

# Inner and outer causatives in Austronesian: A diachronic perspective\*

Daniel Kaufman  
Queens College, City University of New York and  
Endangered Language Alliance

## SUMMARY

Austronesian causatives and transitivity related morphology present several challenges to linguistic theory. As discussed by Travis (2010), some Austronesian languages, like Malagasy, appear to transparently reflect a layered VP structure with an inner “lexical” causative and an outer productive causative. Other languages, like Tagalog, present unexpected morphology-syntax mismatches in this area. I discuss here certain puzzles in transitivity related prefixes from a historical perspective building on the idea that contrast preservation and epenthesis played a large role in the behavior of two key historical affixes: *\*paR-* and *\*paŋ-*.

## RÉSUMÉ

Les causatifs austronésiens et la morphologie liée à la transitivité présentent plusieurs défis à la théorie linguistique. Comme discuté par Travis (2010), certaines langues austronésiennes, comme le malgache, semblent refléter de manière transparente une structure de VP en couches avec un causatif “lexical” interne et un causatif productif externe. D’autres langues, comme le tagalog, présentent des inadéquations inattendues entre la morphologie et la syntaxe dans ce domaine. Je discute ici de certains casse-tête dans des préfixes liés à la transitivité d’un point de vue historique, reposant sur l’idée que la préservation des contrastes et l’épenthèse jouent un rôle important dans le comportement de deux affixes historiques: *\*paR-* et *\*paŋ-*.

---

\* I thank the editors of this volume for allowing me to serve up this humble offering in honor of Lisa Travis, who brought to light the problems discussed here and thus made this work possible.

## 1 INTRODUCTION

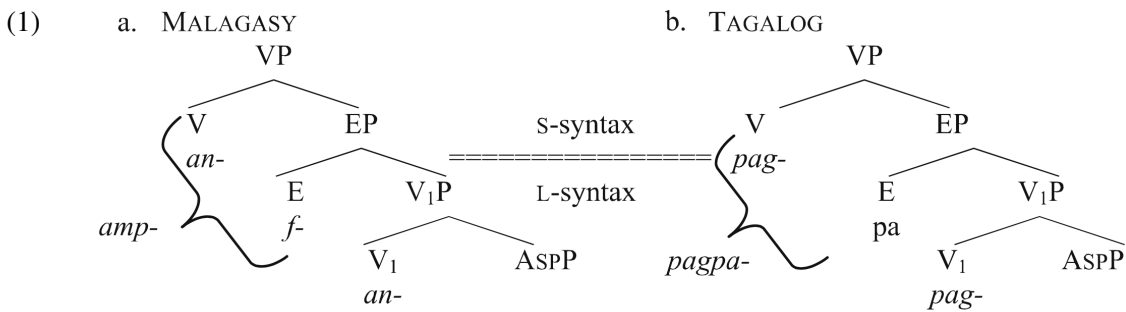
The vast majority of scholarship in Austronesian morphosyntax has centered on the analysis of the complex voice system. Another strand of work, advanced significantly by Lisa Travis and her students over the last several decades, has focused on other aspects of Austronesian morphology which relate to transitivity and causativization. This body of work has brought to light several fascinating problems in the morphology-syntax interface. In particular, we find that many Austronesian languages show evidence for two layers of structure in the verbal domain. An inner layer, which is in some sense more lexical in nature, reflects transitivity distinctions while an outer layer hosts productive causative morphology. Interactions between these layers and their interactions with voice morphology have been subject to scrutiny although only a tiny fraction of Austronesian languages have been investigated in this light.

Here, I will seek to reconcile Travis’s (2010) synchronic analysis of the inner and outer VP in Malagasy and Tagalog with a diachronic analysis of the Austronesian morphology involved in causation and transitivity. Kaufman (2009) posits that Proto-Austronesian (henceforth PAn) *\*pa*-CAUSATIVE underlies both the lexical and productive causative but that descendants of PAn *\*ŋ*-DISTRIBUTIVE and *\*R*- MIDDLE VOICE took on innovative transitivity-related functions by the time of Proto-Malayo-Polynesian, the language that gave rise to all the Austronesian languages outside of Taiwan. The historical perspective on inner and outer causatives leads to a unified segmentation and analysis of the prefix sequence which in turn allows for improved cross-linguistic comparison within Austronesian.

I summarize the crucial points of Travis’s analysis of inner and outer causatives in section 2. For reasons of space, I restrict my focus to the exposition of these ideas in Travis (2010) while setting aside for the time being earlier published and unpublished analyses. A brief history of the proto-morphemes *\*pa*-, *\*ŋ*- and *\*R*- is presented in section 3 while section 4 argues for the continued relevance of certain historical factors in understanding transitivity related morphology in Malagasy, Malay and Tagalog. Section 5 draws out some of the implications of what I analyze here as “*pag*-inversion” and section 6 concludes, noting persistent issues that require further attention.

## 2 BACKGROUND: EVENT STRUCTURE IN SYNTAX (TRAVIS, 2010)

In an extensive series of work, Travis argues for the structures in (1) as part of the Tagalog and Malagasy verb (Travis, 2010, p.168).



There are four layers shown in (1): (i) a higher VP layer, which houses productive causative morphology; (ii) an Event Phrase, which houses the morpheme *f-* in Malagasy and *pa-* in Tagalog; (iii) a lower VP layer which houses the lexical causative, which is identical in form to the productive causative in both languages; and (iv) an Aspect Phrase, involved in telicity distinctions. The Event Phrase in (1) marks the boundary between L(exical)-syntax and narrow syntax.

