

Always far from perfect, yet always good enough

Farrell Ackerman, Ágnes Kalivoda, Robert Malouf

UC San Diego, Hungarian Research Centre for Linguistics, San Diego State University

## Some issues for making sense of morphology

- Specific questions guided by an overarching inquiry
  - What is the nature and extent of cross-linguistic morphological variation, and why does knowing this matter?
  - Reasonable hypothesis: it extends our notion of what is a possible morphological construction and, thereby, defines the appropriate targets of explanation as well as the nature of possible explanations.

# Language systems in the wild

- Languages frequently depart from simple content/form mappings and consistent category clustering (Crysmann and Bonami 2016, Mansfield et. al. 2020, Manova 2022, Mansfield et. al., 2022, Saldana et. al., 2022), some quite dramatically, and this raises questions about the nature of the object we want to understand and model.
  - How are such complex systems organized?
  - What role does this organization play with respect to licensing inferences concerning the existence of one construction given another?
  - What other types of systemic organization, beyond implicational relations, do languages exhibit?
  - What relation do principles of organization have concerning the learnability of such complex (discriminable) systems? (Blevins 2016)
- These are the questions guiding research on inflection and derivation in paradigm based morphology (Blevins 2016, Parker and Sims 2016).

## Guiding generalizations about syntagmatics

- Cross-linguistic principle of syntagmatic organization
  - Language particular inventories of internal elements, segmental and suprasegmental, (recurrently) organized to discriminate constructions from one another, with classic morphemic organization being the simplest case, not a normative type. (Ackerman et. al. 2009, Blevins 2016)

# Guiding generalizations about paradigmatic organization

- Many ways morphological systems can be organized: (Wurzel 1987, Plank, 2011, Parker and Sims 2016 notion of 'work')
- Some systems are organized in terms of implicative organization reflected in (low) conditional entropy values:
  - inflectional systems (Baerman and Palancar 2014; Parker and Sims 2020, Wilmoth and Mansfield 2021, Beniamine, Bonami and Luis 2021)
  - derivational systems ( Bonami and Strnadova 2019; Hathout and Namer 2022)
  - interactions between inflection and derivational systems (Pellegrini and Bonami 2022)

## Guiding generalizations about paradigmatic organization

- Some systems are organized in terms of simple series of analogical inferences within families of inflectional and derivational constructions that operate cross-categorially. Conditional entropy measures are irrelevant, since alternative encodings are largely irrelevant for members of these families
- Access to network of inflectional and derivational constructions for both simple and complex verb constructions explains native speaker competence to generate unencountered forms of constructions:
- If you know one form of a lexeme, you know them all, and if you know one form of one lexeme you know desired forms for other lexemes.

# A complex systems view toward diachronic development of new construction types

- The reuse of old parts for new purposes leads to the development of new categories, constructions, systems of constructions (Bates 1979; Lass 1997) - *exaptation* - (Lass 1990, Traugott 2004; Norde and van de Velde 2016)
- Self-organizing principles in language systems yield construction types (Lass 1997; Dimmendaal 2006, 2011; Ackerman and Nikolaeva 2013) representing recurrent and convergent novelties across languages:
  - Possessive relative constructions (Ackerman and Nikolaeva 2013)
  - Coverb constructions (Dimmendaal 2006)
  - Mixed category constructions (Malouf, Nikolaeva and Spencer)
  - **Complex predicates with relational preverbs** (Nash 1982; Craig and Hale 1988; Booij ed. 2003; Los et. al.,2017; Craig and Imbert 2008; Arkadiev 2015; Weir 1986; Craig and Hale 1998; Kiefer and Honti 2003; Hieber 2018)

# The importance of complex systems

- Guiding analogy from ecological developmental biology: only way to understand the variety of attested phenotypes in an organism is to understand its development over time within a constellation of co-constructing systems.

*At the same time, the organism, as it develops and functions, alters its environment so as to condition the kinds of environmental cues that it and co-occurring organisms will receive at later ontogenetic stages or possibly in subsequent generations. **Development thus arises from a complex network of causal interactions in which organism and environment co-construct each other (Laland et al. 2013a, b, 2014) through reciprocal influences that effectively break the supposed barrier between the internal and the external (Sultan 2015, 2019)***

***Here we come face to face with organisms as integrated wholes, fundamentally not decomposable into independent and separately optimized parts. (Gould and Lewontin 1979:591)***



# The importance of complex systems

- There is no calculably optimal result in grammar, i.e., as in eco-devo, since we don't know all of the relevant interacting systems nor how they affect one another over time.
- Some identifiable morphospaces of possibilities as defined by trajectories of systemic interactions are identifiable, but emergent (and different) encodings are evidently good enough, often yielding dramatically different encodings of the 'same' construction.
- Suggests that variation is not principled deviation from uniformity, but represents contingently guided systemic strategies of grammatical organization, i.e., *bricolage* (Lass 1997; Turner 2007 Sims p.c.)
- Shows the developmental influence of language histories in synchronic constructions and the shaping forces of language specific systems. (See research on ecological evolutionary developmental systems Gilbert and Sarkar 2000; Sultan 2015; on application to language see Balari et al. 2020):
- In other words, the system has to work as a whole, with neither the parts nor the whole reflecting optimization.

