Deconstructing subcategorization: Conditions on insertion vs. position

Laura Kalin and Nicholas Rolle

Princeton and Leibniz-ZAS, Berlin

NELS 2020
Introduction

**Subcategorization** has been utilized for decades to account for idiosyncratic behaviors of individual exponents (morphs)

- A subcat frame expresses elements/structure required in an exponent’s environment, e.g., __V (“be before a vowel”)

(1) **Chamorro:** Actor Focus /um/: __V
   a. V-initial stem: um-epanglo ‘look for crabs’
   b. C-initial stem: tr<um>isti ‘become sad’

(2) **Tzeltal:** 3.POSS /y/: __V
   a. V-initial stem: y-ahwal ‘his ruler’
   b. C-initial stem: *m<y>ul ‘his sin’

- **The puzzle:** If both Ch. um and Tz. y have subcat frame __V, how do we account for their different behavior?

---

Proposal

The point of this flash talk: Exponent subcategorization must be bifurcated into two separate and ordered mechanisms

1. Condition on Insertion (COIN) \((\_\_V \text{ in Tzeltal})\)
   \(\approx\) Can the exponent combine with a given stem?
   \(\rightarrow\) For: suppletive allomorphy,\(^2\) morphological compatibility\(^3\)

2. Condition on Position (COP) \((\_\_V \text{ in Chamorro})\)
   \(\approx\) Where should an exponent be located in a string?
   (note: not for regulating basic linear position)
   \(\rightarrow\) For: unexpected constituency disruption (infixation, second positionhood, ‘special clitics’)\(^4\) and idiosyncratic prosodic effects (and/or rule blocking)\(^5\)

\(^3\)Lieber 1980; Jensen 1990; Inkelas 1990; Booij and Lieber 1993
\(^4\)Zwicky 1977; McCarthy and Prince 1993a,b; Yu 2003, 2007; Bye and Svenonius 2012; Kalin 2020a
\(^5\)Spring 1992; Downing 1998b,a; Chung 2003; Zec 2005; Bickel et al. 2007; Caballero 2010; Bennett et al. 2018; Rolle and Hyman 2019; Rolle and O’Hagan 2019; Tyler 2019
Illustration

Crucial data: An exponent may have a COIN and a distinct COP, and if so, then the COIN is satisfied before the COP

(3) **Nancowry** (Radhakrishnan, 1981; Kalin, 2020b)
   a. INSTNOM /an/: COIN: with monosyllabic stems  
      COP: after first C
   b. INSTNOM /in/: COIN: with disyllabic stems  
      COP: after first V

(4)  a. INSTNOM + *top* (‘to drink’)
      \[ \text{COIN} \Rightarrow \text{an} \]
      \[ \text{COP} \Rightarrow \text{t}<\text{an}>\text{op} (\text{‘a glass’}) \]
   b. INSTNOM + *kurus* (‘to scratch’)
      \[ \text{COIN} \Rightarrow \text{in} \]
      \[ \text{COP} \Rightarrow \text{ku}<\text{in}>\text{rus} \]
      \[ \text{COP} \Rightarrow \text{k}<\text{in}>\text{rus} (\text{‘a rake’}) \]
Take-away: Subcategorization is not a unified phenomenon, even at the fine-grained level considered here (exponents)

- Models employing subcategorization need **two** separate types of frames, one for **insertion** and the other for **position**

Architectural implication: There is a level of representation where **COINs** are evaluated, but **COPs** are not (yet)

- Supports models where morphology precedes phonology\(^6\)
- Parallel P-with-M models\(^7\) and dual-route approaches to phonological **COINs/COPs**\(^8\) make the wrong predictions
  - Kalin (In prep): **COINs** are *never* evaluated in an infixed (post-COP) position; **COINs** always precede **COPs**\(^9\)

---

\(^6\) Halle and Marantz 1993; Paster 2006; Bye 2008; Embick 2010; Bye and Svenonius 2012

\(^7\) McCarthy and Prince 1993a,b; Hyman and Inkelas 1997; Horwood 2002

\(^8\) Mester 1994; Kager 1996; Mascaró 1996, 2007; Tranel 1998; Bonet 2004; Yu 2017

\(^9\) Cases comparable to that in Nancowry: Hunzib verbal plural (van den Berg, 1995), Alabama middle voice (Hardy and Montler, 1991), and Sierra Miwok stem one formation (Bye and Svenonius, To appear).
Thank you!

Thank you to Sharon Inkelas, Mary Paster, several anonymous reviewers, and the attendees of Princeton’s POPCICLE research group for extremely helpful feedback on this project.
Appendix A: Implementation

Consequence: Theoretical models which employ subcategorization must be modified to encode two separate subframe types, one for insertion and the other for position.

(W. Macedonian NEG) (5) $\text{SUBCAT(ne)}: [\omega_{-MIN} \text{ne} [\ldots]]$

- Input: ne mu gi dava . . .
- a. $[\omega \text{ne mu 'GI dava} \ldots$
- b. . . .

(Bennett et al., 2018) $\downarrow \text{SC-COP} \text{ (not SC-COIN)}$

(Serbo-Croatian second position clitics) (Sande et al., 2020) (6) 

$[\text{PRES,3SG}] \longleftrightarrow \begin{cases} \text{Features: /je/} \\ \text{P(SC): } [\omega - X] \\ \text{Ranking: -} \end{cases} \rightarrow \text{SC-COP}$
Appendix B: Frame substance

Argument for differentiating two types of subcat frames: COINs and COPs have different typological profiles with respect to the substance of their subcat frames

COP frames contain... (not an exhaustive list!)

- **Phonological elements**: C, V
- **Prosodic elements**: Syllable, foot, stress, phonological word, phonological phrase

COIN frames contain... (not an exhaustive list!)

- **Phonological elements**: C, V, **specific segments, features**
- **Prosodic elements**: Syllable, foot, stress, phonological word, phonological phrase
- **Lexical elements**: Idiosyncratic (classes of) roots
Appendix C: Optimization

Could the distribution of INSTNOM exponents in Nancowry be analyzed as optimizing, i.e., without COINs/COPs?

No (Kalin, 2020b):

- There is no disyllabic size constraint in the language (minimal or maximal), i.e., no conspiracy for disyllabic).
  - E.g., Another infixal exponent, -am- (AGNOM), builds trisyllabic words from disyllabic stems/roots.
- There is no phonotactic/phonological motivation at all for choosing -in- over -an-.
  - -an- could combine with stems of all sizes (like -am-).
  - A post-first-consonant distribution for any VC infix in Nancowry is more optimal than post-first-vowel, as it avoids vowel hiatus and avoids adding a coda.
  - As noted by Paster (2006, 167-168), there is no reason the distribution of the two INSTNOM exponents shouldn’t be reversed.
Appendix D: One mechanism?

Alternative under Yu (2007, 229): Languages “respond to the failure to satisfy a phonological subcategorization requirement in different ways”

- “when morpheme interruption is prohibited”, an exponent must satisfy its frame in its default position (if it can’t, it is blocked from appearing → gap or allomorphy) = COIN
- when morpheme interruption is allowed, the exponent moves to its desired position (→ infixation) = COP

Our claim: It is not viable to maintain that subcategorization involves one mechanism with two different effects.

- COINs/COPs are not an either/or: A single exponent can have both a COIN and a COP.
- Whether morpheme interruption is allowed is not a language-wide property, but rather is specific to exponents.


Deconstructing subcategorization

Laura Kalin and Nicholas Rolle

Introduction
Proposal
Illustration
Discussion
Appendices
References


