Daniel Kaufman (dkaufman@qc.cuny.edu)  ICAL 16, Manila, Philippines
Queens College, CUNY & Endangered Language Alliance  June 22, 2024

# Resegmentation and the fecundity of Philippine morphology
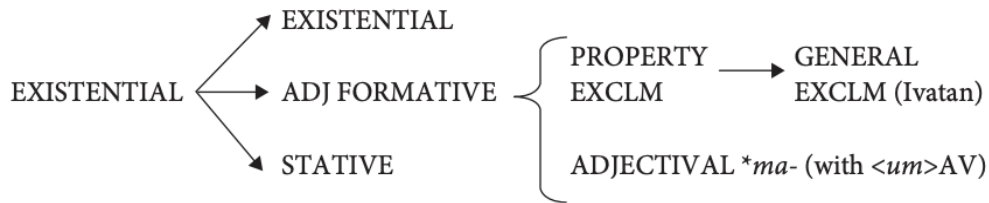
## 1.0  Multifunctionality

- One of the more salient features of Philippine morphology is the massive multifunctionality of common morphemes such as those descending from  PMP *pa-, *ka-, *paR-, *-an, etc.

- In some sense, Philippine languages are the most environmentally friendly of all thanks to their intense recycling (of morphs). Were these combinations compositional, this would not be surprising but in most cases, the relation between a complex morpheme and its component parts is opaque. To take a random Tagalog example, *pa-* CAUSATIVE + *-an* LOCATIVE VOICE yields the meaning of 'competition'.

| (1)a. | pa-ganda-han | b. | pa-lakas-an | c. | pa-taas-an ng  íhì |
|---|---|---|---|---|---|
| | CAU-beauty-LV | | CAU-strength-LV | | CAU-high-LV GEN urine |
| | 'beauty contest' | | 'contest of strength/sports' | | 'pissing contest' |

- Many of these seemingly opaque meanings can be traced back to their component parts through a series of historical reanalyses, although work in this area has hardly begun as "minor morphology" has long been ignored in favor of focusing on the nature of the Austronesian voice system.

- Here, I attempt to shed light on some pathways by which these morphemes expand their functions: (i) resegmentation, by which an affix associated with a stem is reinterpreted as part of a complex affix combination independent of the stem, and (ii) what I term "mood contamination", a process by which an alternation originally based on mood or finiteness is reinterpreted as part of an affix's meaning.

- These processes conspire to create gestalt formations whose meaning is more than the sum of its parts.

## 1.1  PAN *ka-: from existential to stative and beyond

- As argued elsewhere (Kaufman 2011, 2012), there is good evidence that the core function of PAN *ka- was existential (glossed here as HAVE), a meaning that is still reflected in many languages, e.g. Tagalog *mag-ka-bahay* 'to have a house'.

- From this core meaning, it takes on adjectival and stative functions via common pathways of change, ultimately defining a class of predicates.

- Most of the case studies discussed here relate to ways in which *ka- was reanalyzed as belonging to a complex prefix rather than taking the root as a direct complement. This resegmentation gives rise to broad multifunctionality of this prefix and to further degrammaticalization.

## 2.0 Resegmentation, mood contamination and degrammaticalization: four case studies

## 2.1 Exclamative *ka-*

- A class of stative roots in PMP (and PAN) appear to have required the *ka- prefix to form basic verbs (Ross 2015).

    - inherently involuntary roots like PMP *tiduR 'sleep' (i.e. *ka-tiduR 'to sleep')

    - property denoting roots like PMP *paqit 'bitter' (i.e. *ka-paqit 'to be bitter')

- These could then be affixed with the actor voice to form canonical intransitive predicates (i.e. *ma-tiduR, *ma-paqit), in addition to other voices.

- In certain environments (negative, imperative, *inter alia*), reflexes of *<um> disappear in the actor voice, a pattern which can be reconstructed to PAN or thereabouts. This leads to alternations still commonly found in MP languages such as:

Cebuano
(2)a.  masilawon                      b.      kasilaw!
        k<um>a-silaw-en                        ka-silaw
        <AV>HAVE-bright-PV                     HAVE-bright
        'bright'                               'how bright!'