## A recurrent domain for exploration

- Striking similarities in the exaptation profiles of complex predicates consisting of preverb + verb in the languages of the world.
- Grammaticalization of such complex predicates involve tmesis and univerbation with respect to preverbs. (Lehmann 2021, references)
- Adpositions, adverbs and incorporated elements are historical resources for *category genesis* (Hieber) into preverbs.

*...the material exapted is at the point of exaptation doing something else (which it may continue to do): but it still is capable of being remanufactured or restructured... as part of a different kind of coexisting structure. (Lass 1997:320)*

- Cross-linguistic parallels in a common exaptation pattern produces system specific encodings, characteristic for families (Souian) or subset of languages in families (Ugric).
- Following the observations of Maiden these constructions “persist”:

*My aim here is not to describe what might be seen as little more than uninteresting diachronic inertia but rather to identify areas where inertia might almost be described, paradoxically, as ‘dynamic’. I have in mind structures in which the form–meaning relationship was and remains opaque, yet ‘holds out’, through time, despite clear potential for resorting to a more transparent structure. (Maiden. 2011:156)*

## **Two dimensions of attested morphological variation: affix order (see Manova: 2022; Mansfield et. al., 2022) and surface exponence**

- 1. Common morphotactic claim concerning the relative order of derivational and inflectional affixes: Derivational affixes appear 'inside' or 'flanked by' inflectional affixes.  
*If both the derivation and inflection follow the root, or they both precede the root, the derivation is always between the root and the inflection. (Greenberg 1966:93)*
- Numerous (contested) cases where inflectional marking is flanked (either preceding or following) by derivational markers (Bochner 1983; Booij 1993; Rainer 1996; Rice 1985).
- 2. Claims that lexical representations (excepting phrasal idioms) are encoded by synthetic word forms with continuous lexical stems, challenged by periphrastic surface exponence and discontinuous lexical stems.

## Construction theoretic expectations

- Nothing precludes the interleaving of derivation and inflection
- Nothing precludes the multi-word exponence of derivational and inflectional information.
- These are all possible (and attested) ways of encoding information, largely determined by diachrony and systemic organization.
- Caution about overly confident claims concerning universal structure tempered by the *empiricist turn* (Chater et al. 2015)

## Cross-linguistic strategies of exponence

		Inflection relative to derivation	
		External	Internal
Surface exponence	Synthesis	Adyghe (Arkadiev 2015) sə-b-g <sup>w</sup> -e-wac <sup>w</sup> e 1SG.ABS.-2SG.IO.PV-PST-stand 'I stood with you'	Hocank (Helmbrecht & Lehmann 2009:23) ho-ra-gi-sa-rak PV-2SG-APP/BEN-2SG-root 'he sold me'
	Periphrasis	Meskwaki (Jones 1907 cited in LeSourd 2009 ) ke-pyêchi... nân-ene-pwa 2-PV go.get-1/22-22.IND 'I have to come get you'	Hungarian nek-em... rohan-t PV-1SG run-PAST.3SG.INDEF 'He attacked me'

- All logical options in morphospace attested, but the actual networks of constructions and characteristic properties of individual constructions vary.

## Explorations of similarity and difference

- Inflection internal to derivation is independent of surface exponence: occurs in both synthetic (Hocank) and periphrastic (Hungarian).
- (Potentially) discontinuous lexical stems in Siouan synthetic constructions (also see Mithun on Athapaskan and Carter on Ket) and Hungarian periphrastic constructions
- Exaptation produces cumulative systemic effects that are evidently good enough, even if they don't seem "natural".
  - The Siouan configuration is old and persists across the family: similarly in Athapaskan (Mithun) and Ket (Vajda, Carter).
  - The Hungarian pattern is common to Ugric, but most luxuriant and extensive in Hungarian.
  - How do particular patterns arise, why do they persist despite their deviations from the (perfect) alignments that seem natural and why do they keep expanding when this occurs?

