# CAUSATIVES CLUSTER SEMANTICALLY BY JUNCTURE LEVEL ACROSS LANGUAGES

## THE 17TH INTERNATIONAL CONFERENCE ON ROLE AND REFERENCE GRAMMAR (RRG 2023) HHU DÜSSELDORF, AUGUST 14–16, 2023


Erika Bellingham,[1] Jürgen Bohnemeyer,[1] Andrea Ariño-Bizarro,[2] Emanuel Bylund,[3,4] Jing Du,[5] James Essegbey,[6] Stephanie Evers,[1] Saima Hafeez,[1] Iraide Ibarretxe-Antuñano,[2] Pia Järnefelt,[7] Kazuhiro Kawachi,[8] Yu Li,[9] Thomas Fuyin Li,[10] Tatiana Nikitina,[11] Sang-Hee Park,[12] Anastasia Stepanova,[1] and Guillermo Montero-Melis[4,13]

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# TESTING ASSUMPTIONS ABOUT THE SYNTAX–SEMANTICS INTERFACE

- arguably the two top-most distinctive features of Role and Reference Grammar (RRG)
  - a typological focus
    - the theory was developed to fit the facts of a sample of morphosyntactically wildly distinct languages
- universals of grammar are located, not in morphosyntactic form, but in patterns of the syntax-semantics interface

- arguably the two top-most distinctive features of Role and Reference Grammar (RRG) (cont.)

"The specific questions which stimulated the development of Role and Reference Grammar [RRG] were, '**What would a linguistic theory look like if it were based on the analysis of languages with diverse structures, such as Lakhota, Tagalog, Dyirbal and Barai (Papua New Guinea), rather than on the analysis of English and similar languages?**', and '**How can the interaction of syntax, semantics and pragmatics in different grammatical systems best be captured and explained?**' The two questions highlight the profound implications of the analysis of typologically diverse languages for the formulation of a linguistic theory, and they indicate that the resulting theory will be one in which semantics and pragmatics play significant roles. In other words, **RRG is a theory of the syntax–semantics–pragmatics interface.** '" (Van Valin 2023: 18; emphasis *JB*)

- building blocks of RRG
  as a theory of the syntax-semantics interface
  - sentences are licensed by a linking algorithm
    - which describes their form-to-meaning mapping

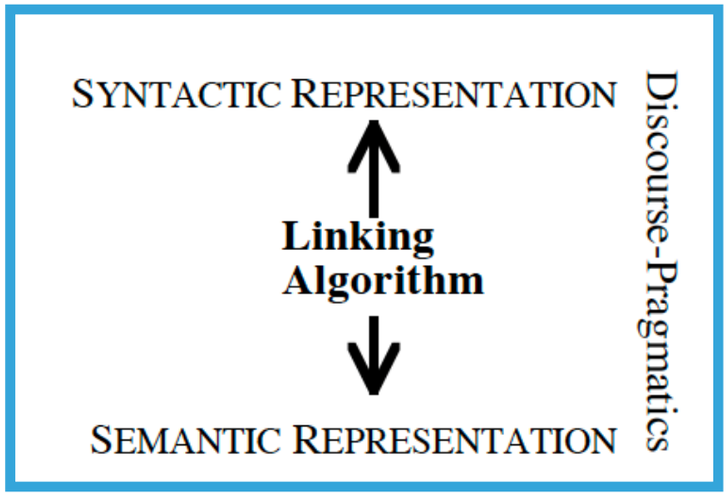



**Figure 1.1.** *The role of the linking algorithm in the parallel architecture of RRG (Van Valin 2023: 1)*

The linking algorithm from semantics to syntax
a. Step 1: Construct the semantic representation of the sentence based on the LS of the main predicator.
b. Step 2: Assign actor and undergoer, following the actor-undergoer hierarchy [AUH].
c. Step 3: Determine the morphosyntactic coding of the arguments of the main predicator.
   i. Select the privileged syntactic argument [PSA], following the PSA selection hierarchy.
   ii. Assign the arguments the appropriate case markers and/or adpositions.
   iii. Assign agreement marking to the main verb or auxiliary, as appropriate.
d. Step 4: Select the syntactic templates for the sentence, following the syntactic template selection principles.
e. Step 5: Assign the nucleus, the arguments and the adjuncts to positions in the syntactic representation of the sentence.