Key transitivity distinctions in the inner VP are exemplified for Tagalog by Maclachlan (1989) in (2) and for Malagasy by Travis (2010) in (3).<sup>1</sup>

- |     |    |  |    |  |
|-----|----|--|----|--|
| (2) | a. | t<um>umba<br><AV>fall<br>'X fall down'       | b. | m-pag-tumba [magtumba]<br>AV-CAUS-fall<br>'Y knock X down'   |
| (3) | a. | m-i-sitrika<br>AV-INTR-hide<br>'X fall down' | b. | m-an-sitrika [manitrika]<br>AV-CAUS-hide<br>'Y knock X down' |

For a substantial set of predicates in Tagalog, the actor voice <um> infix yields a monovalent predicate while another actor voice prefix, *mag-*, forms its bivalent counterpart.<sup>2</sup> Precisely the same alternation can be seen with the monovalent Malagasy predicate with *mi-* in (3a) compared with the bivalent predicate with *man-* in (3b). As Travis (2000, p.159) notes, this alternation, while applying to a large number of predicates, is not entirely regular. On the other hand, the outer VP is completely regular in both languages and serves as the locus of the productive causative. Remarkably, Malagasy preserves the transitivity distinction from the inner VP after the addition of the productive causative, as shown in (4).

(4)			PRODUCTIVE CAUSATIVE
	[-TR]	<i>misitrika</i> 'Y hide'	<i>mampisitrika</i> 'Z make X hide'
	[+TR]	<i>manitrika</i> 'Y hide X'	<i>mampanitrika</i> 'Z make Y hide X'

Travis analyzes the causative transitive form as shown in (5), where actor voice is attached above the structure shown above in (1).<sup>3</sup>

- (5) m- an- f- an- sitrika  
AV- ProdCaus- Event- LexicalCaus- root  
'Z make Y hide X'

Unlike Malagasy, Tagalog does not maintain the transitivity distinctions when the productive causative is applied (Travis, 2010, p.166) and the <um>/mag- distinction is neutralized to *magpa-*. Travis ar-

<sup>1</sup> These facts had long been noted for Tagalog, for example, by Bloomfield (1917, Vol.II p.233) and others, who struggled to characterize the <um>/mag- distinction in general terms. I am less familiar with the development of this generalization in the Malagasy descriptive literature.

<sup>2</sup> Note that *mag-* is generally understood as the combination of *pag-* and actor voice <um>.

<sup>3</sup> Note that the morphophonology of nasal substitution associated with the *aN-* prefix typically deletes the initial segment of the stem when voiceless. This is true not only for Malagasy but also for most other Austronesian languages that maintain a reflex of this prefix (Blust, 2004). Additionally, /f/ regularly surfaces as [p] before a nasal in Malagasy.

gues that the neutralization comes about via deletion of the inner causative, *pag-*, as shown in (6).

- (6) Siya ang **nag-pa-pag**-bukas sa akin ng bentana  
 3S.NOM NOM AV.TR.PRF-CAUS-CAUS-open OBL 1SG GEN window  
 ‘S/he’s the one who made me open the window.’

Travis motivates *pag-* deletion with a purely syntactic constraint relating to the doubly-filled comp filter. Here, I will explore an alternative, historically grounded approach to these facts that makes crucial use of epenthesis rather than deletion.

The morphology involved in the transitivity alternations above can be traced to a small handful of PAN or PMP prefixes although their cognacy and core functions are easily obscured. In section 3, I discuss the relevant morphology from a diachronic perspective. Taking this historical picture seriously, I attempt in 4 to unify the synchronic analysis of Malagasy, Tagalog and Malay/Indonesian (henceforth Malay) transitivity morphology.

### 3 THE PROTO-AUSTRONESIAN CAUSATIVE, DISTRIBUTIVE AND MIDDLE

Building on Ross (2002), I argue in Kaufman 2009 for reconstructing the following morphemes to PAN:

- (7) a. \**ɲ*-                              b. \**R*-                              c. \**pa*-                              d. \**ka*-  
             DISTRIBUTIVE                              MIDDLE                              **do**’/CAUS                              **have**’/STAT

The morphemes in (7) are largely independent of voice morphology, although there are important interactions which will be discussed below. The morpheme \**ɲ*- in (7a) plays a crucial role in pluractional marking in Philippine languages as well as instrumental nominalization but develops into purely transitivity-related functions in Malayo-Polynesian languages outside the Philippines (e.g. anti-passive in Mamuju, see Kaufman 2017). The morpheme \**R*- in (7b) forms reflexives, reciprocals (see Zeitoun, 2002), and typical middle voice predicates. The morpheme \**pa*- in (7c), functioned both as a typical causative but also as a verbalizer for a certain class of verbs (Ross, 2015). For instance, the Mayrinax Atayal roots in (8) cannot become verbs with the mere addition of actor voice <*um*>. Rather, they require prefixation of *pa-* to form a verbal stem that only then can take actor voice <*um*> (Huang, 2000, p.379).<sup>4</sup>

- (8) a. *maqabuɕij*  
       p<*um*>a-qaɕuɕij  
       <AV>CAUS-hat  
       ‘put on a hat’
- b. *maqunas*  
       p<*um*>a-qunas  
       <AV>CAUS-song  
       ‘sing’

The combination of these two morphemes, \**p<um>a-*, was reduced to \**ma-* from the earliest

---

<sup>4</sup> In many of the following examples, I gloss modern languages using the proto-morphemes posited here. This should not be taken to imply that these morphemes are all productive in the modern language. This is simply meant to clarify the historical development of these morphemes in each language and highlight certain cross-linguistic similarities that would otherwise be obscured.

reconstructable stages of Austronesian.<sup>5</sup> As can be gleaned from the Atayal examples above, *\*pa-* appears as a verbalizer on entity-denoting roots.

