The morphological causative in Panãra: A Grammar Matrix implementation

●●●

Carly Crowther, Emily Luedke, Emily M. Bender
University of Washington

Road Map

- Panãra language background

- Grammar Matrix & Customization system implementation

- Causative morpheme problem

- Solution & analysis

- Conclusion

Motivation

- We are computational linguistics masters students at the University of Washington
- LING 567: Knowledge Engineering for NLP
  - Graduate level grammar engineering class
  - Using the Grammar Matrix Customization System (Bender et al. 2010) to build an implemented grammar for a language
- Class goals:
  - Incremental development of a hand-built implemented grammar
  - Test new Grammar Matrix libraries
  - Explore if the Grammar Matrix itself holds up given new data from a new language
- Panãra is studied by other researchers at UW (Lapierre)

Language Background

- Panãra is part of the Jê language family
- ~630 speakers, largely monolingual
- Speakers live between the states of Mato Grosso and Pará in Brazil
- Previous work by Dourado, Bardagil, and Lapierre
- Work in this presentation comes from Bardagil's 2018 dissertation: *Case and Agreement in Panará*

Terminology

- "subject/object" refer to the HPSG distinction between SUBJ and COMPS

- S/A/O refer to core arguments
  - S = intransitive "subject"
  - A = transitive "subject" or agentive argument
  - O = transitive "object" or patientive argument

Language Overview - Alignment

- Ergative-Absolutive syntactic alignment
- **A** argument is case marked with the ergative morpheme *hẽ*
- **S/O** arguments are unmarked and have absolutive case
- Dual and plural pronouns have a slightly different form of ergative case marking (*-ân*)

(1) Jyrawâ inkjẽ.
Jy-ra-wâ   inkjẽ
INTR-1SG.S-born   1SG
'I was born.' (Bardagil 2018:103)

(2) Karân kamẽrânpun inkjẽ.
Ka-rân   ka-mẽ-r-ânpun   inkjẽ
2SG-DU.ERG   2SG.A-DU-1SG.O-see   1SG
'You two saw me.' (Bardagil 2018:121)

(3) Inkjẽ hẽ rêsunpa nãkãã.
Inkjẽ hẽ   rê-s-unpa   nãkãã
1SG   ERG   1SG.A-3SG.O-fear   snake
'I'm scared of snakes.' (Bardagil 2018:59)

Language Overview - Verbal Morphology

- All core arguments are optionally expressed
- Transitive and intransitive verbs are obligatorily inflected with a set of agreement prefixes that agree in person and number with the arguments

A agreement paradigm

Person	SG	DUAL	PL
1	rê	rê...mẽ	nẽ
2	ka	ka...mẽ	ka rê
3	ti	ti...mẽ	nẽ

(Bardagil 2018:113)

S/O agreement paradigm

Person	SG	DUAL	PL
1	ra (r)	mẽ...ra (r)	ra
2	a (k)	mẽ...a (k)	rê a (k)
3	∅ (s/j)	mẽ...∅ (s/j)	ra (r)

(Bardagil 2018:117)

Language Overview - Verb Structure

- S/O agreement prefixes always appear adjacent to the verb root
- A agreement prefixes occur to the left of S/O in transitive constructions
- If one or both referents are dual number, the dual prefix *mẽ-* will appear directly to the left of the absolutive prefixes (S or O)
- Intransitive verbs are further inflected with the intransitive morpheme *jy-*

Intransitive Verbs

Intransitive-	(Dual-)	S/O-	Verb root

Transitive Verbs

A-	(Dual-)	S/O-	Verb root

Language Overview - Data

Dual S argument: dual + 2SG S; intransitive morpheme *jy-*

(4)   Kara jymẽapôô
       Kara   jy-mẽ-a-pôô
       2DU   INTR-DU-2SG.S-arrive
       'You two arrived.' (Bardagil 2018:27)

