Two Types of States: A Cross-linguistic Study of Change-of-State Verb Roots

John Beavers*, Michael Everdell*, Kyle Jerro*, Henri Kauhanen†, Andrew Koontz- Garboden†,
Elise LeBovidge*, and Stephen Nichols†
* = The University of Texas at Austin, † = The University of Manchester
Corresponding author email: jtbeavers@utexas.edu

1 Introduction

- It is typically assumed that verb meanings consist of an "event structure" constraining the events described by the verb, consisting of (a) a template built from basic event-denoting predicates (e.g. via functional *v* heads; **?**) and (b) idiosyncratic roots filling in real world meanings (e.g. manner, state; **?**):

  (1)   a.   John flattened the rug $\approx$ [$_{vP}$ John [$_{v'}$ v_{cause} [$_{vP}$ the rug [-$en_{v_{\text{become}}}$ $\sqrt{flat}$]]]]
        b.   Kim cracked a vase $\approx$ [$_{vP}$ Kim [$_{v'}$ v_{cause} [$_{vP}$ a vase [$_{v'}$ v_{become} $\sqrt{cracked}$]]]]

- The template defines the verb's lexical aspectual properties, argument structure, and regular derivational morphology; the root just determines the verb's idiosyncratic morphology.

- An underexplored question (though see **???**, a.o., for related discussion) is whether there is a clean divide between meanings encoded by roots and by templates, e.g. are CAUSE and BECOME *only* introduced templatically?

- The null hypothesis is that if templates determine grammatical behavior semantically, such meanings should be excluded from roots, Embick's (**?**) "Bifurcation Thesis for Roots" (BTR) and Arad's (**?**) "Root Hypothesis" (also **??**).

- If this is true, then morphologically all change-of-state verbs should have the same templatic structures (e.g. that introduce an entailment of change) and when the roots of change-of-state verbs are used in templates lacking entailments of change no such reading should arise.

- We present a cross-linguistic study into this question. We show that across languages certain classes of roots are always associated with an entailment of change regardless of the template, and furthermore show morphosyntactic properties indicating that they are derived in distinct ways not predicted by the BTR.

- We argue that some roots entail templatic meaning, which furthermore has grammatical effects, ultimately arguing against the BTR (**?** in prep).

2 Change-of-State Roots

- As discussed, usually a change-of-state verb is built around a state denoting root. Most accounts are inexplicit as to the precise meaning of the root; the BTR entails, however, that it is purely stative, with no templatic meaning packaged into it.

- While this may be correct for some change-of-state verbs, it is not for others (cf. **??**). We distinguish the roots of **?**'s (**?**: 245) deadjectival change-of-state verbs in (**??**) (the basic Property Concepts of **?**) from those of her various non-deadjectival change-of-state verbs in (**??**) (for full lists see the appendix).

- Henceforth, we call these **property concept roots** and **result roots**.

  (2) **Property Concept**

  a. *Dimension*: large/big/enlarge, small/shrink/shrunken, short/shorten, ...

  b. *Age*: old/aged/age

  c. *Value*: bad/worsen/worse, good/improve/improved

  d. *Color*: white/whiten, black/blacken, red/redden, green/make green, ...

  e. *Physical Property*: cool/cool, dirty/dirty, dry/dry, wet/wetted, ...

  f. *Speed*: fast/speed up, slow/slow down

  g. *Human Propensity*: angry/anger, calm/calmed, sick/sicken, ...

  (3) **Result Roots**

  a. *Entity-specific Change of State*: burned/burn, melted/melt, frozen/freeze, ...

  b. *Cooking Verbs*: cooked/cooked, baked/bake, fried/fry, roasted/roast, ...

  c. *Breaking Verbs*: broken/break, cracked/crack, crushed/crush, shattered/shatter, ...

  d. *Bending Verbs*: bent/bend, folded/fold, wrinkled/wrinkle, creased/crease

  e. *Verbs of Killing*: dead/killed/kill, murdered/murder, drowned/drown

  f. *Destroying Verbs*: destroyed/destroy (ruined/ruin)

  g. *Verbs of Calibratable Change of State*: go up, go down, ...

  h. *Verbs of Inherently Directed Motion*: come/came, gone/go, go in, go out, ...

