivots**.
  ◦ The set of possible pivots is given in (31):

(31) **Possible pivots** (Yu 2007:67, adapted from Kalin & Rolle 2021:7; parentheses = uncommon)

Edge pivots		**Prominence pivots**
First consonant	(Last consonant)	Stressed foot
First vowel	Last vowel	Stressed vowel
(First syllable)	Last syllable	Stressed syllable

- According to this table, there are four units that can function as pivots:

(32) **Pivot units:** consonant (non-syllabic segment), vowel (syllabic segment), syllable, foot

- And there are three features that can identify these units:

(33) **Pivot features:** first (leftmost), last (rightmost), stressed

- Given that (non-syllabic) consonants can't be stressed, the cross-classification of the units and features is nearly fully fleshed out, with the only exception being first/last foot.

- Per Yu, languages can employ subcat frames aligning morphemes to either the left or the right of any of these pivots, but only these pivots.
  ◦ Absent are other conceivable phonological entities, like a specific consonant or a vowel with specific features, etc. (Yu 2007:218ff.).
  → This is noteworthy because Paster (2006, 2009) says that these kinds of entities *can* define subcat frames for PCSA (Kalin & Rolle 2021:§4.1).

⋆ While Yu identifies the pivots in (31) through his typological survey, the way he implements subcategorization (inviolable opposite-edge alignment) doesn't always actually refer to that pivot.
  → For example, he formalizes "pre-V_1" position using moras:

(34) **Tagalog** (19): ALIGN(/um/$_{\text{ACTOR.FOCUS}}$, R; μ_1, L) (cf. Yu 2007:91 on Leti)

- Another conceptual problem with his account is the status of first/last.
  → Kalin & Rolle (2021) are able to recast this as "closest" (which may or may not be conceptually stronger) by positing a step of edge-selection *before* infixation.

- Putting aside the subcategorization implementation questions, we should consider how this notion of pivots might relate to P $\gg$ M.

→ The fact that first/last consonant/vowel defines most cases of infixation, this seems largely compatible with the P $\gg$ M alignment-based view I introduced above.
  ◦ Displacement from the edge should be minimal (because of gradient alignment), so we should observe mostly first/last positions.
  ◦ In phonology-optimizing cases involving syllable structure, it should position itself with respect to consonants and vowels.
  ◦ In alignment-optimizing cases, it should position itself immediately inside the stem in terms of segments. If a language has relatively consistent phonological structures for roots/stems, this is likely to look like positioning relative to a consonant or vowel.

- The prominence pivots are a little trickier. Here's some data from Samoan in (35):
  - The Samoan plural is marked by infixal reduplication.
  - This morpheme is always CV, copying the stressed syllable, which it immediately precedes.
  - Stress is rigidly on the penultimate mora.

(35) **Samoan plural** (Yu 2007:24, citing Mosel & Hovdhaugen 1992:221–222)

tóa	'brave'	<to>tóa
máː	'ashamed'	<ma>máː
alófa	'love'	aː<lo>lófa
galúe	'work'	gaː<lu>lúe
aːvága	'elope'	aː<va>vága
atamái	'clever'	ata<ma>mái
maʔalíli	'cold, feel cold'	maʔa líli
toʔúlu	'fall, drop'	to<ʔu>ʔúlu

- Yu would account for this with the following subcategorization constraint:

(36) **Samoan:** ALIGN(/RED/$_{\text{PL}}$, R; $\acute{\sigma}$, L) [could also use (stressed/final) foot]

∗ Can we do this with P ≫ M? Yes, if *stress constraints* outrank ALIGN-PL-R.

- We can derive penultimate stress w/ foot-free stress constraints *LAPSER ([*$\sigma\sigma$#]) and NONFIN ([*$\acute{\sigma}$#]):

(37) **Simplex stress**

/alofa/			*LAPSER	NONFIN
a.		alofá		*!
b.	☞	alófa		
c.		álofa	*!	