Dual A argument: 3SG A + dual; transitive verb, no intransitive morphology

(5)   Perankô mẽ Mĩkre hẽ timẽkre kwy
       Perankô   mẽ   Mĩkre   hẽ   ti-mẽ-ɸ-krẽ   kwy
       Perankô   and   Mĩkre   ERG   3SG.A-DU-3SG.O-cook   manioc
       'Perankô and Mĩkre cooked manioc.' (Bardagil 2018:107)

Grammar Matrix & Customization System

- LinGO Grammar Matrix: open-source starter-kit for building implemented grammars in the HPSG formalism (Bender, Flickinger & Oepen 2002)
    - Abstracted from the implemented HPSG grammars of English, Japanese, German, & Spanish
    - TDL code that will be useful/applicable to many, if not all, languages
    - Over 130 languages modeled by students in LING 567 since 2004

- Customization system: web-based questionnaire to enhance coverage of phenomena specific to a user's language (Bender et al. 2010)
    - Series of questions of typological information about a language
    - Outputs even more detailed TDL customized to a language

Customization System - Verb Inflection Position Classes

▼ A-pn (verb-pc3)

X **Verb Position Class 3**:
Position Class Name: [A-pn]
Obligatorily occurs: ☑
Appears as a prefix or suffix: [Prefix ▾]
Possible inputs: [dual-transitive (verb-pc2)] [▾]

Morphotactic Constraints:

[Add a Require constraint]

[Add a Forbid constraint]

Lexical Rule Types that appear in this Position Class:

▶ first-non-plural-A (verb-pc3_lrt1)
▶ second-non-plural-A (verb-pc3_lrt2)
▶ third-non-plural-A (verb-pc3_lrt3)
▶ first-or-third-plural-A (verb-pc3_lrt4)
▶ second-plural-A (verb-pc3_lrt5)
▶ first-sg-irr (verb-pc3_lrt7)
▶ first-dual-irr (verb-pc3_lrt8)

[Add a Lexical Rule Type]

A-pn : agent's person and number

- Position classes determine the order in which morphemes appear (Goodman 2013)
- Each position class has multiple LRTs for different person/number combinations

▼ first-non-plural-A (verb-pc3_lrt1)

X **Lexical Rule Type 1**:
Name: [first-non-plural-A]
Supertypes: [ ] [▾]

Features:

X Name: [Person ▾] Value: [1st] [▾] Specified on: [The subject ▾]
X Name: [Subtype {i} number ▾] Value: [singular, dual] [▾] Specified on: [The subject ▾]
X Name: [Mood ▾] Value: [realis] [▾] Specified on: [The verb ▾]

[Add a Feature]

Valence-changing operations may modify the valence structure of
a verb by adding or removing either a subject or object, possibly
including changes to e.g. case frames or adding predicates. (**Experimental**)

[Add a valence-changing operation]

Morphotactic Constraints:

[Add a Require constraint]

[Add a Forbid constraint]

Lexical Rule Instances:

X | Instance 1 ○ No affix ● Affix spelled [rê-]

[Add a Lexical Rule Instance]

Lexical rule instances provide correct spelling of the morpheme from the A and S/O paradigms

Customization System - Verb Inflection Position Classes

Verb Inflection

▼ O-pn (verb-pc1)

▼ first-non-plural-O (verb-pc1_lrt1)

[X] **Lexical Rule Type 1:**
Name: [first-non-plural-O]
Supertypes: [　　　　▼]

Features:

[X] Name: [Person ▼] Value: [1st　　　▼] Specified on: [The object ▼]
[X] Name: [Subtype {i} number ▼] Value: [singular, dual ▼] Specified on: [The object ▼]

[Add a Feature]

Valence-changing operations may modify the valence structure of a verb by adding or removing either a subject or object, possibly including changes to e.g. case frames or adding predicates. (**Experimental**)

[Add a valence-changing operation]