- There are morphological and semantic arguments that while PC change-of-state verbs are built on state-denoting roots lacking templatic entailments, result root verbs are not, contra the BTR and assumptions in the literature (e.g., **??**).

  > At least some roots encode at least some kinds of templatic meaning.

3 Morphological observations

- The morphological predictions of the BTR are:

  (4) a. Statives based on these two root types are identical in morphological complexity.

  b. Intransitive change-of-state verbs based on these two root types are identical in morphological complexity.

  c. Transitive change-of-state verbs based on these two root types are identical in morphological complexity.

- Here, we deal just with (**??**a) (see **??** for observations about (**??**b,c)). Two adjectival structures—basic states and result states (the latter deverbal).

  (5) a. Basic states (cp. **?**: 363): $[_{\text{AspP}} \text{ Asp } \sqrt{\text{Root}}\,]$

  b. Result states (cp. **?**: 367): $[_{\text{AspP}} \text{ Asp}_R [_{v\text{P}} \text{ DP } v_{\text{BECOME}} \sqrt{\text{Root}}\,]\,]$

- The contrast is transparently seen in English with adjective/deverbal adjectives derived from PC verbs, which show the verbalizing morphology –*en*.

  (6) a. Look at the bright picture on your left. (=camera took a bright picture)

  b. Look at the brightened picture on your left. (=camera took a bad picture, brightened with e.g. software)

- By contrast, with result root verbs, there is a single morphological form. If one assumes the BTR, and that any stative root can appear in either of the structures, this single form must realize both structures (**?**: 358).

  (7) broken, chipped, cracked, crashed, crushed, fractured, ripped, shattered, smashed, snapped, splintered, split, torn, baked, barbecued, blanched, boiled, braised, ...

- Embick claims that -*ed/en* realizes Asp and Asp_R with *break*-type roots, but only Asp_R for *red*-type roots, with Asp null. This is considered an accident of English morphology.

- Some data published in the literature though suggests that this generalization might occur in other languages as well, though the data are gappy (on Eastern Armenian see **?**: 98, on Ulwa see **??**, on Tongan see **?**, on O'odham, **?**: 92, on Pima see **?**: 3).

 | Adjectives derived from PC and result roots are not morphologically uniform, contra the BTR. |
  | --- |

4 The lexical semantics of the two kinds of roots

- The BTR gives rise to specific predictions in relation to the derivatives of roots of change of state verbs:

  (8) a. Simple adjectives (e.g., *red*) will not entail a prior change.

  b. Adjectives from deadjectival verbs (e.g., *reddened*) will entail a prior change.

  c. *boiled*, *split*, *cracked*, etc., since they realize both (**??**) and (**??**), will not entail prior change (since in any context, the adjective could be realizing (**??**)).

- The predictions in (**??**a,b) seem to be borne out:

  (9) a. The red dirt has never been reddened.

  b. The bright photo has never been brightened.

  (10) a. #The reddened dirt has never been reddened.

  b. #The brightened photo has never been brightened.

- The prediction in (**??**) is not — such adjectives entail a change of the kind named by the verb they are derivationally related to (**???**).

  (11) a. #The shattered vase has never been shattered.

  b. #The cooked chicken has never been cooked.

- NB: There are also the "derived statives" of **?**. However, recent work has shown these uses describe *atemporal* change along a spatial or even non-spatial scale and also that there are corresponding verbal uses as well (**???**), i.e., these uses are *semantically* deverbal (**?**).

- Now, this could just be a fluke of English. Maybe some roots for some reason always require a v_{BECOME} head for categorization, or something like that.

- But we also see a split in restitutive *again* modification, which has been analyzed as entailing modification of the root itself without any heads or decompositional operators.

  (12) Kim flattened the rug again.
    a. $[_{v\text{P}}$ John $[_{v'}\ v_{\text{cause}}\ [_{v\text{P}}$ the rug $[\ \text{-}en_{v_{\text{become}}}\ [\ \sqrt{flat}\ $ again $]\ ]\ ]\ ]\ ]$
    b. $[\ [_{v\text{P}}$ John $[_{v'}\ v_{\text{cause}}\ [_{v\text{P}}$ the rug $[\ \text{-}en_{v_{\text{become}}}\ \sqrt{flat}\ ]\ ]\ ]\ ]$ again $]$

- As **?**: 7 and **?**: 358 observes, *broken* roots do not allow pure restitutive readings:

  (13) a. John thawed the meat again (necessarily two 'defrostings')
        b. John melted the soup again (necessarily two 'defrostings')
        c. John drowned the zombie again (necessarily two 'drownings')

- This is very difficult to analyze on a decompositional approach.