Some key morphophonological dependencies explain certain quirks in the distribution of *\*pa-* CAUS, *\*ŋ-* DISTRIBUTIVE and *\*R-* MIDDLE. No attested language allows reflexes of *\*ŋ-* and *\*R-* to attach to lexical roots without further prefixation. Rather, these mono-segmental prefixes serve as codas to preceding prefixal material. This gives rise to what can be called “dummy *pa-*”, whose only function is to serve as a host to these affixes. This can be seen in the Tagalog comparison in (9), where the plain actor voice is expressed with <um> but the distributive, *maŋinom*, contains a dummy *pa-* to host *ŋ-*.

- |     |    |  |    |   |
|-----|----|--|----|---|
| (9) | a. | <i>uminom</i><br><um>inom<br><AV>drink<br>'to drink' | b. | <i>maŋinom</i><br>p<um>a-ŋ-inom<br><AV>CAUS-DIST-drink<br>'to drink repeatedly/excessively' |
|-----|----|--|----|---|

The second quirk is that the reduction *\*p<um>a-* → *\*ma-* produced an uncomfortable homophony with a form having a near opposite interpretation that stemmed from the combination of actor voice with the stative-like prefix *\*ka-*, whose meaning I reconstruct as **have'**: *\*k<um>a-* → *\*ma-*. In several Formosan languages, e.g. Isbukun Bunun (Ross 2015, p.314), this results in two homophonous *ma-* prefixes, one that serves to mark active verbs and the other which marks stative verbs. In other morphological contexts (i.e. imperative, irrealis, non-actor voices) these functions are distinguished, thus creating paradigmatic pressure to avoid homophony throughout. Just as we find a dummy *\*pa-* to host the distributive and middle voice infixes, we also find an epenthetic use of these infixes to differentiate active *ma-* (from *\*p<um>a-*) from stative *ma-* (from *\*k<um>a-*). It appears that in PMP, active *ma-* may have already been eliminated in this way. Most Philippine languages employ *\*R-* in this epenthetic function, yielding a prefix like Tagalog *mag-* (from *\*p<um>aR-*) that has a valency increasing function but no discernible middle voice interpretation.

These twin dependencies create two sources for prefixes like Tagalog *mag-*. One contains a dummy *pa-* with a meaningful middle voice *g-* (from *\*R-*). The other contains a meaningful *pa-* CAUSATIVE with a dummy *\*R-* to maintain the distinction with stative *ma-* throughout the voice/aspect/mood paradigm. This leads to a paradoxical set of functions for prefixes like *mag-*, with the core problem illustrated by Tagalog in table 1, following Pittman (1966).

Some uses of *mag-* in Tagalog, notably the reflexive and reciprocal, are typically associated with valency decreasing morphology. Other uses of *mag-*, however, such as the causative and deliberate action, are typically associated with valency increasing morphology cross-linguistically. I suggest this is a natural outcome of the two types of morphological epenthesis found in *mag-*. The middle voice functions in (a) and (b) of table 1 involve dummy *\*pa-* while the causative functions in (c) and (d) involve a dummy middle infix. In other languages, the DISTRIBUTIVE *\*ŋ-* plays the epenthetic

<sup>5</sup> Ross (2015) reconstructs a verb class to PAN which he terms  $\emptyset$ /ma and which shows *\*ma-* prefixation in the actor voice and null prefixation in other voices. This class of verbs poses special challenges to reconstruction (Ross, 2015, p.333-4) but I believe it can be shown to derive from a verbalizing *\*pa-* in combination with the actor voice. The disappearance of the verbalizing *\*pa-* in the non-actor voice of many Austronesian languages may be tied to the inherently higher transitivity of those voices. More on this below.

Table 1: Tagalog *-um-* vs. *mag-* on identical stems (Pittman, 1966)

	<i>&lt;um&gt;</i>	<i>mag-</i>
a. Non-reflexive	<b>Reflexive</b>	(dummy * <i>pa-</i> )
b. Non-reciprocal	<b>Reciprocal</b>	(dummy * <i>pa-</i> )
c. Non-causative	<b>Causative</b>	(dummy * <i>R-</i> )
d. Impulsive	<b>Deliberate</b>	(dummy * <i>R-</i> )

role seen for \**R-* above. Malay and Malagasy, where reflexes of \**mag-* become generic transitive markers are particularly clear examples of this. We can compare the dummy \**ŋ-* in Malay (10a) with the meaningful middle \**R-* in (10b). In (10a), it is the \**pa-* which does the work, whereas in (10b) we see a reflex of the dummy \**pa-* whose only function is to host the middle voice.

(10)	a.	<i>məmbunuh</i>	b.	<i>bərbunuh</i>
		p<um>a-ŋ-bunuh		p<um>a-R-bunuh
		<AV>CAUS-DIST-kill		<AV>CAUS-MID-kill
		‘to kill s.o.’		‘to kill each other’

With this introduction to the key features of the relevant prefixes, we now explore the cognate morphology in Malagasy, Malay and Tagalog from a historical perspective.

#### 4 A DIACHRONIC APPROACH TO TRANSITIVITY MORPHOLOGY

As discussed above, \**pa-* functioned both as a productive causative and as a verbalizer for certain predicates from a very early point in Austronesian. Nonetheless, I adopt the analysis of *pa-* as a causative prefix regardless of where it appears in the structure. I locate modern reflexes of the two prefixes \**ŋ-* DISTRIBUTIVE and \**R-* MIDDLE in *v*, as they are transitivity related but clearly beneath voice in the verbal structure. The two layers within the prefix sequence examined below are represented as [CAUS<sub>1</sub> *v*<sub>1</sub>] for the outer layer and [CAUS<sub>2</sub> *v*<sub>2</sub>] for the inner layer.