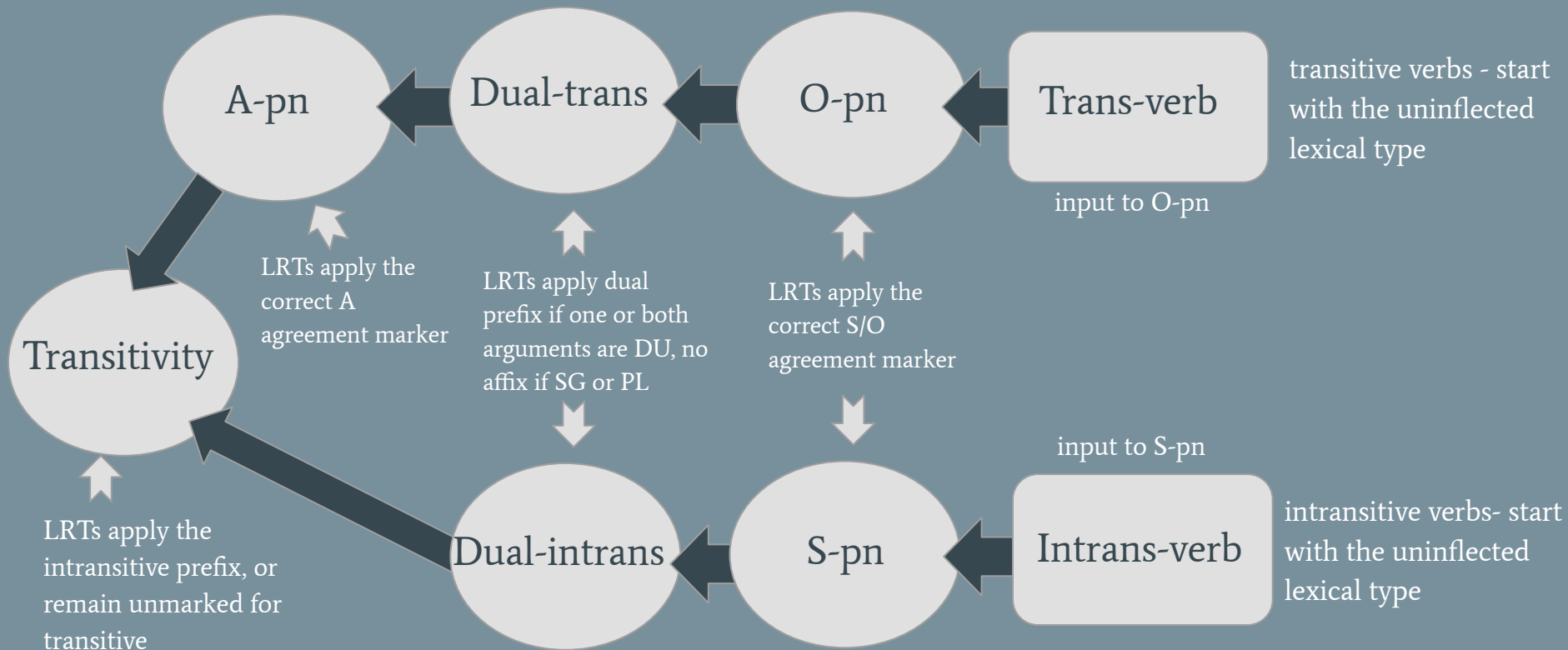Morphotactic Constraints:

[Add a Require constraint]

[Add a Forbid constraint]

Lexical Rule Instances:

[X] Instance 1 ○No affix ●Affix spelled [ra-]

[Add a Lexical Rule Instance]

▼ S-pn (verb-pc4)

▼ first-non-plural-S (verb-pc4_lrt2)

[X] **Lexical Rule Type 2:**
Name: [first-non-plural-S]
Supertypes: [　　　　▼]

Features:

[X] Name: [Person ▼] Value: [1st　　　▼] Specified on: [The subject ▼]
[X] Name: [Subtype {i} number ▼] Value: [singular, dual ▼] Specified on: [The subject ▼]

[Add a Feature]

Valence-changing operations may modify the valence structure of a verb by adding or removing either a subject or object, possibly including changes to e.g. case frames or adding predicates. (**Experimental**)

[Add a valence-changing operation]

Morphotactic Constraints:

[Add a Require constraint]