  > Result roots and PC roots are not semantically uniform, contra the BTR

- The data above though are largely English-centric. Could it be that these are just quirks or exceptions that do not point to a general trend? We explore this next.

5 Cross-Linguistic Study
5.1 Morphology
- Our methodology was to take a balanced language sample and check for the existence of simple statives based on the root meanings above in those languages.

- We targeted the WALS 100 Language list (**?**), being a manageable balanced corpus. In cases where we lacked sufficient grammatical and dictionary resources, we substituted languages from the sister WALS 200 list, and also added in a few others based on available resources or native speakers/fieldworkers. The final total was 87 languages, mostly covering the areas and families of the WALS 100 (see the appendix for a list).

- We collected paradigms of simple state-inchoative-causative-result state (e.g. *red-redden-redden-reddenned*) for translations of the roots we examined in English, plus also any underlying root for languages in which such paradigms are based on a shared bound root.

- We privileged lexicalized terms over productively derived terms, and morphologically derived terms over periphrastically derived terms, motivated by a general tendency across languages for more lexical(ized) forms to be "normal" or "default" expressions (see e.g. the literature on causatives such as **???**).

- If a form was unattested but our grammatical resources made clear that there were productive processes for deriving it, we constructed a hypothetical form (but kept these marked so they could be left out of the analysis if need be). The ultimate preference rankings were:

  (14)   listed wholly lexicalized form > listed productive morphological form > hypothetical productive morphological form > listed periphrastic form > hypothetical periphrastic form > no data

- For simple states we looked for predicative forms, including those with a possessive predicational strategy (**?**).

- For consistency, we always took the dictionary translation seriously, e.g. if a morphologically simple stative term was translated as *broken* we actually listed it as a result state term.

- We have to date collected full data on 75 languages, and the main overall pattern holds: from a simple visual inspection it is clear that PC roots overwhelmingly tend to have simple stative forms (that usually serve as input to the rest) and result roots overwhelmingly tend not to (though there are certainly exceptions in both directions).

- A statistical analysis confirms this, though we had to make the following assumptions owing to the fact that a missing form could either be unattested or just not in our resources:

  - We ignored any root meanings for which we had no data in a given language; this we assumed was not having the data.
  - Otherwise, the methodology was to look at each remaining root and calculate for it the percentage of languages for which the simple state was attested.

- We then checked to see if the mean percentages of the PC and result roots (and the deviation from the mean) were significantly different. The results were striking and statistically significant (on a Mann Whitney U-test on the proportion of simple statives for PC and result roots; $p < 0.001$) (the results did not change if we threw out hypotheticals):

proportion filled simple states across language sample
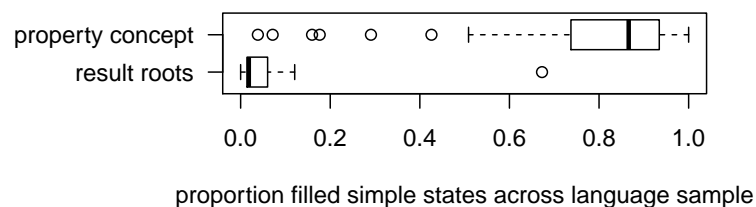
- We also did a similar comparison across root subclasses (and the difference was again significant on a one-way ANOVA across all subclasses; $p < 0.001$):