**Figure 1.2.** *How the linking algorithm works: the jet perspective (Van Valin 2023: 97)*

- building blocks of RRG
  as a theory of the syntax-semantics interface (cont.)
  - the linking algorithm makes reference to key elements of RRG's theory of the syntax-semantics interface
    - the macro-role hierarchy

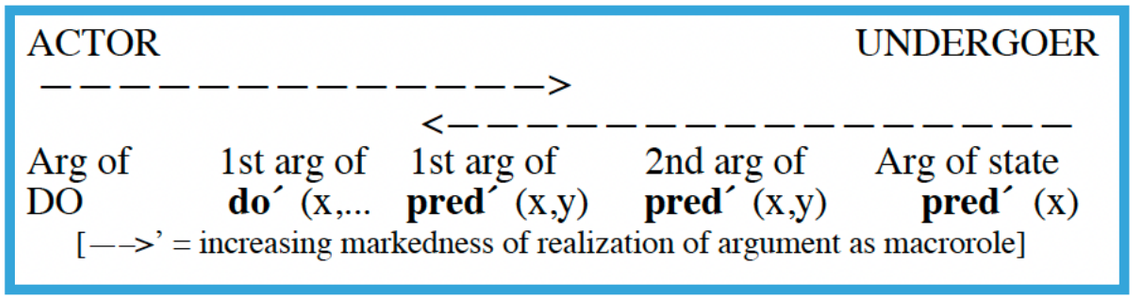



**Figure 1.3.** *The macro-role hierarchy (Van Valin 2023: 89)*

  - the layered structure of the clause (LSC)
    - a semantically "bootstrapped" syntactic architecture
      - the layers of which are defined in terms of communicative functions

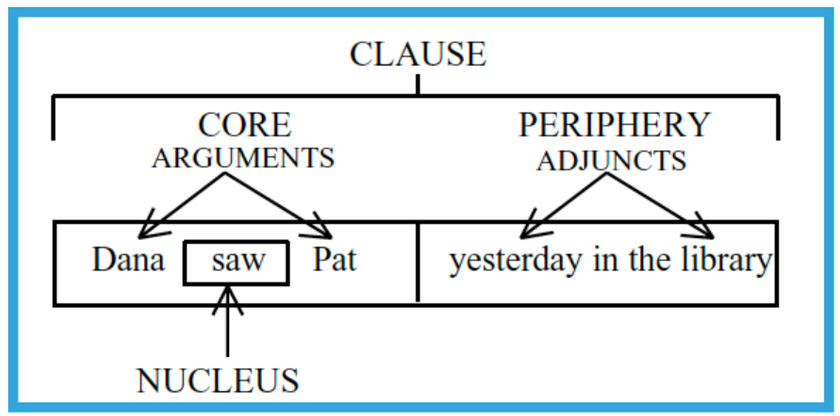



**Figure 1.4.** *The Layered Structure of the Clause schematically (Van Valin 2005: 4)*

## TESTING ASSUMPTIONS ABOUT THE SYNTAX-SEMANTICS INTERFACE (CONT.)

- the goal of the present study
  - an empirical test of the LSC
    - in one particular semantic domain: **causation**
  - the question:
    - *Do causatives semantically cluster across languages around the three juncture layers*
      - *of nuclear-, core-, and clause-layer junctures?*

| | |
|---|---|
| [CORE …[NUC … ]… + …[NUC … ] …] | Nuclear juncture |
| [CLAUSE …[CORE …]… + …[CORE …] …] | Core juncture |
| [SENTENCE …[CLAUSE …]… + …[CLAUSE …] …] | Clausal juncture |
| [TEXT…[ SENTENCE…]… + …[ SENTENCE…] …] | Sentential juncture |

**Figure 1.5.** *The LSC and the juncture levels in the theory of clause combination (Van Valin 2023: 50)*

## TESTING ASSUMPTIONS ABOUT THE SYNTAX-SEMANTICS INTERFACE (CONT.)

- motivating the question
  - the layers of the LSC are semantically bootstrapped
    - nuclei are predicators
    - cores express predicate-argument combinations
      - clauses express propositions = objects of speech acts
        - as arguably reflected in the operator hierarchy (Bohnemeyer 2019)
- this predicts a certain amount of crosslinguistic semantic uniformity associated with the layers