[Add a Forbid constraint]

Lexical Rule Instances:

[X] Instance 1 ○No affix ●Affix spelled [ra-]

[Add a Lexical Rule Instance]

Visualizing the Verb Inflection Position Classes

A-pn

Dual-trans

O-pn

Trans-verb

transitive verbs - start with the uninflected lexical type

input to O-pn

Transitivity

LRTs apply the correct A agreement marker

LRTs apply dual prefix if one or both arguments are DU, no affix if SG or PL

LRTs apply the correct S/O agreement marker

LRTs apply the intransitive prefix, or remain unmarked for transitive

Dual-intrans

S-pn

Intrans-verb

input to S-pn

intransitive verbs- start with the uninflected lexical type

Causative Morpheme

- The causative morpheme *ho-* can inflect an intransitive verb
  - Adds a subject and increases valence from 1 to 2 core arguments
  - Appears between the two argument agreement affixes
- These constructions behave exactly like transitives:
  - Ergative case marking on the overt A argument
  - Two argument agreement affixes– one that agrees with the ergative case of the A argument & one that agrees with the absolutive case of the other argument
  - No intransitive morphology is present on the resulting verb
- The presence of this morpheme is ungrammatical with underlyingly transitive verb roots.

Causative Morpheme - Examples

(6) Ka jõpãã jysõti.
     Ka   jõpãã   jy-ɸ-sõti
     2SG   child   INTR-3SG.S-sleep
     'Your child sleeps.' (Bardagil 2018: 108)

(7) Ka hẽ kahosõti ka jõpãã.
     Ka   hẽ   ka-**ho**-ɸ-sõti   ka   jõpãã
     2SG   ERG   2SG.A-CAUS-3SG.O-sleep   2SG   child
     'You made your child sleep.' (Bardagil 2018: 108)

(8) *Inkjẽ hẽ rêhokuri inkjẽ jõpãã suasĩra jĩ
     Inkjẽ hẽ   rê-**ho**-ɸ-kuri   inkjẽ   jõpãã   suasĩra   jĩ
     1SG   ERG   1SG.A-CAUS-3SG.O-eat   1SG   child   peccary   meat
     'I made my child eat peccary meat.' (Bardagil 2018: 174)

The Problem...
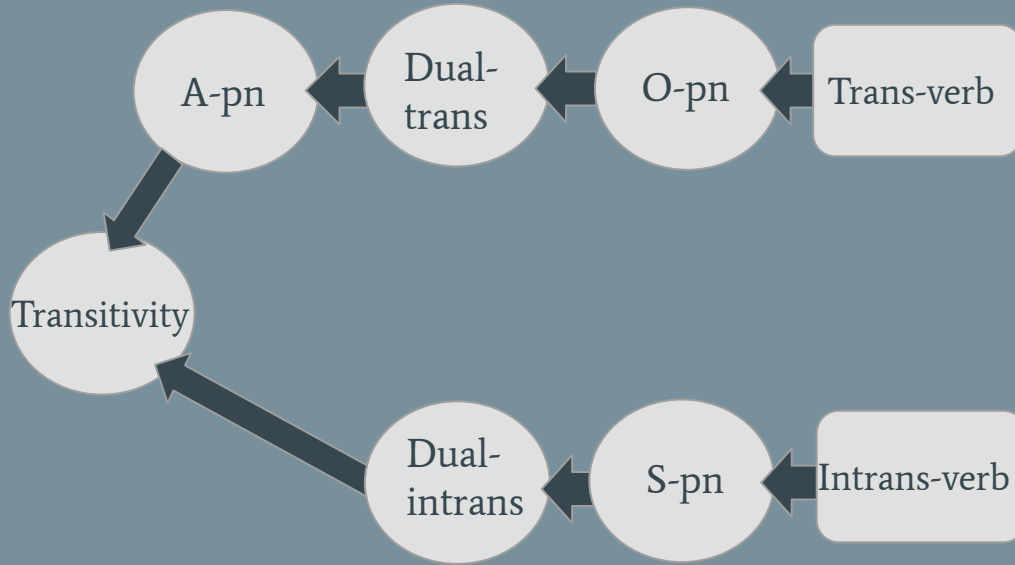
Causative verb construction