proportion filled simple states across language sample
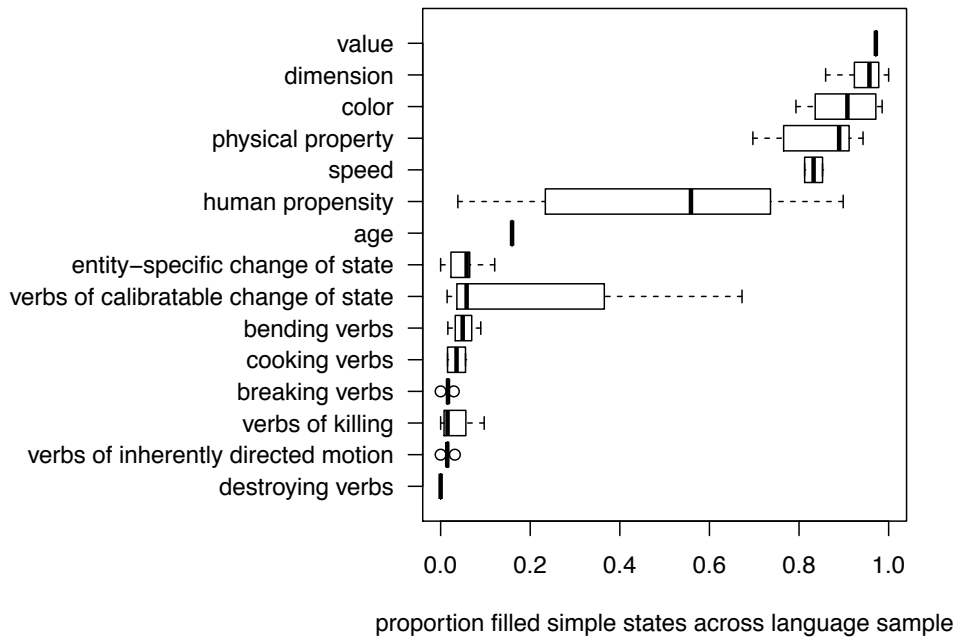
- Some noteworthy observations:

  – Age is the one PC root that patterns like a result root. But this also had one main member, *old*. In hindsight, this is not really a PC root semantically — one cannot be *old* without having been first younger. It's a rare adjective in English that entails change. This is not surprising: *old* is historically deverbal, from Old High German *aulden* "grow".

  – The other stand out is human propensity. But here there was a semantic category issue: for many of these verbs the simple state is derived (e.g. English *frightened*) but the verb it is derived from is itself stative. We made a decision to only look at eventive verbs, and so these forms were left out since they were unrelated to the rest of the paradigm.

- Despite this, the trend seems clear:

  (15) **The crosslinguistic morphological generalization**: State-denoting words based on result roots don't exist in the morphological form that PC roots do. The former tend to lack simple stative forms, the latter have them.

5.2 Semantics

- To test the semantic predictions cross-linguistically we choose opportunistic in-depth studies, using Kinyarwanda and Kakataibo as case studies (see **?** and **?**, respectively, both at this conference), to which we can also add the English data above.

- As a brief illustration, PC vs. result roots show the distinction regarding change entailments, and the latter resist restitutive modification, as in the following from Kakataibo:[1]

---

[1] 3=third person, A=subject of transitive verb, EMPH=emphatic, FACT=factitive, INDF=indefinite,

6

(16) a. *báinka   ani ˈikë   ˈaibika   uini   abi   ni*
        báin=ka=a   ani ˈikë   ˈai=bi=ka=a   uini   a=bi   ni
        hill=VAL=3A/S big be.3.IPFV then=EMPH=VAL=3A/S INDF.PRO 3=EMPH nor
        Diosabi   ni uni yubë   unibi   anioima.
        Diosabi   ni uni yubët uni=bi   ani-o-i-i=ma.
        God=EMPH nor man sorcerer man=EMPH big-FACT-IPFV=PROX=NEG
        'The hill is big, but nobody nor God nor a sorcerer made it big.'

   b. **ñu   nami tëakë   ˈikë   ˈaibika   uini*
        ñu   nami tëa-kë   ˈikë   ˈaibika   uini
        thing flesh cut=NFUT.NMLZ be.3.IPFV then=EMPH=VAL=3A/S INDF.PRO
        abi   tëakëma   ˈikë.
        abi   tëa-kë=ma   ˈikë.
        3=EMPH cut=NFUT.NMLZ=NEG be.3.IPFV
        'The meat is cut but nobody cut it.'