Muna (van den Berg 1989: 173)
(3)a.  no-pana lalo-ku                      b. ka-pana-no       lalo-ku!
        3s.RL-hot heart-1s.GEN                 HAVE-hot-3s.GEN  heart-1s.GEN
        'I am angry.' (Lit. 'My heart is angry')    'How angry I am!'

- The null exponence of actor voice *<um> in exclamatives reveals the hidden *ka- that was there all along.

- Due to the surface alternation between *ma- (declarative) and *ka- (exclamative), the stage is set for **mood contamination**: reanalysis of ka- as an exclamative marker itself.

- Once this reanalysis takes place, *ka- can appear in new environments, which are neither stative nor adjectival, for instance:

Ivatan (Hidalgo & Hidalgo 1971: 148)

(4)a.  machi-nanaw sa        du  kayskuyan          b.      su  ka-pachi-nanaw  da!
    AV.SOC-study   3p.NOM OBL school                      ACC  EXCLM-SOC-study   3p.GEN
    'They study in school.'                              'How hard they study!'

## 2.2    Voxogenesis and ika- the voice of reason

- Psychological predicate roots select *ka- and form an intransitive clause with the actor voice, as in (5). (The combination of *ka- and *<um> is idiosyncratically truncated to *ma- from its earliest reconstructible stage.)

(5)a.  *malungkot*                    b.      *magálit*                    c.      *matuwà*
    k<um>a-lungkot                          k<um>a-gálit                          k<um>a-*tuwà*
    <AV>HAVE-sadness                        <AV>HAVE-anger                        <AV>HAVE-joy
    'X to be sad'                           'X to be angry'                       'X to be happy'

-  Such predicates may also select the theme of the psychological predicate using a reflex of *Si- the circumstantial voice, as shown in (6b). As seen in (6a), at no stage were ka- predicates allowed to co-occur with the patient voice *-en.

(6)a.  **kalungkutin*                   b.      *ikalungkot*
    **[**ka-lungkot**]-en                   Si-[**ka-lungkot**]
    <AV>HAVE-sadness                        CV-HAVE-sadness
    'X to be sad'                           'to be sad about X'

- Crucially, in formations such as (6b), *the theme of a psychological predicate is identical with the cause or reason for the state*.

(7)a.  Na-lungkot=siya                      dahil      namatay              ang    kanyang aso.
    k<um><in>a-lungkot=siya             dahil      k<um><in>a-matay     ang    kanyang aso
    <AV><PRF>have-sadness=3s.NOM because  <AV><PRF>have-die   NOM   3s.OBLaso
    'S/he was sad because his/her dog died.'

  b.  i-k<in>a-lungkot=niya               ang    pagkamatay ng    kanyang      aso.
    CV-<PRF>have-sadness=3s.GEN          NOM  death        GEN  3s.OBL        dog

- Because of the semantic simplicity of 'reason' and its sufficiently distinct interpretation from ordinary themes, ika- is ripe for resegmentation and reanalysis as the 'voice of reason'

- The entire development can be schematized as in (8).
  - The psych predicate source construction, exemplified in (8a), simply consists of the CV prefix attaching to a complex *ka- stem.
  - This is reanalyzed in (8b), where *Si-ka- is now resegmented as a complex prefix identifying a 'reason'.
  - The reanalysis is actuated in formations such as (8c), with roots that do not belong to the *ka- class.

|  | SOURCE | | REANALYSIS | | ACTUATION |
|---|---|---|---|---|---|
| (8)a. | *ikalungkot* | b. | *ikalungkot* | c. | *ikapunta* |
|  | Si-**[ka-lungkot]** | | Sika-**[lungkot]** | | Sika-**[punta]** |
|  | CV-HAVE-sadness | | CV.REAS-sadness | | CV.REAS-go |
|  | 'to be sad about X' | | 'to be sad because X' | | 'to go because of X' |