A-	(Dual-)	Causative-	S/O-	Verb root

- The causative morpheme appears outside of the S/O prefix
- The resulting sentence should be inflected like a transitive verb
- But, since inflection is handled via ordered position classes, the intransitive verb will already be inflected for the S-pn position class
- It's impossible for a verb to be inflected with A-pn after inflecting for S-pn

The Solution!
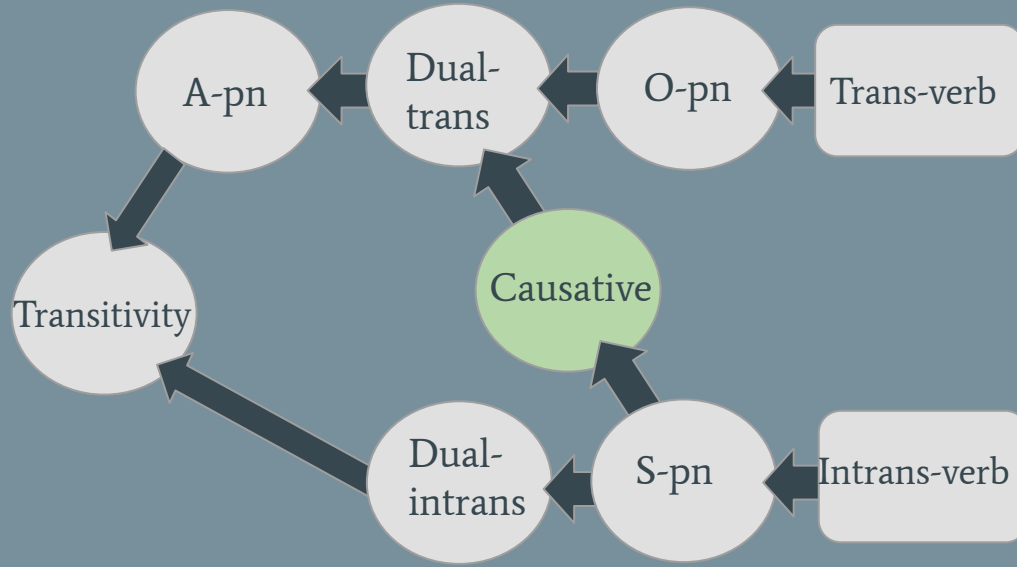
Causative verb construction

A-	(Dual-)	Causative-	S/O-	Verb root

- The S and O morphemes are homophonous, as they both mark agreement with absolute arguments
- A new causative rule moves information from the intransitive "subject" to the resulting transitive "object"
- S/O orthographic identicality allows us to obtain the correct surface form for the O argument agreement through the S-pn position class

Customization System - Causative Position Class

▼ Causative (verb-pc9)

X **Verb Position Class 9**:
Position Class Name: `Causative`
Obligatorily occurs: ☐
Appears as a prefix or suffix: `Prefix ▼`
Possible inputs: ? `S-pn (verb-pc4)` `▼`

- Possible inputs limited to S-pn lexical types, so this PC will only apply to intransitives
- Changed the Dual-trans position class to accept output of the Causative PC as its input
- Subject-adding operation in the Valence-changing library (Curtis 2018) handles the rest

▼ causative-marker (verb-pc9_lrt1)

X **Lexical Rule Type 1**:
Name: `causative-marker`
Supertypes: `▼`

Features:

X Name: ? `case ▼` Value: `erg ▼` Specified on: `The subject ▼`

`Add a Feature`

Valence-changing operations may modify the valence structure of a verb by adding or removing either a subject or object, possibly including changes to e.g. case frames or adding predicates. (**Experimental**)

X Type: `Subject-adding (e.g. causative) ▼`

Most valence-changing operations currently must operate on a known input valence. If separate lexical rule types. (This may change in the future).