(17) a. [The desert starts off dry. Then, it is made non-dry. Then it turns dry again.]

        madin   papanka   ëdkitëkënia.
        madi=n   papa=n=ka=a   ëd-ki-tëkën-i-a.
        sand=POSS father=A/S=VAL=3A/S dry-INTR-ITR-IPFV=N.PROX

        'The desert is getting dry again.'

   b. *kana   ˈasa   ˈarutëkëa.*
        =ka=na   ˈasa   ˈa-ru-tëkë-a.
        =VAL=1A/S manioc do-UP-ITR=PFV
        I cook manioc again.   (requires two cooking events)

- **Upshot**: While PC roots behave as predicted by the BTR having a purely stative meaning, result roots always have an entailment of change, a fact that holds up across languages.

6 Conclusions and Consequences

- PC and result roots differ consistently across languages in terms of the inferences they generate and their morphological behavior. This clearly violates the BTR:

    - Under the BTR the roots of result verbs lack entailments of change and a form based on them would have to acquire one through addition of a v_{become} head.

    - This would predict that in contexts without it the change entailment will not arise.

    - Statives are such contexts, yet the entailment is there.

    - Furthermore, on the analysis consistent with the BTR *again* modifiers should have access to the root below v_{become}, but this also does not arise.

    - Morphologically, in at least *some* languages result roots should have simple stative forms, i.e. where the stativizer is null. This rarely happens, and clearly not to the degree it does with PC roots.

INTR=intransitive, IPFV=imperfective, ITR=iterative, NEG=negation, NFUT.NMLZ=non=future nominalizer,
POSS=possessive, PRO=pronoun, PROX=proximate, S=subject of intransitive verb, UP=up, VAL=validational

- So what would explain these data? A simple analysis would be to assume that roots of PC verbs and result root verbs differ in that the former describe simple states and the latter states for which it is also entailed that there exist a cause:

(18) a. $[\![\sqrt{\text{flat}}]\!] = \lambda x \lambda s [flat'(x,s)]$

b. $[\![\sqrt{\text{crack}}]\!] = \lambda x \lambda s [has.fissure'(x,s) \wedge \exists e'[become'(e',s)]]$

- Defining v_{become} as in (**??**a) and combining it with (**??**) derives the inchoatives in (**??**b,c).

(19) a. $[\![v_{become}]\!] = \lambda P \lambda x \lambda e \exists s [become'(e,s) \wedge P(x,s)]$

b. $[\![v_{become} \sqrt{\text{flat}}]\!] = \lambda x \lambda e \exists s [become'(e,s) \wedge flat'(x,s)]$

c. $[\![v_{become} \sqrt{\text{crack}}]\!] = \lambda x \lambda e \exists s [become'(e,s) \wedge has.fissure'(x,s) \wedge \exists e'[become'(e',s)]]$

- The change-of-state use of the PC root only entails change by virtue of v_{become}, but the root of the *crack* root has the entailment itself. This predicts that other uses of the root without v_{become} will give rise to the entailment, and that *again* modification, even if applying to the root alone, will still scope over an entailment of change.

- This does not capture the morphological asymmetry. However, default Spell-Out rules for Asp and *v* heads might explain this (overridden by root-specific rules as with *old*):

(20) Default Spell-Out for v_{become} for root $\sqrt{\text{R}}$:
  a. If $\sqrt{\text{R}}$ entails change, then $\emptyset$
  b. If $\sqrt{\text{R}}$ does not entail change, then *-en/ed*

(21) Default Spell-Out for Asp:
  a. If $\sqrt{\text{R}}$ does not entail change, then $\emptyset$
  b. If $\sqrt{\text{R}}$ entail change, then *-en/ed*

- This has a functional motivation: the two categories are unmarked for opposite meanings.

- However, these rules go against the grain of event structures: roots should only matter arbitrarily in regular morphology. Yet it seems clear that the pattern governing the overt realization must be contingent on root semantics, suggesting the overall correctness of this analysis, arguing against that aspect of event structural theories.

- In sum, the root encodes an entailment that can be elsewhere introduced templatically, and it can matter grammatically. This is inconsistent with the BTR.

7   Acknowledgments

A Roots and Languages

- Here is the full list of root meanings we used, including some of the synonyms we considered. These served as target terms for dictionary searches. The assumption was not that the translations would be perfect; rather, our goal was to target many root meanings within the same broader classes.