(9)  *I-k<in>a-sigaw=niya        sa   sakit ang  pag-tulak  nang  ka-tabi=niya      sakaniya.*
  CV-\<PFV\>REAS-shout=3S.GEN  OBL pain  NOM NMLZ-push GEN   RECP-side=3S.GEN 3S.OBL
  'He shouted from pain because of his seatmate's pushing him' (Klimenko & Endriga 2016:483)

- As noted by Chen, Gallego, Kuo et al. (this conference), not all Philippine languages employ *Si-ka- in this way. As seen in (5), Cebuano employs the bare CV form where Tagalog typically uses ika-.

(10)  <u>19th c. Cebuano</u>
  *Ang pagcamatay ni nanay maoy igahilac mo.*
  ang  pag-ka-matay  ni=nánay      mao=y         i-ga-hilak=mo
  NOM GER-STA-die       GEN=mother  AFF=INDEF      CV-IMPRF-cry=2s.GEN
  'You will cry over the death of your mother.' (Guillén 1898:69)

- Kankanaey, among other Cordilleran languages, appears to have gone in the other direction and removed *ka-* in the conveyance voice of original ka- stems:

|  | <u>Tagalog</u> |  | <u>Kankanaey</u> |
|---|---|---|---|
| (11)a. | ang  i-k<in>a-tabà=niya | (12)a. | han i-n-lames=na |
|  | NOM CV-\<PRF\>have-fat=3s.GEN |  | NOM CV-PRF-fat=3s.GEN |
|  | 'his/her reason/cause for gaining weight' |  | "                                " |
| b. | ang   i-k<in>a-matay=niya | b. | han i-n-katey=na |
|  | NOM  CV-\<PRF\>have-die=3s.GEN |  | NOM CV-PRF-die=3s.GEN |
|  | 'his/her cause of death' |  | "                                " |

## 2.3  *Magpaka-* the stative reflexive

- In constructions such as (2), we find a stative *ka- root combining with causative *pa-, combining with *pag-*, which contains the causative *pa- and a middle voice *R- (Kaufman 2009), followed by the actor voice *<um>.

(13)  <u>Tagalog</u>
*magpakamatay*
p<um>a-R-pa-[**ka-matay**]
<AV>CAU-MID-CAU-HAVE-dead
'to cause self to have death'
'to kill oneself'

(14)  <u>Ilokano</u>
*ag-paka-lukmeg*
pa-R-pa-[**ka-lukmeg**]
CAUS-MID-CAUS-HAVE-fat
'to cause self to have fat'
'to fatten oneself'

- Where does the reflexive meaning come from? Originally, it could only come from the middle voice *-R-, as seen in below:

(15)a. *umahit*
<um>ahit
<AV>shave
'to shave s.o.'

  b. *mag-ahit*
p<um>a-R-ahit
<AV>CAU-MID-shave
'to shave oneself'

<u>Ivatan</u> (PAn *R > y)  (Hidalgo & Hidalgo 1971, Kaufman 2009)

(16)a. *mandiman*
p<um>a<ŋ>-diman
<AV>CAUS<PL>-kill
'to kill s.o.'

  b. *maydiman*
p<um>a<R>-diman
<AV>CAUS<MID>-kill
'to kill each other'

(17)a. *manweswes*
p<um>a<ŋ>-weswes
<AV>CAUS<PL>-turn
'to turn s.o.'

  b. *mayweswes*
p<um>a<R>-weswes
<AV>CAUS<MID>turn
'to turn oneself'

- But because of *ka-*'s opacity and the fact that *paka-* does not show up frequently elsewhere in Tagalog, the stage has been set for reanalyzing *magpaka-* as a reflexive inchoative.

| SOURCE | REANALYSIS | ACTUATION |
|---|---|---|
| (18)a. *magpakamatay* | b. *magpakamatay* | c. *magpakatao* |
| p<um>a-R-pa-[**ka-matay**] | p<um>a-R-paka-**[matay]** | p<um>a-R-paka-**[tao]** |
| <AV>CAU-MID-CAU-HAVE-die | <AV>CAU-MID-REFL.INCH-die | <AV>CAU-MID-REFL.INCH-person |
| 'X to cause oneself to die' | 'X to cause oneself to die' | 'X cause oneself to be a person/humane' |

## 2.4 *pinaka-* the superlative

- Property denoting roots in Philippine languages form large paradigms with various adjectival, causative, intensive, moderative and superlative functions, among others, e.g.