Should apply to: `Intransitive ▼` targets

Object-adding operations currently only support strict transitive verbs as inputs. For subject- and object-adding operations, also specify (ignored for other operations):
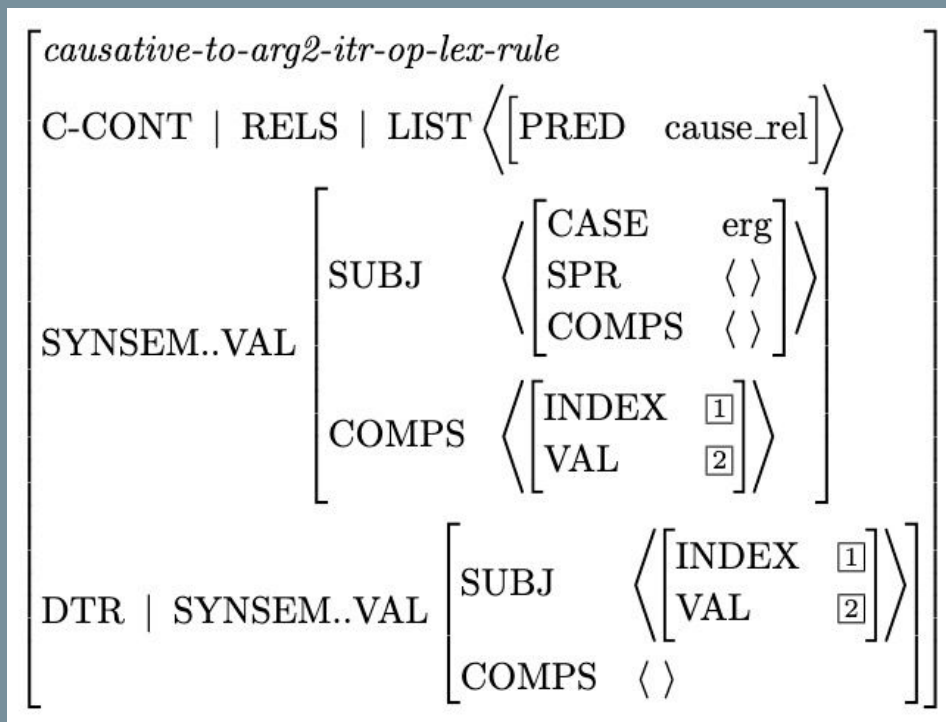
Predicate: `cause_rel`
The added argument/erstwhile subject is at the: `front ▼` of the complements list.
The added argument must be a(n): `NP ▼`

Lexical Rule Instances:

X Instance 1 ○ No affix ● Affix spelled `ho-`

Causative Rule Feature Structure

$$\begin{bmatrix} \textit{causative-to-arg2-itr-op-lex-rule} \\ \text{C-CONT} \mid \text{RELS} \mid \text{LIST} \left\langle \begin{bmatrix} \text{PRED} & \text{cause\_rel} \end{bmatrix} \right\rangle \\ \text{SYNSEM..VAL} \begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \text{CASE} & \text{erg} \\ \text{SPR} & \langle\,\rangle \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \right\rangle \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{VAL} & \boxed{2} \end{bmatrix} \right\rangle \end{bmatrix} \\ \text{DTR} \mid \text{SYNSEM..VAL} \begin{bmatrix} \text{SUBJ} & \left\langle \begin{bmatrix} \text{INDEX} & \boxed{1} \\ \text{VAL} & \boxed{2} \end{bmatrix} \right\rangle \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \end{bmatrix}$$

- Subject-adding valence-changing inflectional rule
- Based on the TDL produced by the valence-changing operation library (Curtis 2018)
- Identifies the INDEX and VAL values of the daughter's SUBJ with the mother's COMPS
- New SUBJ constrained to have ergative case
- Contributes a cause_rel predicate in the semantics