- motivating the question (cont.)
  - these semantic profiles of the layers express themselves in
    - the operator hierarchy
    - the interclausal relations hierarchy

Nuclear operators:
- Aspect
- Negation
- Directionals (only those modifying orientation of action or event without reference participants)

Core operators:
- Directionals (only those expressing the orientation or motion of one participant with reference to another participant or to the speaker)
- Event quantification
- Modality (root modals, e.g. ability, permission, obligation)
- Internal (narrow scope) negation

Clausal operators:
- Status (epistemic modals, external negation)
- Tense
- Evidentials
- Illocutionary Force [IF]

**Figure 1.6.** *The LSC and the operator hierarchy (Van Valin 2023: 15)*

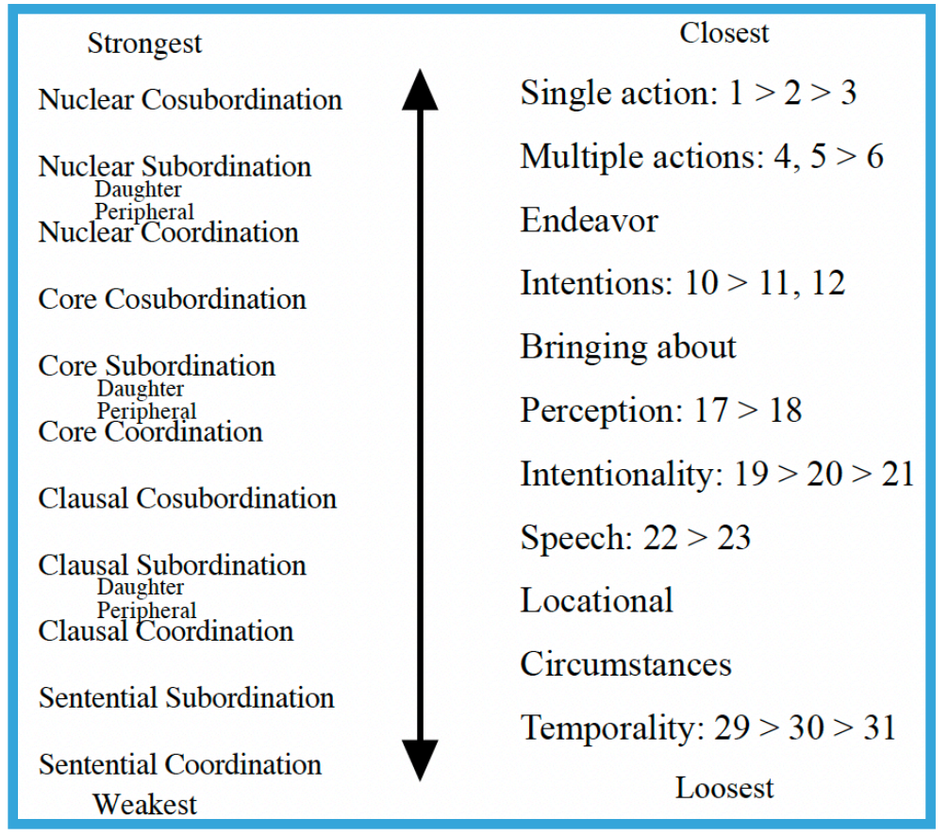



**Figure 1.7.** *The LSC and the interclausal relations hierarchy (Van Valin 2023: 68)*

- causation as a test case
  - causal relations should be expressible at all layers
    - simplex and complex nuclei express causation as part of the meanings of lexical event descriptors