# Hocank (Souian) synthetic verbal template – Rankin 2002; Helmbrecht and Lehmann 2009;

Kasak 2019; Marsault 2021

pron I		outer applicatives		outer instrumentals	pron II		benefactive applicative*	pron III A	inner instrumentals	verbal root	suffixes
		instrumental	locative		U	A					
<i>hj-</i> 1DLA/ 1PLA	<i>wa-</i> 3PL OBJ	<i>hi-</i> APPL. INST	<i>ha-</i> APPL. SUPESS	<i>boo-</i>	<i>hj-</i> 1E.U	<i>ha-</i> 1E.A	<i>gi-</i> APPL.BEN	second conjugation	<i>gi-</i>	X	<i>-kje</i> FUT
<i>wq'qga-</i> 1DI.U/ 1PI.U			<i>ho-</i> APPL. INESS	<i>nqq-</i>	<i>nj-</i> 2U	<i>ra-</i> 2A	<i>kii-</i> REFL		<i>ra-</i>		
				<i>mqq-</i>	<i>nji-</i> 1→2		<i>kiki-</i> RECP		<i>ru-</i>		
				<i>taa-</i>			<i>kara-/kV-</i> POSS. REFL		<i>wa-</i>		
<b>-7a</b>	<b>-7b</b>	<b>-6a</b>	<b>-6b</b>	<b>-5</b>	<b>-4a</b>	<b>-4b</b>	<b>-3</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>

\* /reflexive/reciprocal/possessive reflexive

# Challenges of Hocank (Souian) verbal template (Helmbrecht and Lehmann 2009)

pron I		outer applicatives		outer instrumentals	pron II		benefactive applicative*	pron III A	inner instrumentals	verbal root	suffixes
		instrumental	locative		U	A					
<i>hj-</i> 1DLA/ 1PLA	<i>wa-</i> 3PL. OBJ	<i>hi-</i> APPL. INST	<i>ha-</i> APPL. SUPESS	<i>boo-</i>	<i>hj-</i> 1E.U	<i>ha-</i> 1E.A	<i>gi-</i> APPL.BEN	second conjugation	<i>gi-</i>	X	<i>-kje</i> FUT
<i>wq̣q̣ga-</i> 1DI.U/ 1PI.U			<i>ho-</i> APPL. INESS	<i>nqq-</i>	<i>nj-</i> 2U	<i>ra-</i> 2A	<i>kii-</i> REFL		<i>ra-</i>		
				<i>mqq-</i>	<i>nji-</i> 1→2		<i>kiki-</i> RECP		<i>ru-</i>		
							<i>kara-/kV-</i> POSS. REFL				<i>wa-</i>
<b>-7a</b>	<b>-7b</b>	<b>-6a</b>	<b>-6b</b>	<b>-5</b>	<b>-4a</b>	<b>-4b</b>	<b>-3</b>	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>

\* /reflexive/reciprocal/possessive reflexive

- the ambivalent morphemic or submorphemic status of preverbs,
- the discontinuous nature of many roots,
- the existence of sizable paradigms of internal affixes,
- inflectional, derivational, and lexical material are interdigitated in a synchronically idiosyncratic, but diachronically explicable order,
- their syntagmatic interdependence and multiple exponence.
- Similar observations have been made about Athapaskan languages (Mithun, K. Rice) and Ket (Vajda, Carter).



## **Hungarian PV V constructions**

- Diachronic development from an early system with few (and marginal) constructions to a synchronic situation of enormous abundance and categorial variability of PVS.

# The perennial challenge of Hungarian Preverb Verb (PV V) constructions

PV	Gloss	Types	Tokens	
el	'away'	5 321	3 748 134	established preverbs: high type and token frequencies
meg	'perfective'	5 149	5 734 400	
le	'down'	4 528	1 096 673	
ki	'out'	4 418	2 521 026	
...				
hozzá	'towards'	525	347 442	debated preverb-status: medium type frequency, varied token frequencies
agyon	'over, to death'	438	11 207	
körül	'around'	410	23 180	
haza	'(to) home'	401	79 725	
...				
egyet	<i>egyet-ért</i> 'agree'	1	72 765	preverb-like lexical items: co-occurring with a single verb, varied token frequencies
utol	<i>utol-ér</i> 'catch up'	1	5 215	
cserben	<i>cserben-hagy</i> 'let sy down'	1	1 202	
nyilván	<i>nyilván-tart</i> 'keep track of st'	1	997	
...				

based on the Hungarian Gigaword Corpus (Oravecz et al. 2014)