##### 4.1 MALAGASY

Malagasy has a remarkably transparent morphology that faithfully indicates both the transitivity of the inner VP as well as the presence of a higher causative head. In table 2, we see the relevant fragment of Malagasy prefixal morphology beginning with the actor voice and followed by two sequences of [CAUS *v*].<sup>6</sup> In the first two rows of table 2, we find intransitive and transitive morphology without the productive outer causative. In the next two rows, we find the same distinction but with the causative layer. The relevant morphemes are shown in their historical form and the actual

<sup>6</sup> In the following discussion, I reduce a complex set of patterns to simplistic feature [±TRANSITIVE]. The truth, of course, is messier than that but for present purposes I take the simple transitivity alternation as a core function of the inner causative. Note also that in the following tables, as elsewhere, I represent the relevant formants using the reconstructed forms.

resulting forms are shown on the far right.<sup>7</sup>

Table 2: Malagasy prefix sequence (componential)

	VOICE	[CAUS <sub>1</sub>	v <sub>1</sub> ]	[CAUS <sub>2</sub>	v <sub>2</sub> ]	
[-CAUS, -TR]	<um>	∅	∅	pa-	R-	→ <i>mi-</i>
[-CAUS, +TR]	<um>	∅	∅	pa-	ŋ-	→ <i>maN-</i>
[+CAUS, -TR]	<um>	pa-	ŋ-	pa-	R-	→ <i>mampi-</i>
[+CAUS, +TR]	<um>	pa-	ŋ-	pa-	ŋ-	→ <i>mampaN-</i>

Both types of epenthetic morphology in the inner VP are indicated in boxes in the paradigm tables.<sup>8</sup> On this analysis, the *pa-* formant in *mi-* has no semantic content; it is rather the \**R-* that marks the predicate as [-TRANSITIVE]. In contrast, it is the historical distributive infix \**ŋ-* which is semantically empty in *maN-*, as this prefix now marks transitive predicates without carrying distributive or pluralational meaning. Here, the *pa-* does the work, making the predicate [+TRANSITIVE].

Note that \**ŋ-* is always epenthetic on this historical account, as it bears no trace of its original distributive function in Malagasy. As such, there is little reason to believe that \**ŋ-* has not been reanalyzed with preceding material as part of a general transitivity marker. Table 3 shows a simpler, more canonical analysis that treats the reflexes of \**paŋ-* and \**paR-* as unitary prefixes. The former marks intransitivity (at the relevant layer) and the latter adds an object.

Table 3: Malagasy prefix sequence (fused)

	VOICE	[CAUS+v] <sub>1</sub>	[CAUS+v] <sub>2</sub>	
[-CAUS, -TR]	<um>	∅	paR- INTR	→ <i>mi-</i>
[-CAUS, +TR]	<um>	∅	paŋ- +OBJ	→ <i>maN-</i>
[+CAUS, -TR]	<um>	paŋ- +OBJ	paR- INTR	→ <i>mampi-</i>
[+CAUS, +TR]	<um>	paŋ- +OBJ	paŋ- +OBJ	→ <i>mampaN-</i>

## 4.2 MALAY

At first glance, the Malay paradigm, shown in 4, forms a minimal pair with Malagasy.<sup>9</sup>

<sup>7</sup> Correspondences like PMP \**maR-* and Malagasy *mi-* may appear distant but all the historical reflexes shown here are regular and should not be controversial (Dahl, 1995, p.174). Compare *rami* ‘k.o. tree’ < PAn \**damaR*; *reni* heard < PMP \**deŋeR*; *turi* ‘sleep’ < \**turuy* < PAn \**tuduR* (Blust and Trussel, ongoing).

<sup>8</sup> For reasons discussed below, I remain agnostic as to what components of the outer VP are epenthetic.

<sup>9</sup> The path by which PMP \**maR-* became Malay *bər-* is discussed by Adelaar (1992) and Van den Berg (2004), who disagree on certain details but agree on cognacy, which is uncontroversial.

Table 4: Malay prefix sequence (componential)

	VOICE	CAUS <sub>1</sub>	v <sub>1</sub>	CAUS <sub>2</sub>	v <sub>2</sub>		
[-CAUS, -TR]	<um>	∅	∅	pa-	R-	→	<i>bər-</i>
[-CAUS, +TR]	<um>	∅	∅	pa-	ŋ-	→	<i>məŋ-</i>
[+CAUS, ±TR]	<um>	pa-	ŋ-	pa-	R-	→	<i>məmpər-</i>

Cognate forms correspond to the Malagasy except that the productive causative forms do not make any transitivity distinction in the inner VP. Unlike in Malagasy, the inner VP shows a reflex of \**paR-* in the [+CAUS, +TR] paradigm.

We could attempt to impose a Malagasy analysis on Malay so that \**paR-* is an intransitive marker and \**paŋ-* adds an object, but if Malagasy represents the expected state of affairs, what prevents a similarly transparent form *məmpəŋ-* for causative transitives in Malay (cf. Malagasy *mampaN-*)? In fact, what appears to be a minimal difference between Malay and Malagasy represents a deeper divergence between the two systems. Specifically, it is the *inner* layer which hosts the productive causative in Malay, not the outer layer. This emerges clearly in the patient voice causative, where *məŋ-* is replaced completely by *di-*, as seen in (11).