(1.1) *I fed John the cookies*

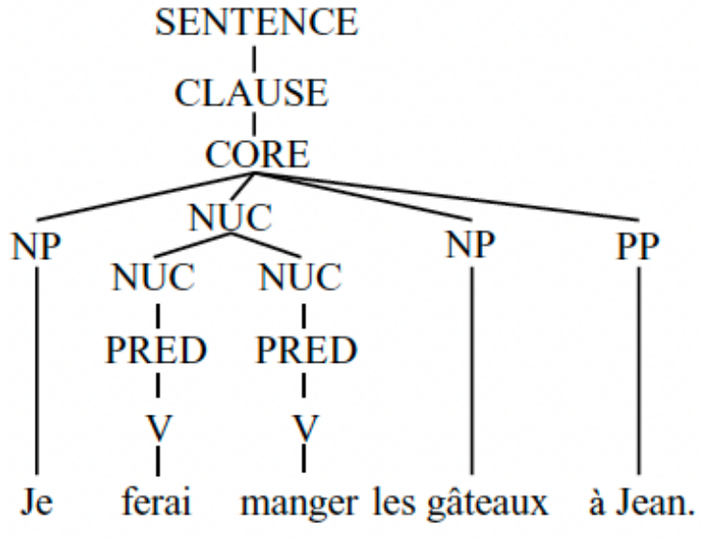
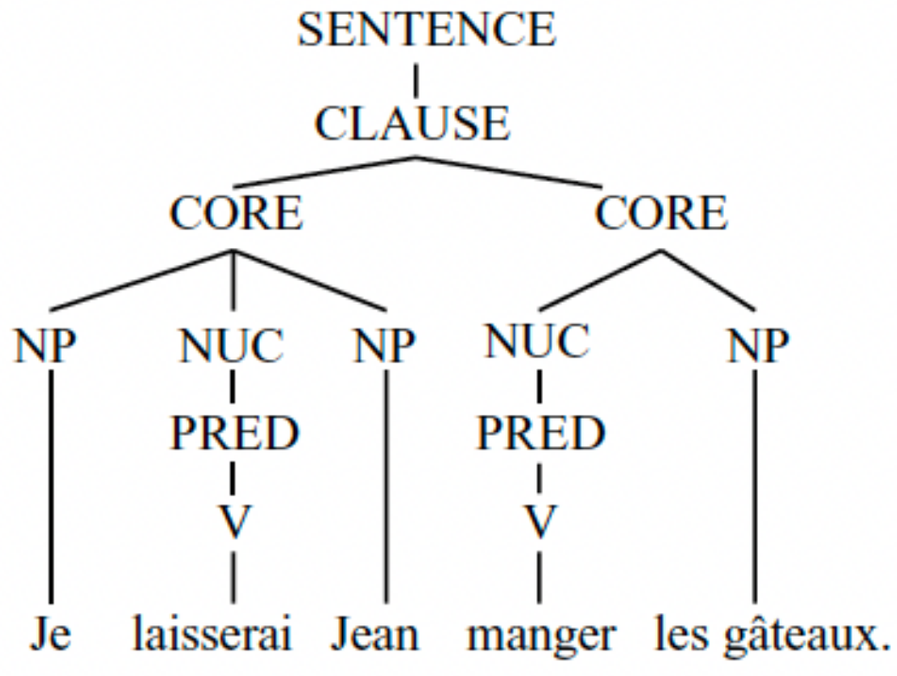



**Figure 1.8.** *French nuclear juncture causatives (Van Valin 2023: 191)*

- causation as a test case (cont.)
  - causal relations should be expressible at all layers (cont.)
    - core-layer junctures express causation as part of complex event representations

(1.2) *I made John eat the cookies*




(Van Valin 2005: 189)

- causation as a test case (cont.)
  - causal relations should be expressible at all layers (cont.)
    - clause–layer junctures express causation as a relation between propositions

(1.3) a. *The street is wet because it rained last night*
a′. *The rain caused the street to be wet*
b. *It must have rained last night, because the street is wet*
b′. #*The wetness of the street caused the rain last night*
b″. *The wetness of the street makes me think that it rained last night*

- in principle, any given event can be framed at all three layers

(1.4) [John is known to dislike cookies]
a. *John ate the cookies because I intimidated him*
b. *I must have intimidated John, because he ate the cookies*

- causation as a test case (cont.)
  - the predictions we tested:
    - causal relations can be expressed by
      - simplex nuclei and nuclear-layer junctures
        - core junctures
        - clause-layer junctures
  - but **each of these construction types should be associated with a unique semantic profile**