- If these stress constraints, plus a constraint demanding that stress be identical between the derivative and its base (IDENT[stress]-BD), outrank ALIGN-PL-R, we derive the outcome where the reduplicant tucks in right before the stressed syllable (38e).
  - If it comes any further to the right, it will either displace the stress too far to the left (38a) or cause stress to fall on a different syllable than in the base (38b–d).
  - The pre-stress position (the antepenult) is the rightmost position that does not disrupt the original stress pattern ((38e) ≻ (38f)).

(38) **Stress and infixation**

BASE: [alófa] INPUT: /RED$_{\text{PL}}$, alofa/			*LAPSER	NONFIN	IDENT[stress]-BD	ALIGN-PL-R
a.		alófa<fa>	*!			
b.		alofá<fa>			*!*	
c.		alo<fa>fá		*!	*!*	**
d.		alo<fá>fa			*!	**
e.	☞	a<lo>lófa				****
f.		<a>alófa				*****!

→ This works well because we can say that the infix is oriented towards the same edge where stress is regulated (the right edge).
  - In most of the cases that Yu (2007:Ch. 4.7) identifies, it seems like the two edges match up.
  - Not all of them are amenable to such a simple analysis (but somebody should try...).

5 Infixation and allomorphy (Kalin 2020)

- To my mind, the best argument *against* P ≫ M for infixation comes from Kalin's (2020) work on the interaction between infixation and allomorphy.
  - Here are her findings, as summarized in her Appendix B (pp. 43–44):

(39) **On suppletive allomorphy involving an infix**

a. Suppletive allomorphs may differ with respect to pivot/placement (§3.1)
b. Suppletion involving an infix may be lexically, morphologically, phonologically, or prosodically conditioned (§3.2)
c. Conditions on exponent choice are distinct from an exponent's pivot/placement (§3.3)
d. Suppletive allomorphs share an edge orientation (§3.4)
e. Suppletion is conditioned based on the underlying form of the stem, at the stem edge identifiable via edge orientation (§3.5)
f. The surface (infixed) environment of an infix cannot condition suppletion (§3.6)

(40) **On non-suppletive infix allomorphy**

a. Non-suppletive infix allomorphy is conditioned only in surface (infixed) positions (§4.1)
b. No hypothetical position for an infix apart from its surface (infixed) position can (§4.2) induce non-suppletive allomorphy
c. An infix may condition phonological stem changes only in its surface (infixed) position (§4.3)

- The conclusions regarding non-suppletive allomorphy are completely consistent with a P ≫ M model, because they say that phonologically-*driven* allomorphy is local and transparent.
  - Some of the suppletive allomorphy conclusions are consistent too, especially (39d) given a system where alignment is sensitive to morphosyntactic features (and thus will apply equally to different exponents of the same morpheme).

∗ However, as Kalin (2020) points out, most of the conclusions about suppletive allomorphy do not appear to be consistent with P ≫ M.
  - E.g., if PCSA is governed by P ≫ M via something like Priority (Bonet, Lloret, & Mascaró 2007, Mascaró 2007), then PCSA should be able to be conditioned by infix location, not just the edge (39e,f).

- Many of Kalin's analyses need to be made more precise, and certain P ≫ M-based alternative analyses should be pursued further, but overall her results seem fairly strong.

⋆ Coupled with her refinements of subcategorization into "Conditions on Insertion" and "Conditions on Position" (Kalin & Rolle 2021), this seems like a compelling theory of the phonology-morphology interface (as much as I don't want to admit it).

References

Blevins, Juliette. 2014. Infixation. In Rochelle Lieber & Pavol Štekauer (eds.), *The Oxford Handbook of Derivational Morphology*, 1–33. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199641642.013.0009.

Bonet, Eulàlia, Maria-Rosa Lloret & Joan Mascaró. 2007. Allomorph Selection and Lexical Preferences: Two Case Studies. *Lingua. International review of general linguistics. Revue internationale de linguistique générale* 117:903–927.

Egerod, Søren. 1965. Verb Inflexion in Atayal. *Lingua* 15:251–282. doi:10.1016/0024-3841(65)90015-X.

French, Koleen Matsuda. 1988. Insights into Tagalog Reduplication, Infixation and Stress from Nonlinear Phonology. University of Texas at Arlington, Master's Thesis. https://www.proquest.com/docview/193705670.