- (11) a. mem-per-kecil-kan                      b. di-per-kecil-kan  
 AV-CAUS-small-APPL                      PV-CAUS-small-APPL  
 ‘X to make smaller’                      ‘X to be made smaller’

The outer *məŋ-* thus seems to express no more than the actor voice in causatives. This contrasts with Malagasy where the outer reflex of PMP \**paŋ-* is present independent of voice and co-occurs with the patient voice, as can be seen in (12) (modifying Randriamasimanana’s 1986, p.146, segmentation for consistency).

- (12) n-am-pi-asa-in                      i Paoly i Jeanne  
 PAST-CAUS-INTR-work-PV Paul Jean  
 ‘Paul was exploiting Jean’s work.’

The historical prefix combination \**pa-ŋ-* CAUS-DIST is thus primarily a base for the agent voice in Malay but a voice-independent transitivity marker in Malagasy. In neither language has the distributive/pluractional semantics of \**ŋ-* been maintained.

The fact that Malay *məŋ-* is primarily an actor voice morpheme is key to understanding its position. Both on the basis of cross-linguistic comparison and scope considerations, it not surprising that voice is positioned outside the productive causative. We can say then that the reanalysis of *məŋ-* as a basic exponent of the actor voice is a precondition for its external position.

### 4.3 TAGALOG

Tagalog, shown in table 5, differs in several respect from the previous languages. First, the historical \*<um> ACTOR VOICE infix is still productive as an independent morpheme. Second, there are two very distinct functions for *mag-*, as discussed above in section 3. It appears in some contexts to



Table 5: Tagalog prefix sequence (componential)

	VOICE	CAUS <sub>1</sub>	v <sub>1</sub>	CAUS <sub>2</sub>	v <sub>2</sub>	→	
[-CAUS, -TR]	<um>	∅	∅	∅	∅	→	<um>
[-CAUS, -TR]	<um>	∅	∅	pa-	R-	→	mag <sub>1</sub>
[-CAUS, +TR]	<um>	∅	∅	pa-	R-	→	mag <sub>2</sub>
[+CAUS, ±TR]	<um>	pa-	R-	pa-	∅	→	magpa-

decrease the valency of the predicate (*mag<sub>1</sub>*) and in other cases to increase its valency (*mag<sub>2</sub>*). Middle voice \*R- attaches to the dummy host \*pa- in *mag<sub>1</sub>* but in *mag<sub>2</sub>*, causative \*pa- has fused with \*R- to avoid homophony with stative *ma-*. As in the above tables, the epenthetic morphemes in the inner VP are highlighted in boxes.<sup>10</sup> Similar to Malay but unlike Malagasy, Tagalog neutralizes transitivity distinctions in the causative. Both types of causative, intransitive and transitive, are expressed with *magpa-* in Tagalog, as mentioned earlier.

In order to better understand the argument for epenthesis, it is necessary to step back and look at a slightly larger slice of Tagalog morphology. In (13), we see the Tagalog reflexes of the PMP morphemes relevant here (see Ross, 1995, 2002; Wolff, 1973; Blust, 2003; Zeitoun and Huang, 2000; Kaufman, 2009, 2012, for discussion).

(13)	PMP	Tagalog
CAUSATIVE	*pa-	pa-
STATIVE ( <b>have'</b> )	*ka-	ka-
ACTOR VOICE	*<um>	<um>
MIDDLE	*R-	g-

In (14), we see the Tagalog outcomes of common combinations of the morphemes in (14), with the epenthetic morphemes glossed as such.

(14) a.	*k<um>a-	→	ma-	e.	*p<um>a-R-pa-	→	magpa-
	<AV>STA-				<AV>CAUS-EPEN-CAUS-		
b.	*p<um>a-R-	→	mag <sub>1</sub>	f.	*p<um>a-ka-	→	maka-
	<AV>EPEN-MID-				<AV>CAUS-STA		
c.	*p<um>a-R-	→	mag <sub>2</sub>	g.	*k<um>a-pa-R-	→	mapag-
	<AV>CAUS-EPEN-				<AV>STA-CAUS-MID-		
d.	*k<um>a-pa-	→	mapa-				
	<AV>STA-CAUS						

When the inner or “lexical” \*pa- combines with the actor voice, Tagalog shows *mag-*, as seen in (14c). When a productive causative is added to this, the result is *magpa-*, as seen in (14e). In

<sup>10</sup>Tagalog also has a reflex of \*p<um>ag-, namely *mag-*, but this retains its original distributive/pluractional semantics and, unlike Malay and Malagasy, is not implicated in basic transitivity alternations. It is thus excluded from the following discussion.

both (14c) and (e), it is only the semantically empty *g-* which differentiates these active forms from stative forms, as aphaeresis (dating all the way to PAN) applies to the combination of actor voice *\*<um>* and stative *\*ka-* yielding *ma-* (Ross, 1995). Perhaps the strongest evidence for contrast preservation as the basis of this pattern is the lack of *\*R-* epenthesis in (14f). Here, *\*p<um>a-* uncharacteristically yields *ma-* rather than *mag-* but, crucially, this only occurs before the stative *ka-*. Because stative *ka-* can only occur once in the sequence, there is no ambiguity here between *ma-* as the result of *p<um>a-* versus *k<um>a-*, it can only be the former.

Nonetheless, *\*R-* epenthesis is clearly fossilized in Tagalog: the inner causative regularly appears as *pag-* and the outer causative as *pa-* with very few exceptions.<sup>11</sup> On this account, we must explain why the “inner” causative, *mag-*, appears external to the “outer” causative in (14e), *pa-*. We cannot rely completely on *mag-* being a basic exponent of ACTOR VOICE because, unlike Malay, Tagalog preserves *<um>* as the basic exponent of actor voice.