## TESTING ASSUMPTIONS ABOUT THE SYNTAX-SEMANTICS INTERFACE (CONT.)

- history of research
  - 50 years of typological research on causatives has focused on the broad division of labor
    - between simple and complex causatives
  - particularly the iconicity it involves and the underlying causes of this iconicity
    - Bohnemeyer et al (2010); Comrie (1981); Dixon (2000); Haiman (1983); Haspelmath (2008); Kemmer & Verhagen (1994); Levin & Rappaport-Hovav (1995); Levshina (2015), (2016), (2017); McCawley (1976, 1978); Shibatani ed. (1976); Shibatani & Pardeshi (2002); Talmy (1976); Verhagen & Kemmer (1997); Wolff (2003); *inter alia*

- history of research (cont.)
  - simple ‘direct’ causal chains favor simple causative constructions

(1.5) Le=máak=o’ t-u=**nik**-ah le=bàaso-s-o’b=o’
YUC DEF=person=D2 PRV-A3=scatter-CMP(B3SG) DEF=cup-PL-PL=D2
‘The man, he scattered the cups’

**Figure 1.2.** *HO5_cuptower*

  - more complex constructions/descriptions are preferred for more complex, ‘indirect’ chains

(1.6) a. #Le=x-ch’úupal=o’ t-u=**nik**-ah le=bàaso-s-o’b=o’
YUC DEF=female:child=D2 PRV-A3=shatter+slap-APP-CMP(B3SG) DEF=cup-PL-PL=D2
‘The girl, she scattered the cups’

b. Le=x-ch’úupal=o’ t-u=**mèet-**ah
DEF=F-female:child=D2 PRV-A3=make-CMP(B3SG)
**u=nik-ik** le=bàaso-o’b le=máak=o’
A3=scatter-INC(B3SG) DEF=cup-PL DEF=person=D2
‘The girl, she made the man scatter the cup’

**Figure 1.3.** *HUO2_cups*

## TESTING ASSUMPTIONS ABOUT THE SYNTAX-SEMANTICS INTERFACE (CONT.)

- history of research (cont.)
  - mostly missing so far: a comprehensive typological examination of the causative ecology based on primary data
    - yielding a semantic map of the domain for each language
      - exceptions
        - Bohnemeyer et al. (2010) (pilot study, data from just four languages; highly unbalanced stimulus set)
        - Levshina (2022) (movie subtitle data from 22 languages (13 Indo-European))
- our goal: contribute toward closing this gap based on a new methodology for semantic typology

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
  - The language sample
  - The cluster analysis
  - Further analyses
  - Implications for RRG