Testing the Implementation

- Created a test suite with 230 grammatical and ungrammatical examples
    - Some items drawn from Bardagil's (2018) dissertation, 142 edited/simplified by us to isolate relevant features for the purpose of incremental development
- Used [incr tsdb()] grammar profiling software to measure coverage and over-generation (Oepen & Flickinger 1998)
    - 84.3% coverage (107/127); 13.6% overgeneration (14/103)
- 8 test suite items addressed the causative morpheme
    - 100% coverage; 0% overgeneration
- Successful implementation without introducing added ambiguity

Tree Structure

(9) Ka hẽ kahosõti jõpãã.
Ka   hẽ   ka-ho-ø-sõti   jõpãã
2SG ERG 2SG.A-CAUS-3SG.O-sleep   child
'You made the child sleep.' (Based on Bardagil 2018: 108)
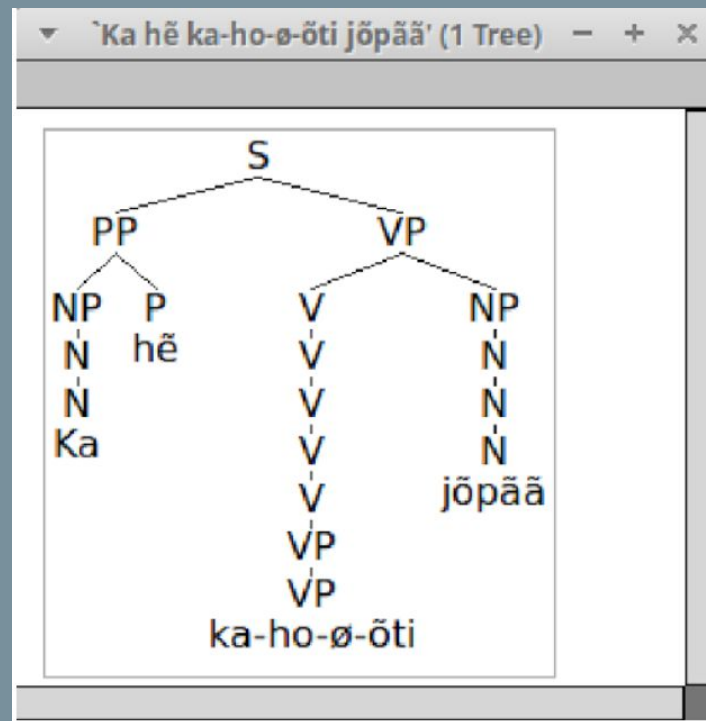
MRS

The ARG0 of the pronoun (2SG) is identified with the ARG1 of the cause relation

(9)   Ka hẽ kahosõti jõpãã.

Ka   hẽ   ka-ho-ø-sõti   jõpãã

2SG   ERG   2SG.A-CAUS-3SG.O-sleep   child

'You made the child sleep.' (Based on Bardagil 2018: 108)

MRS

The ARG0 of the child relation is identified with the ARG1 of the sleep relation and the ARG2 of the cause relation

(9) Ka hẽ kahosõti jõpãã.