(19)a. **ganda**  b. ma**ganda**  c. nápaka**ganda**  d. pinakama**ganda**
  'beauty'     'beautiful'     'very beautiful'   'most beautiful'

- Here we are interested in *pinaka-* the Tagalog superlative, which shows evidence of the following development, despite the first two stages not being attested historically:

|  | SOURCE | REANALYSIS | ACTUATION |
|---|---|---|---|
| (20)a. | ?*pinakaganda* | b. ?*pinakaganda* | c. *pinakamaganda* |
|  | p<in>a-[**ka-ganda**]-∅ | pinaka-[**ganda**] | pinaka-[**ma-ganda**] |
|  | <PRF>CAU-HAVE-beauty-PV | SUPER-beauty | SUPER-ADJ-beauty |
|  | ('made beautiful') | ('most beautiful') | 'most beautiful' |

- In stage (20a), a stative root taking *ka-, combines with a perfective patient voice causative to take on a non-compositional, intensive meaning.

- Because of the semantic simplicity of the superlative meaning, it is easily resegmented, as in (20b). At this point, the stage is set to attach to full stems, as in (20c), because *ka- is no longer treated as a root class prefix.

- A strikingly similar phenomenon is seen in the causatives of Pazeh, a Formosan language, which shows typical causative alternations, as in (21).

Pazeh (Li and Tsuchida, 2002; Wolff, 2009)
(21)a. ini  ma-ngesel  aku     b.  pa-ka-ngesel-i
   NEG  <AV>HAVE-fear 1S.NOM      CAU-HAVE-fear-IMP
   'I am not afraid.'           'Cause (him) to be afraid.'

- But as noted by Chen (2016:517), Zeitoun and Huang (2000:394) and Blust (2003:464), Pazeh attaches *paka-* to words that already have *ma-*, suggesting *paka-* is now monomorphemic.

Pazeh (Blust 1999:348)
(22)a. *ma-busuk*      b.  *paka-ma-busuk*
   k<um>a-busuk       paka-k<um>a-busuk
   <AV>HAVE-busuk      CAU.STA-<AV>HAVE-busuk
   'drunk'          'to make someone drunk'

- In other languages, similar prefixes are similarly reanalyzed and applied far beyond the usual *ka- stems seen above. In Polangui Bikol, the intensifier *laʔka-* now attaches to any stem:

West Albay Bikol (Polangui)
(23)a. *ma-pagal=siya*        b.  *laʔ-ka-pagal=niya*
k\<um>a-pagal=siya           laʔ-ka-pagal=niya
\<AV>HAVE-tired=3s.GEN     INTNS-HAVE-tired=3s.GEN
'for him/her to be tired.'      'How tired s/he is.'

(24)a. *abusero*             b.  *laʔka abusero*
abusive                INTNS  abusive
'abusive'               'very abusive'

- The looser attachment of the prefix can then lead to further **debonding** (Norde 2009). In the case of Pazeh, Chen argues that contact with Chinese languages led to *paka* becoming an independent word, as seen in (26b).

Kaxabu Pazeh (Chen 2016:516)
(26)a. *mu-payak    laladan.*     b.  *Adunu paka laladan mu-payak*
AV-wet       desk         Adunu CAUS desk    AV-wet
'The desk is wet.'        'Adunu made the desk wet.'