# A NEW STUDY DESIGN FOR SEMANTIC TYPOLOGY

- a new approach

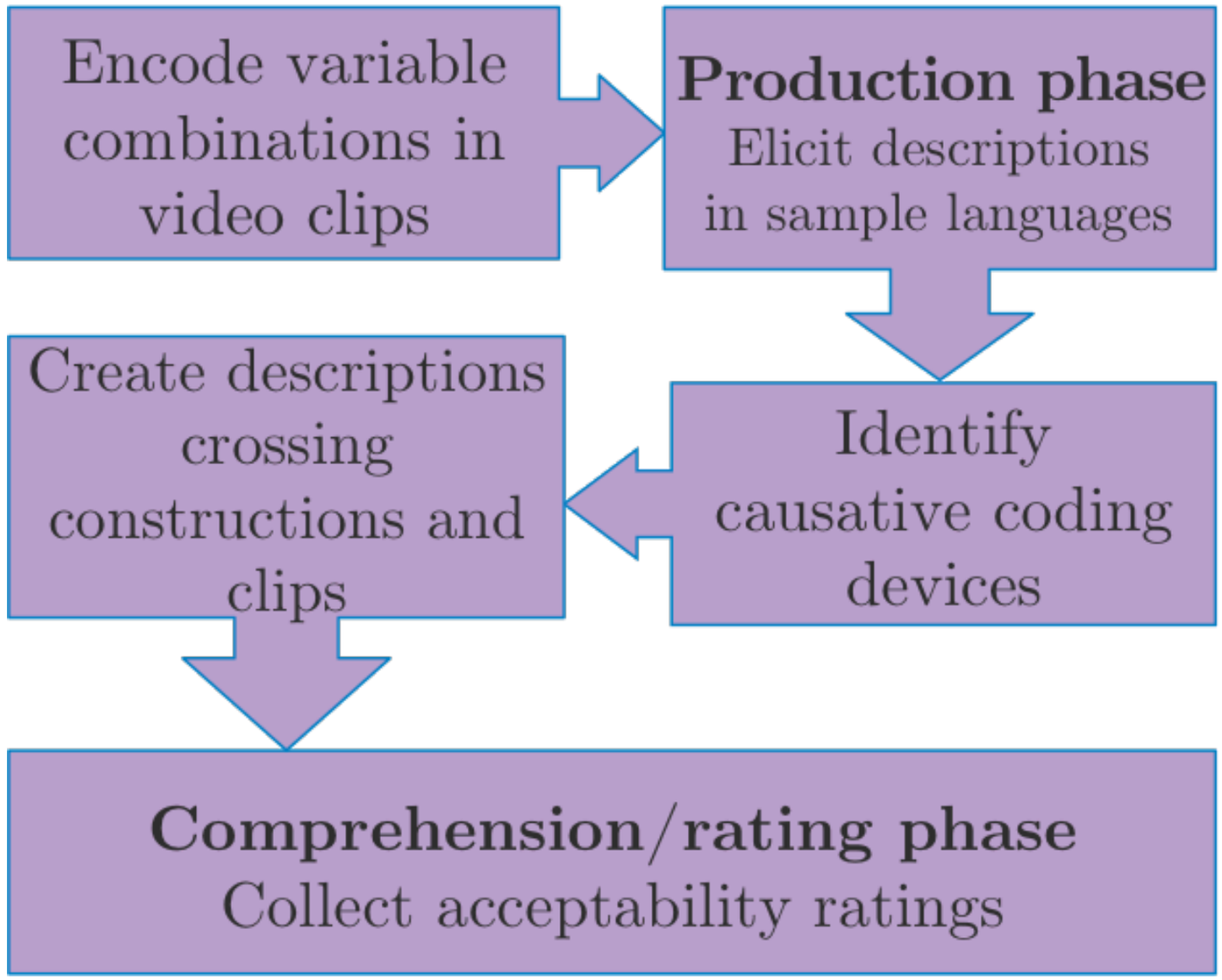



**Figure 2.1.** *A hybrid study design for semantic typology*

- advantages of this hybrid design type
  - vis-à-vis corpus studies
    - applicable to languages
      for which (large) corpora are unavailable
    - provides both positive and negative evidence
      - gives direct access to the scene being described
- vis-à-vis traditional elicited production studies
  (the staple in contemporary semantic typology)
    - allows rapid data collection and analysis
      from a larger number of speakers
    - provides both positive and negative evidence

- the rating scale
  - after some experimentation,
    we settled on a four-point qualitative scale
  - we trained the participants with the help of additional stimuli to distinguish among
    - ungrammatical utterances (1)
    - well-formed but inaccurate descriptions (2)
    - accurate but misleading descriptions (3)
    - accurate and appropriately informative descriptions (4)

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# VARIABLES AND STIMULI: THE CAL CLIPS