Ka   hẽ   ka-ho-ϕ-sõti   jõpãã

2SG   ERG   2SG.A-CAUS-3SG.O-sleep   child

'You made the child sleep.' (Based on Bardagil 2018: 108)
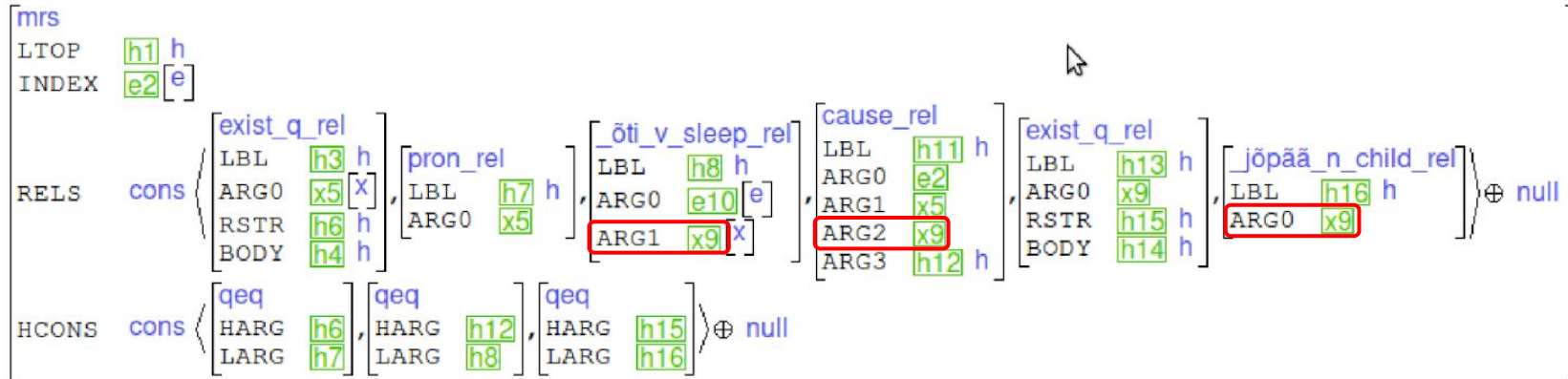
MRS

The Label of the sleep relation is identified with the ARG3 of the cause relation via a qeq identity

(9) Ka hẽ kahosõti jõpãã.

Ka	hẽ	ka-ho-ɸ-sõti	jõpãã
2SG	ERG	2SG.A-CAUS-3SG.O-sleep	child

'You made the child sleep.' (Based on Bardagil 2018: 108)

Conclusion

- Causative constructions take an intransitive verb and add a second core argument. The way that we modeled A/S/O agreement prefixes made the combination of an intransitive verb with the A-pn position class impossible.

- We take advantage of the orthographic and syntactic parallels between the S and O agreement affixes, modeling a "switch" from the intransitive position class inflection chain to the transitive inflection chain upon the presence of the causative morpheme.

- This analysis has been validated against a testsuite of 230 grammatical and ungrammatical sentences, 8 of which include the causative morpheme.

Thank you!

Special thanks to the anonymous reviewers for their helpful feedback.

References

Bardagil, Bernat. 2018. *Case and agreement in Panará*. University of Groningen PhD dissertation. LOT Publications n. 511.

Bender, Emily M., Scott Drellishak, Antske Fokkens, Laurie Poulson & Safiyyah Saleem. 2010. Grammar Customization. *Research on Language & Computation,* 8(1), 23–72.

Bender, Emily M., Dan Flickinger & Stephan Oepen. 2002. The Grammar Matrix: An Open-Source Starter-Kit for the Rapid Development of Cross-linguistically Consistent Broad-coverage Precision Grammars. In *COLING-02: Grammar Engineering and Evaluation.*

Copestake, Ann. 2002. *Implementing Typed Feature Structure Grammars*. CSLI publications.

Copestake, Ann, Dan Flickinger, Carl Pollard & Ivan Sag. 2005. Minimal Recursion Semantics: An Introduction. *Research On Language And Computation*. 3. 281-332.

Curtis, Christian Michael. 2018. *A Parametric Implementation of Valence-changing Morphology in the LinGO Grammar Matrix*. University of Washington MA thesis.

Goodman, Michael Wayne. 2013. Generation of Machine-Readable Morphological Rules from Human-Readable Input. *University of Washington Working Papers in Linguistics*. 30.

Lapierre, Myriam. 2023. The Phonology of Panãra: A Segmental Analysis. *International Journal of American Linguistics* 89(2). 183–218. The University of Chicago Press.

Oepen, Stephan & Dan Flickinger. 1998. Towards systematic grammar profiling: Test suite technology ten years after. *Computer Speech and Language: Special Issue on Evaluation* 12. 411-436.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar.* The University of Chicago Press.

Zamaraeva, Olga, Chris Curtis, Guy Emerson, Antske Fokkens, Michael Goodman, Kristen Howell, TJ. Trimble & Emily M. Bender. 2022. 20 years of the Grammar Matrix: Cross-linguistic hypothesis testing of increasingly complex interactions. *Journal of Language Modelling* 10(1). 49-137.