Goldwater, Sharon & Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120. http://homepages.inf.ed.ac.uk/sgwater/papers/OTvar03.pdf.

Hayes, Bruce & Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39(3):379–440. doi:10.1162/ling.2008.39.3.379.

Huang, Hui-chuan J. 2018. The Nature of Pretonic Weak Vowels in Squliq Atayal. *Oceanic Linguistics* 57(2):265–288. doi:10.1353/ol.2018.0012.

Kalin, Laura. 2020. Infixes Really are (Underlyingly) Prefixes/Suffixes: Evidence from Allomorphy on the Fine Timing of Infixation. Ms., Princeton. https://ling.auf.net/lingbuzz/005581.

Kalin, Laura & Nicholas Rolle. 2021. Deconstructing Subcategorization: Conditions on Insertion versus Position. Ms., Princeton & Leibniz - ZAS, May 2021. https://ling.auf.net/lingbuzz/005975.

Klein, Thomas B. 2005. Infixation and Segmental Constraint Effects: UM and IN in Tagalog, Chamorro, and Toba Batak. *Lingua* 115(7):959–995.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky. 1990. Harmonic Grammar – A Formal Multi-Level Connectionist Theory of Linguistic Well-Formedness: Theoretical Foundation. ICS Technical Report 90-5, University of Colorado, Boulder, CO.

Mascaró, Joan. 2007. External Allomorphy and Lexical Representation. *Linguistic Inquiry* 38(4):715–735.

McCarthy, John J. 1981. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry* 12(3):373–418.

———. 1998. Morpheme Structure Constraints and Paradigm Occultation. In M. Catherine Gruber, Derrick Higgins, Kenneth Olsen & Tamra Wysocki (eds.), *Proceedings from the Annual Meeting of the Chicago Linguistic Society 32, Part II: The Panels*, 123–150. Chicago: Chicago Linguistic Society.

———. 2003. OT Constraints are Categorical. *Phonology* 20(1):75–138. doi:10.1017/S0952675703004470.

McCarthy, John J. & Alan Prince. 1986. Prosodic Morphology. *Linguistics Department Faculty Publication Series* 13 (1996 version).

———. 1993a. Generalized Alignment. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1993*, 79–153. Kluwer. doi:10.1007/978-94-017-3712-8_4.

———. 1993b. Prosodic Morphology I: Constraint Interaction and Satisfaction. *Linguistics Department Faculty Publication Series* 14 (2001 version).

Mosel, Ulrike & Even Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.

Orgun, Cemil Orhan & Ronald L. Sprouse. 1999. From "MParse" to "Control": Deriving Ungrammaticality. *Phonology* 16(2):191–224. doi:10.1017/S0952675799003747.

Paster, Mary. 2006. Phonological Conditions on Affixation. PhD Dissertation, University of California, Berkeley.

———. 2009. Explaining Phonological Conditions on Affixation: Evidence from Suppletive Allomorphy and Affix Ordering. *Word Structure* 2(1):18–37. doi:10.3366/E1750124509000282.

Prince, Alan & Paul Smolensky. [1993] 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, MA: Blackwell Publishing.

Rau, Der-Hwa Victoria. 1992. A Grammar of Atayal. PhD Dissertation, Cornell University.

Schachter, Paul & Fe T. Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: University of California Press.

Smolensky, Paul & Géraldine Legendre. 2006. *The Harmonic Mind*. Cambridge, MA: MIT Press.

Wolf, Matthew. 2008. Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model. PhD Dissertation, University of Massachusetts, Amherst.

Yu, Alan C. L. 2007. *A Natural History of Infixation*. Oxford: Oxford University Press.

Zukoff, Sam. to appear. The Mirror Alignment Principle: Morpheme Ordering at the Morphosyntax-Phonology Interface. *Natural Language & Linguistic Theory*. https://ling.auf.net/lingbuzz/005374.

Zuraw, Kie. 2007. The Role of Phonetic Knowledge in Phonological Patterning: Corpus and Survey Evidence from Tagalog Infixation. *Language* 83(2):277–316.