- variables that have been shown to impact causative choice

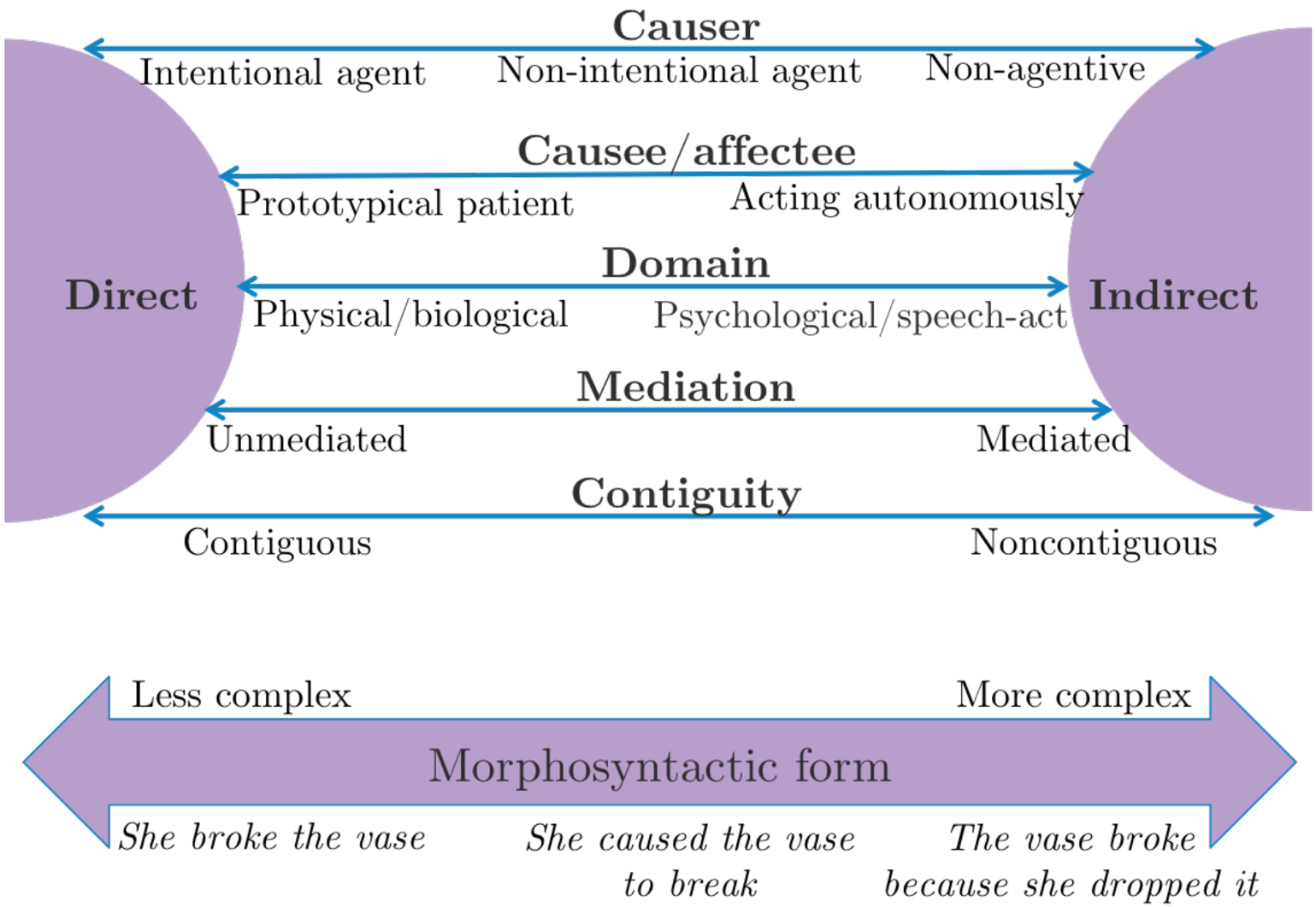



**Figure 3.1.** *A multidimensional continuum model of causation directness*

# VARIABLES AND STIMULI: THE CAL CLIPS (CONT.)

- design: E. Bellingham; J. Bohnemeyer
- 58 short video clips featuring everyday causal chains
  - most staged/enacted, a few found on the internet
- variables manipulated
  - **causer** (**CR**) type: volitional vs. accidental vs. force
  - **causee** (**CE**; = intermediate participant in the chain) type
    - volitional/controlled
    - vs. involuntary response to psychological impact
    - vs. involuntary response to mechanical impact
    - vs. no CE

- **affectee** (**AF**) type
  - volitional/controlled
  - vs. involuntary response to psychological impact
  - vs. involuntary response to mechanical impact
    - vs. physical object
- **resulting event type**
  physical state change vs. location change vs. process
- **force dynamics**
  - causation (43 core + 10 sup.) vs. letting (5 sup. scenes)

- stimuli: the CAL Clips (cont.)
  - examples
    - CR = force; CE = none; AF = mechanically impacted; resultant event = location change; FD = causation


**Figure 3.1.** *NM2_reporter*

- stimuli: the CAL Clips (cont.)
  - examples (cont.)
    - CR = accidental; CE = volitional/controlled; AF = object; resultant event = location change; FD = letting

**Figure 3.2.** *UCO1_ball*

- stimuli: the CAL Clips (cont.)
  - examples (cont.)
    - CR = volitional; CE = psychologically impacted; AF = object; resultant event = physical change; FD = letting