## Subset of Pvs inflects for pronominal person/number argument governed by the PV V combination

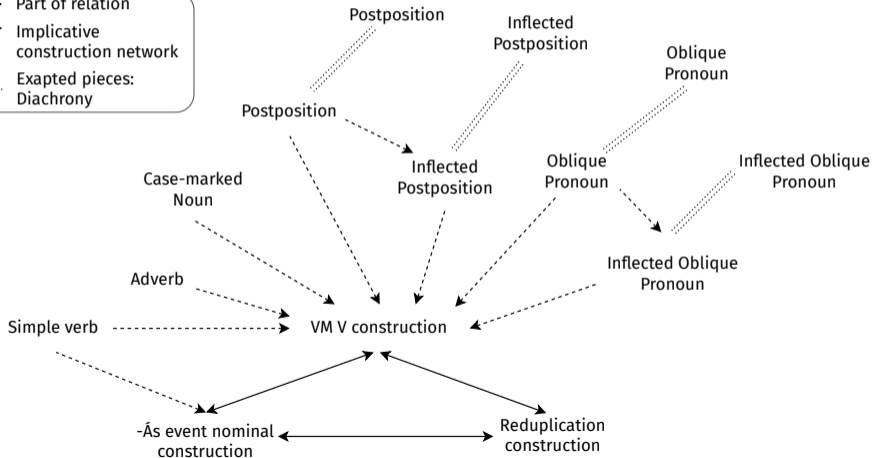
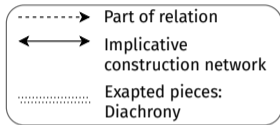
preverb	1SG	2SG	3SG UNM.	3SG M.	1PL	2PL	3PL	sum	gloss
rá	12,187	4,507	176,586	93	5,360	677	5,352	204,762	'onto'
hozzá	2,566	1,104	188,523	10	1,454	180	1,875	195,712	'to, toward'
bele	1,953	890	146,492	95	717	55	1,004	151,206	'into'
<b>neki</b>	<b>3,165</b>	<b>1,347</b>	<b>21,183</b>	—	<b>686</b>	<b>186</b>	<b>1,189</b>	<b>27,756</b>	'to, against'
rajta	481	401	2,065	—	421	63	368	3,799	'on'
alá	68	34	43,473	78	27	4	50	43,734	'under.to'
utána	719	223	8,416	—	304	29	852	10,543	'after'
mellé	742	162	2,785	368	350	20	398	4,825	'beside.to'
elé	974	223	144	302	1,304	33	298	3,278	'before.to'
közé	1	1	145	8	235	36	531	957	'between.to'
föle	135	19	23	112	82	5	49	425	'above.to'

Preverbs above the middle line are formally identical to oblique pronouns, while the ones below the line are homonymous with postpositions in the synchronic system. Only preverbs occurring in direct order preverb – finite verb constructions are counted. **Stems indicate the oblique relation while the person/number suffixes (PNMS) function as pronominals.**

## Partial construction distributions for inflected PVS

- (1) a. Bocs, hogy megint **belé-d**-köt-ök.  
sorry that again PV-2SG-tie-1SG  
'I'm sorry to provoke you again.'
- b. A vita és a **belé-m**-köt-és nem ugyanaz!  
the argument and the PV-1SG-tie-NMLZ not same  
'Arguing and provoking me is not the same!'
- c. Biztos akad néhány **belé-m**-köt-nivaló.  
certainly occur some PV-1SG-tie-NMLZ  
'There must to be some reason to provoke me.'
- d. a piros ruhá-s, **belé-m**-köt-ős hapsi  
the red dress-ADJZ PV-1SG-tie-ADJZ guy  
'the guy dressed in red who has the habit of provoking me'

# Exaptation - the diachronic redeployment of existing constructions for new purposes



## Concluding observations

- Similar trajectories across languages that eventuate in complex predicates produce structures that violate (somewhat simple) expected distributions and behaviors of constitutive pieces.
- These unexpected structures persist within and across families, suggesting that satisficing, i.e., being good enough, within whole, specific grammar systems leads to necessary imperfectability.

# Ket verbal template

Templatic model of the Ket verb (simplified from Nefedov & Vajda 2015, 36)

P8	P7	P6	P5	P4	P3	P2	P1	P0	P-1
subject class/ person	stem	subject or object features	stem	subject or object features or tense	subject or object features or empty morph	tense/ mood	subject or object features or resultative	stem	plural subject (in some verbs that use P8 for subject)

# Optimization

- Optimization refers to the attempt to find the best option out of a set of potential courses of action. Simon (1955, 1972) contrasted optimization with satisficing. He defined optimization as the selection of best choice, the one with the highest expected utility. (A more technically accurate term would be “maximizing”, as in maximizing expected utility, rather than “optimizing.”) Unfortunately, according to Simon, the computational requirements for optimizing are usually too high for decision makers. Therefore, people will satisfice, which means selecting the first course of action that appears to be successful, even if it is not the best. Satisficing is a heuristic that allows decision makers to overcome their limited information-processing capacities. (Klein)