- In the case of Tagalog, we now find *pinaka-* attaching to fully inflected words, as in (25), and even to negation, as in (26). While these may be calques from English, *pinaka-* was already primed for degrammaticalization by earlier developments.

(25)   *ang  pinaka-nakaka-boring na linggo EVER.*
      NOM  SUPER-STA.CAUS-boring LNK week ever
      'the most boring week ever'

(25)   *pinaka-hindi=mo    ma-li~limut-ang         karanasan*
      SUPER-NEG=2s.GEN   POT-IMPRF~forget-LV:LNK   experience
      'your most unforgettable experience'

## 3.0  Conclusion

- We began by looking for reasons behind morphological recycling. One motivation is the reanalysis of simple stem forming prefixes like *ka-* as parts of larger prefixes. This was especially likely in conservative PMP languages because of the opacity of certain alternations (e.g. "disappearing" or null exponence *ka-, *-en, *\<in>, *\<um>).

- We also saw that resegmenting often led to more promiscuous attachment of morphemes and in some cases to total debonding and degrammaticalization into independent words.

- This is an extremely rich area for further exploration, since each Philippine language has numerous unique morphological formations based on familiar parts. What is especially intriguing are recurring patterns of degrammaticalization that have largely gone unnoticed and which can contribute much to the relatively poor inventory of clearly attested degrammaticalization phenomena (Norde 2009).

- Hopefully, this can spur new scholars to look at Philippine morphologies from a historical lens.

## References

Blust, Robert. 2003. 'Three notes on early Austronesian morphology.' *Oceanic Linguistics* 42(2):438--478.

Chen, Victoria. 2016. When Synthetic Meets Analytic: A Note on Structural Borrowing in Kaxabu Pazeh. *Oceanic Linguistics* 55(2), 669-677.

Guillén, Félix. 1898. *Gramática bisaya para facilitar el estudio del dialecto bisaya cebuano*. Malabón: Estab. Tipo-Lit. del Asilo de Huérfanos.

Hidalgo, Cesar and Araceli Hidalgo. 1971. *A tagmemic grammar of Ivatan*. Manila: Linguistic Society of the Philippines.

Kaufman, Daniel. 2011. 'Exclamatives and temporal nominalizations in Austronesian' In: Foong Ha Yap and Janick Wrona (eds.), *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, pp.721-754. Amsterdam: John Benjamins.

Kaufman, Daniel. 2018. "Inner and outer causatives in Austronesian: A diachronic perspective." In Ileana Paul, Laura Kalin and Jozina Vander Klok (eds), Festschrift for Lisa Travis. McGill Working Papers in Linguistics. pp. 204-218

Kaufman, Daniel. 2012. Predicate classes and Proto-Austronesian *ka-. Austronesian Formal Linguistics Association 19 (AFLA 19), Academia Sinica, Taiwan.

Klimenko, Sergei B. and Divine Angeli P. Endriga. 2016. Semantic Verb Classes and Regularity of Voice Paradigms in Tagalog. *Oceanic Linguistics* 55(2):480-499

Li, Paul and Shigeru Tsuchida. 2001. *Pazeh Dictionary*. Institute of Linguistics, Academica Sinica,

Norde, Muriel. 2009. *Degrammaticalization*. Oxford: OUP

Palmer, Gary B. Russel S. Rader and Art D. Clarito. 2009. The metonymic basis of a 'semantic partial': Tagalog lexical constructions with ka-. In Klaus Uwe Panther, Linda L. Thornburg and Antonio Barcelona (eds.), *Metonymy and Metaphor in Grammar*, pp.111-146. Amsterdam: John Benjamins.

Ross, Malcolm. 2015. Reconstructing Proto-Austronesian verb classes. *Language and Linguistics* 13(3):279–315.

Wolff, John U. 2009. Reconstructing PAN morphology by analyzing commonalities between Pazih and Tagalic languages. Paper presented at ICAL XI, Aussois, 2009

Zeitoun, Elizabeth and Huang, Lillian M. 2000. Concerning ka-, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics* 39: 391-414.