(22) **Property Concept**

a. *Dimension*: large/big/enlarge, small/shrink/shrunken, short/shorten, long/lengthen, deep/deepen, wide/widen, tall/height/heighten

b. *Age*: old/aged/age

c. *Value*: bad/worsen/worse, good/improve/improved

d. *Color*: white/whiten, black/blacken, red/redden, green/make green, blue/make blue, brown/make brown

e. *Physical Property*: cool/cool, cold/make cold, warm/warm, hot/heat up, dirty/dirty, dry/dry, wet/wetted, straight/straighten, hard/harden (tough/toughen), soft/soften, tight/tighten, clear/clear, clean/clean, smooth/smooth, sharp/sharpen, sweet/sweeten, weak/weaken, strong/strengthen

f. *Speed*: fast/speed up, slow/slow down

g. *Human Propensity*: angry/anger, calm/calmed, scared/scare (frightened/frighten), sick/sicken, sad/sadden (depressed/depress), hurt/hurt, tired/tire, embarrassed/embarrass, entertained/entertain, surprised/surprise, worried/worry, pleased/please

(23) **Result Roots**

a. *Entity-specific Change of State*: burned/burn, melted/melt, frozen/freeze, decayed/decay (rotten/rot), swollen/swell, grown/grow, bloomed/bloom (flowered/flower, blossomed/blossom), withered/wither (wilted/wilt), fermented/ferment, sprouted/sprout (germinated/germinate), rusted/rust, tarnished/tarnish

b. *Cooking Verbs*: cooked/cooked (baked/bake, fried/fry, roasted/roast, steamed/steam), boiled/boil

c. *Breaking Verbs*: broken/break, cracked/crack, crushed/crush, shattered/shatter, split/split, torn/tear (ripped/rip), snapped/snap

d. *Bending Verbs*: bent/bend, folded/fold, wrinkled/wrinkle, creased/crease

e. *Verbs of Killing*: dead/killed/kill, murdered/murder, drowned/drown

f. *Destroying Verbs*: destroyed/destroy (ruined/ruin)

g. *Verbs of Calibratable Change of State*: go up (raised/rise, ascended/ascend, increased/increase, gained/gain), go down (fallen/fall, dropped/drop, descended/descend, decreased/decrease, declined/decline)

h. *Verbs of Inherently Directed Motion*: come/came, gone/go, go in (entered/enter), go out (exited/exit), returned/return

- Here is the full of list languages we investigated:

(24) Abkhaz, Acholi, Acoma, Alamblak, Anejom, Amele, Arabic (Egyptian), Arapesh (Mountain), Bagirmi, Bariai, Basque, Berber (Middle Atlas), Burmese, Burushaski, Canela-Krahô, Carib, Chamorro, Chukchi, Cree (Plains), Daga, Dani (Lower Grand Valley), English, Fijian, Finnish, French, Georgian, German, Gooniyandi, Grebo, Greek (Modern), Guarani, Gújjolaay Eegimaa, Hausa, Hawaiian Hebrew (Modern), Hindi, Hmong Njua Hopi, Huitoto (Minica), Imonda, Indonesian, Jakaltek, Japanese, Kakataibo, Kamano, Kannada, Karok, Kayardild, Kewa, Khalkha (Mongolian), Khoekhoe, Kinyarwanda, Kiowa, Koasati Koiari, Korean, Koyraboro Senni, Krongo, Kutenai, Kwoma, Lakhota, Lezgian, Lavukaleve, Luvale, Makah, Malagasy, Mandarin, Mangarrayi, Mapudungun (Mapuche), Martuthunira, Maung, Maybrat Meithei, Mixtec (Chalcatongo), Mocoví, Murrinh-Patha Navajo, Ngiyambaa, Oksapmin, Oneida, Oromo (Harar), Otomi (Mezquital) Paiwan, Paumarí, Persian, Pintupi, Quechua Huallaga, Rama, Rotokas, Russian, Sango, Sanuma, Spanish, Supyire, Swahili, Tagalog, Tenango Tzeltal, Thai, Tiwi, Turkish, Vietnamese, Wari', Warao, Wichita, Yagua, Yaqui, Yoruba, Yup'ik, Zoque, Zulu