We can start with the fact that the expected output of *\*p<um>a-pa-R-*, which is *mapag-*, would lose the contrast with the output of the stative/abilitative *\*k<um>a-pa-R-*, shown in (14g); we face the same problem here of distinguishing between *pa-* and *ka-* in combination with *<um>*. If homophony avoidance is the motivation, we do not expect to see an epenthetic *\*R-* in non-actor voices, as it is the actor voice infix which triggers the potentially contrast destroying aphaeresis (deletion of initial *pu* in *\*p<um>a-* and *ku* in *\*k<um>a-*). This prediction is partially met, as seen in (15).

- |      |    |                   |    |               |
|------|----|-------------------|----|---------------|
| (15) | a. | <i>magluto</i>    | b. | <i>lutuin</i> |
|      |    | p<um>ag-luto      |    | lutu-in       |
|      |    | <AV>CAUS/MID-cook |    | cook-PV       |
|      |    | ‘X to cook’       |    | ‘to cook X’   |

However, anti-homophony only predicts that the *\*R-* disappears while in fact the entire *\*paR-* disappears in the patient voice. This pattern is highly regular across Austronesian languages: predicates that employ inner *\*pa-* to form simple transitives in the actor voice, typically do not employ it in this function in non-actor voices. Tagalog employs a reflex of (plain) *\*pa-* in non-actor voices only for the productive “outer” causative. The transitivity distinction seen earlier between *<um>* and *mag-* in the actor voice is subsequently often lost in the non-actor voices. This is most likely connected to the inherently higher transitivity of the non-actor voices, a phenomenon which has been explored in the ergativity literature (De Guzman, 1988; Liao, 2004; Aldridge, 2004, inter alia). The actor voice, because of its inherently lower transitivity, requires the valency increasing *pa-*, while this is not required by the bona fide transitive voices.

Recall, however that the complex prefix *pag-* can also consist of the middle voice with *pa-* as a dummy host, corresponding to *mag<sub>-1</sub>* above. We expect the underlying *pag-* of *mag<sub>-1</sub>* to remain in all voices, as it is motivated not by transitivity *pa-* but rather by the middle voice. This is borne out in cases such as the following. In (16) and (17), we see middle *mag<sub>-1</sub>* and transitive *mag<sub>-2</sub>*.

<sup>11</sup>I adopt the convention of treating Tagalog *pag-* as a single morpheme with a bipartite gloss CAUS/MID. Strikeout indicates which component is epenthetic. Thus, ~~CAUS~~/MID is the middle with dummy *pa-* and CAUS/~~MID~~ is causative *pa-* with an epenthetic *g-*.

- (16) *Magsama kayo!*  
 p<um>ag-sama kayo  
 <AV>CAUS/MID-together 2P.NOM  
 ‘Be together (with each other)!’
- (17) *Magsama ka ng kaibigan!*  
 p<um>ag-sama ka ng kaibigan  
 <AV>CAUS/MID-together 2S.NOM GEN friend  
 ‘Bring a friend!’

In the patient voice of this predicate, we only find *pag-* with the middle function, as in (18). There is no transitive patient voice counterpart of (17) with *pag-*, as demonstrated by the ungrammaticality of (19).

- (18) *Pagsamahin mo sila!*  
 pag-sama-hin mo sila  
 CAUS/MID-together-PV 2S.GEN 3P.NOM  
 ‘Join them together!’
- (19) \**Pagsamahin mo sila ng kaibigan!*  
 pag-sama-hin mo sila ng kaibigan  
 CAUS/MID-together-PV 2S.GEN 3P.NOM GEN friend  
 (For, ‘Make them bring a friend!’)

Were this pattern completely regular, the mystery of the disappearing and reappearing *pag-* would be largely solved. When *pag-* functions as a valency increaser, it is omitted in the non-actor voices because these voices are inherently transitive. When *pag-* functions as a middle voice marker, it may appear in non-actor voices providing they too bear the middle voice meaning. Unfortunately, things are not so simple. Other predicates appear to maintain transitivity *pag-* (corresponding to *mag-<sub>2</sub>*) in non-actor voices. For instance, as discussed by Maclachlan (1989) and Travis (2000), the root *sabog*, glossed here as ‘disperse’, obtains two different meanings depending on whether it takes *pag-* or not. Without *pag-*, derivations of *sabog* relate to ‘explode’; with the *pag-* prefix, *sabog* obtains the meaning ‘to scatter (something)’, as seen in (20).

- |      |    |   |    |   |
|------|----|---|----|---|
| (20) | a. | <i>sumabog</i><br>s<um>abog<br><AV>disperse<br>‘X to explode’ | b. | <i>magsabog</i><br>p<um>ag-sabog<br><AV>CAUS/MID-disperse<br>‘X to scatter something’ |
|------|----|---|----|---|

The non-actor voice counterpart of (20b) is *i-sabog* (CV-disperse) and does not contain *pag-*. However, in the patient voice causative counterparts, shown in (21), the transitivity *pag-* plays the same role as it does in the actor voice, contrary to our previous pattern (Travis, 2010, p.166).

- (21) a. *pasabugin*  
 pa-sabog-in  
 CAUS-disperse-PV  
 ‘cause X to explode’
- b. *pagsabugin*  
 pag-sabog-in  
 CAUS/MHD-disperse-PV  
 ‘cause X to scatter something’  
 ‘cause someone to scatter X’

The reappearance of *pag-* in the causative is accounted for by K. Ross (1993) (cited by Travis 2010, p.183) as a type of agreement with a moved argument. When the outer (causer) agent moves, *pag-* surfaces in the higher VP shell. Movement of the inner agent to the position of the nominative phrase triggers the appearance of *pag-* in the inner VP shell. Travis (2010) builds on this idea but analyzes it as deletion of *pag-* in the shell containing an unmoved argument due to the Doubly Filled Comp Filter. On this view, the prefix sequence appears as in table 6.

Table 6: The distribution of Tagalog *pag-* following K. Ross (1993) and Travis (2010, p.183)

	PROD CAUS	EVENT	LEX CAUS	NOM phrase	unmoved arg
AV, L-CAUS	∅	∅	pag-	Agt	Pat
PV, L-CAUS	∅	∅	∅	Pat	Agt
AV, P-CAUS	pag-	pa-	∅	outer Agt	inner Agt
PV, P-CAUS	∅	pa-	pag-	inner Agt	outer Agt

The movement/non-movement based approach is elegant but faces a difficulty in the claim that *pa-* is merely an event marker while the true outer causative is *pag-*. Tagalog also makes use of *pag-* to form gerunds. The only difference between the non-causative and causative gerund is the absence and presence of *pa-*. The same holds true for circumstantial voice causatives.

- (22) a. *pagluluto*  
 pag-lu~luto  
 GER-TR~cook  
 ‘cooking’
- b. *pagpapaluto*  
 pag-pa~pa-luto  
 GER-TR~CAUS-cook  
 ‘making someone cook’

There are also bare *pa-* forms that contain a predictable causative meaning and are clearly derived via the outer causative. Schachter and Otnes (1982, p.105) cite the following forms:

- (23) a. *pa-dala* CAUS-bring ‘something caused to be brought’  
 b. *pa-luto* CAUS-cook ‘something caused to be cooked’  
 c. *pa-gawa* CAUS-make ‘something caused to be made’  
 d. *pa-abot* CAUS-hand.over ‘something caused to be handed over’

These data, among others, make it difficult to avoid the conclusion that *pa-* is the true outer causative in Tagalog and that *pag-* is the inner causative.

It seems that the distribution of *pag-* in non-actor voice causatives has never been fully described and it is here that we find considerable hesitation and variation across native speakers. The structure cited by Travis (2010) (see also Ramos 1974 and Schachter and Otnes 1982, p.326) is shown in (24a) but the form in (24b) is far more common and many speakers accept it as the only grammatical

possibility for certain verbs in the patient voice causative. Surprisingly, we also find the form in (24c), which reflects the inverted order of inner and outer causative seen in the actor voice.<sup>12</sup> The one form that is consistently rejected is that with *pag-* sandwiched by two *pa-* morphemes, as in (24d).<sup>13</sup>

- |      |    |   |    |  |
|------|----|---|----|--|
| (24) | a. | pa-pag-aral-in<br>CAUS-CAUS/MID-study-PV<br>'to let/make X study' | b. | pag-aral-in<br>CAUS/MID-study-PV<br>'to let/make X study'                  |
|      | c. | pag-pa-aral-in<br>CAUS/MID-CAUS-study-PV<br>'to let/make X study' | d. | *pa-pag-pa-aral-in<br>CAUS-CAUS/MID-CAUS-study-PV<br>'to let/make X study' |

The preferred form, (24b), shows that the deletion posited by Travis must be part of the current analysis as well, although it appears to apply to *pa-* rather than *pag-*. Assuming that the semantically and syntactically transparent form in (24a) is the basis for the other forms, what triggers *pa-* deletion in (24b) and inversion in (24c)? Recall that lexical *pag-* is typically omitted in the plain non-actor voice forms, e.g. *mag-luto* AV-cook, but *lutu-in* cook-PV. It can thus serve double duty as both inner and outer causative in the patient voice without merging contrasts in the larger paradigm. Whatever the ultimate analysis, it seems that the quirky behavior of *pag-* in causative and non-causative contexts should receive a unified explanation.<sup>14</sup> As for the inverse form in (24c), I have argued that this order appears in the actor voice to maintain contrast with the stative/abilitative *mapag-*. Although we've seen that both the epenthesis and the inversion are often absent in the non-actor voice, it is possible that the tendency towards paradigm uniformity has spread inversion into certain non-actor voice constructions as well. For instance, we already saw in (22) how *pag-* precedes *pa-* in the gerund, which does not employ actor voice morphology at all.

## 5 CONCLUSION

I have explored transitivity related morphology in Malagasy, Malay and Tagalog from a historical perspective while considering certain consequences for a unified synchronic analysis. I have not attempted here a definitive formal analysis but I suggest that certain puzzles may obtain an explanation through the lens of the historical components of the *\*paR-* and *\*paŋ-* prefixes, which are so central to understanding transitivity alternations in Austronesian. I have also shown that Tagalog

<sup>12</sup>An attestation of this construction from a newspaper article can be seen in (i) (with simplified glossing):

(i) p<in>ag-pa-pa-bayad-∅ ng piyansang may total na P45,000 si H.  
<BEG>CAUS-MID-CAUS-IMPRF-pay-PV GEN bail:LNK EXT total LNK P45,000 NOM H.  
'H. is being made to pay a bail of a total 45,000 pesos.'

From: <http://www1.bomboradyo.com/boc-official-asawa-guilty-sa-kaso-ukol-sa-saln-ombudsman/>

<sup>13</sup>A small handful of attestations of this were found but were assumed to be errors. Similarly, we do not find *pa-* sandwiched between two *pag-* prefixes.

<sup>14</sup>On the other hand, the transparent form in (ia) is ambiguous between prefixation of the causative and CV-reduplication of *pag-* in the imperfective. This makes the infinitive form in (24a) similar to the imperfective of form (24b) (although vowel length distinctions may differentiate the two in dialects that maintain it).

contains more variation and complexity than previously acknowledged. While a proper analysis of Malagasy and Malay appears well within reach, much work remains to sort out the variation in Tagalog with regard to the combinatorics of *pag-* and *pa-* in the non-actor voices. It may ultimately be the case that rather than insertion or deletion of *pag-* in Tagalog, we have movement from the inner position to an outer position, but motivating this movement in gerunds and certain non-actor voice variants remains to be explored. This is clearly a rich area for further cross-linguistic research within Austronesian and a fascinating laboratory for morphological change.

## REFERENCES

- Adelaar, A. K. (1992). *Proto Malayic: The reconstruction of its phonology and parts of its lexicon and morphology*. Number 119. Dept. of Linguistics, Research School of Pacific Studies, the Australian National University.
- Aldridge, E. (2004). *Ergativity and Word Order in Austronesian Languages*. PhD thesis, Cornell University.
- Bloomfield, L. (1917). *Tagalog Texts*. University of Illinois Press, Urbana.
- Blust, R. (2003). Three notes on early Austronesian morphology. *Oceanic Linguistics*, 42(2):438–478.
- Blust, R. (2004). Austronesian nasal substitution: a survey. *Oceanic Linguistics*, (43):73–148.
- Blust, R. and Trussel, S. (ongoing). The Austronesian Comparative Dictionary, web edition.
- Dahl, O. C. (1995). Predicate, subject, and topic in Malagasy. *Oceanic Linguistics*, 35(2):169–179.
- De Guzman, V. (1988). Ergative analysis for Philippine languages: an analysis. In McGinn, R., editor, *Studies in Austronesian Linguistics*, number 76 in Monographs in International Studies Southeast Asia Series, pages 323–346. Ohio University Center for International Studies, Athens, Ohio.
- Huang, L. M. (2000). Verb classification in Mayrinax Atayal. *Oceanic Linguistics*, 39:364–390.
- Kaufman, D. (2009). On the Proto-Austronesian morphemes \*pa-, \*<R> and \*<ŋ>. Paper presented at ICAL XI, Aussois. Available at: <https://www.academia.edu/1919923>.
- Kaufman, D. (2012). Pan \*ka- and predicate classes. Paper presented at AFLA 19, Academia Sinica, Taiwan.
- Kaufman, D. (2017). Lexical category and alignment in Austronesian. In Travis, L., Coon, J., and Massam, D., editors, *Oxford Handbook of Ergativity*. Oxford University Press, Oxford.
- Liao, H.-C. (2004). *Transitivity and Ergativity in Formosan and Philippine Languages*. PhD thesis, Department of Linguistics, University of Hawai'i, Mānoa.
- Maclachlan, A. E. (1989). The morphosyntax of tagalog verbs: The inflectional system and its interaction with derivational morphology. *McGill Working Papers in Linguistics*, 6:65–84.
- Pittman, R. (1966). Tagalog -um- and mag-: an interim report. *Linguistic Circle of Canberra Publications*, A(8):9–20.
- Ramos, T. V. (1974). *The Case System of Tagalog Verbs*. Number 27 in Pacific Linguistics B. Dept. of Linguistics, Australian National University, Canberra.
- Randriamasimanana, C. (1986). *The Causatives of Malagasy*. Oceanic Linguistics Special Publication 21. University of Hawai'i Press, Manoa.

- Ross, K. (1993). Causatives in Tagalog. McGill University.
- Ross, M. (1995). Proto-Austronesian verbal morphology: Evidence from Taiwan. In Li, P. J.-k., Tsang, C.-w., Ho, D.-a., Huang, Y.-k., and Tseng, C.-y., editors, *Austronesian Studies relating to Taiwan*, number 3 in Symposium Series of the Institute of History and Philology. Academia Sinica.
- Ross, M. (2002). The history and transitivity of western Austronesian voice and voice marking. In Wouk, F. and Ross, M., editors, *The History and Typology of Western Austronesian Voice Systems*, pages 17–63. Pacific Linguistics, Research School of Pacific Asian Studies, ANU, Canberra.
- Ross, M. (2015). Reconstructing Proto-Austronesian verb classes. *Language and Linguistics*, 13(3):279–315.
- Schachter, P. and Otnes, F. T. (1982). *Tagalog Reference Grammar*. University of California Press.
- Travis, L. (2000). Event structure in syntax. In Tenny, C. and Pustejovsky, J., editors, *Events as Grammatical Objects: the Converging Perspectives of Lexical Semantics and Syntax*, pages 145–185. CSLI Publications.
- Travis, L. d. (2010). *Inner Aspect: The Articulation of VP*, volume 80 of *Studies in Natural Language and Linguistic Theory*. Springer, Dordrecht.
- Van den Berg, R. (2004). Some notes on the origin of Malay *di-*. *Bijdragen tot de Taal-, Land- en Volkenkunde*, 160(4):532–554.
- Wolff, J. U. (1973). Verbal inflection in Proto-Austronesian. In Gonzalez, A., editor, *Parangal kay Cecilio Lopez*, number 4 in Philippine Journal of Linguistics Special Monograph, pages 71–91. Linguistic Society of the Philippines, Quezon City.
- Zeitoun, E. (2002). Reciprocals in the Formosan languages: a preliminary study. Paper presented at ICAL 9, Canberra, January 8-11.
- Zeitoun, E. and Huang, L. M. (2000). Concerning *ka-*, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics*, 39:391–414.