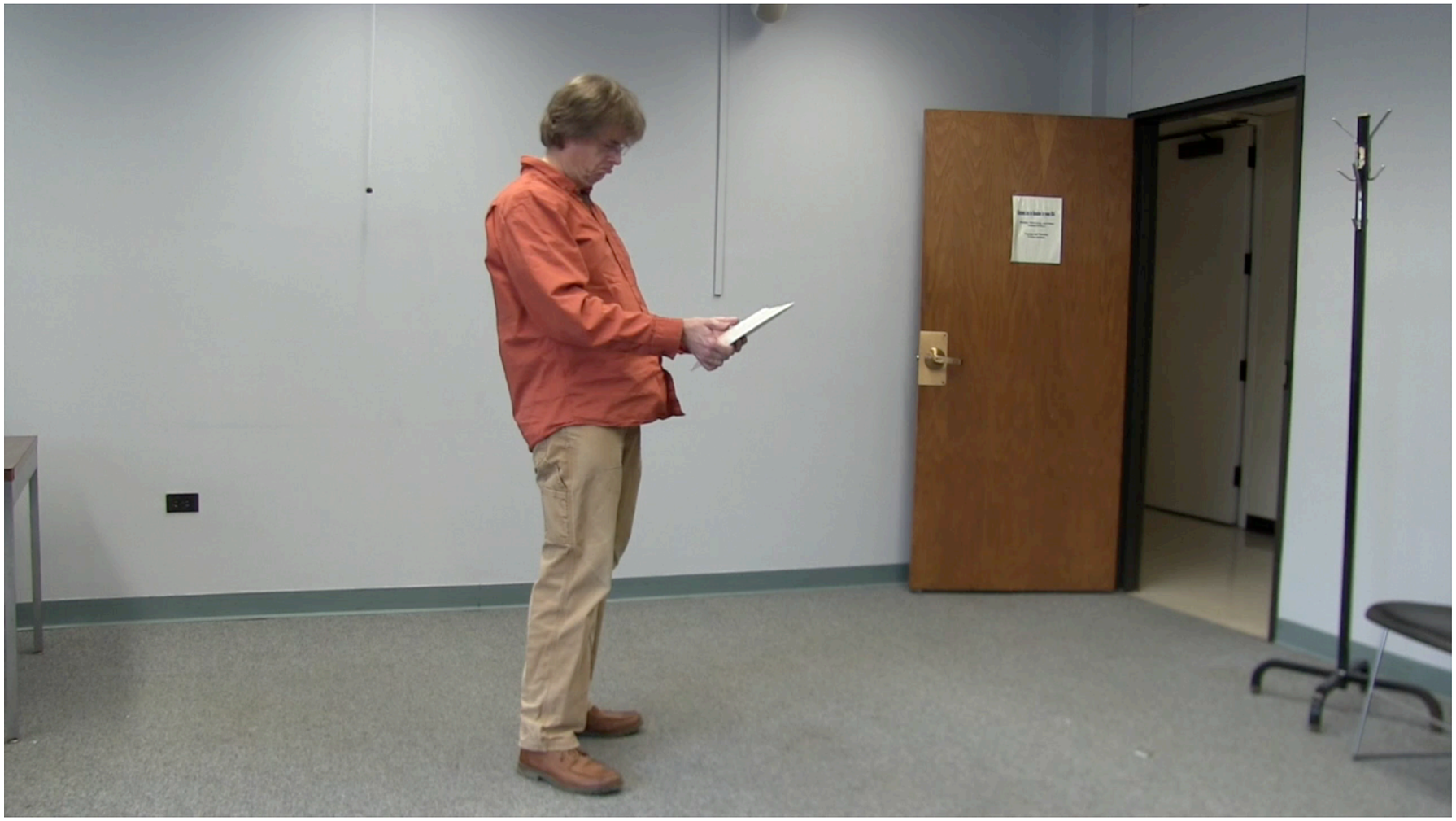


**Figure 3.3.** *HUO1_plate*

- stimuli: the CAL Clips (cont.)
  - examples (cont.)
    - CR = volitional; CE = volitional/controlled; AF = object; resultant event = process; FD = causation

**Figure 3.4.** *HCOproc1_swing*

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# THE LANGUAGE SAMPLE

- the languages from which data has been collected for the Semantic Typology subproject so far

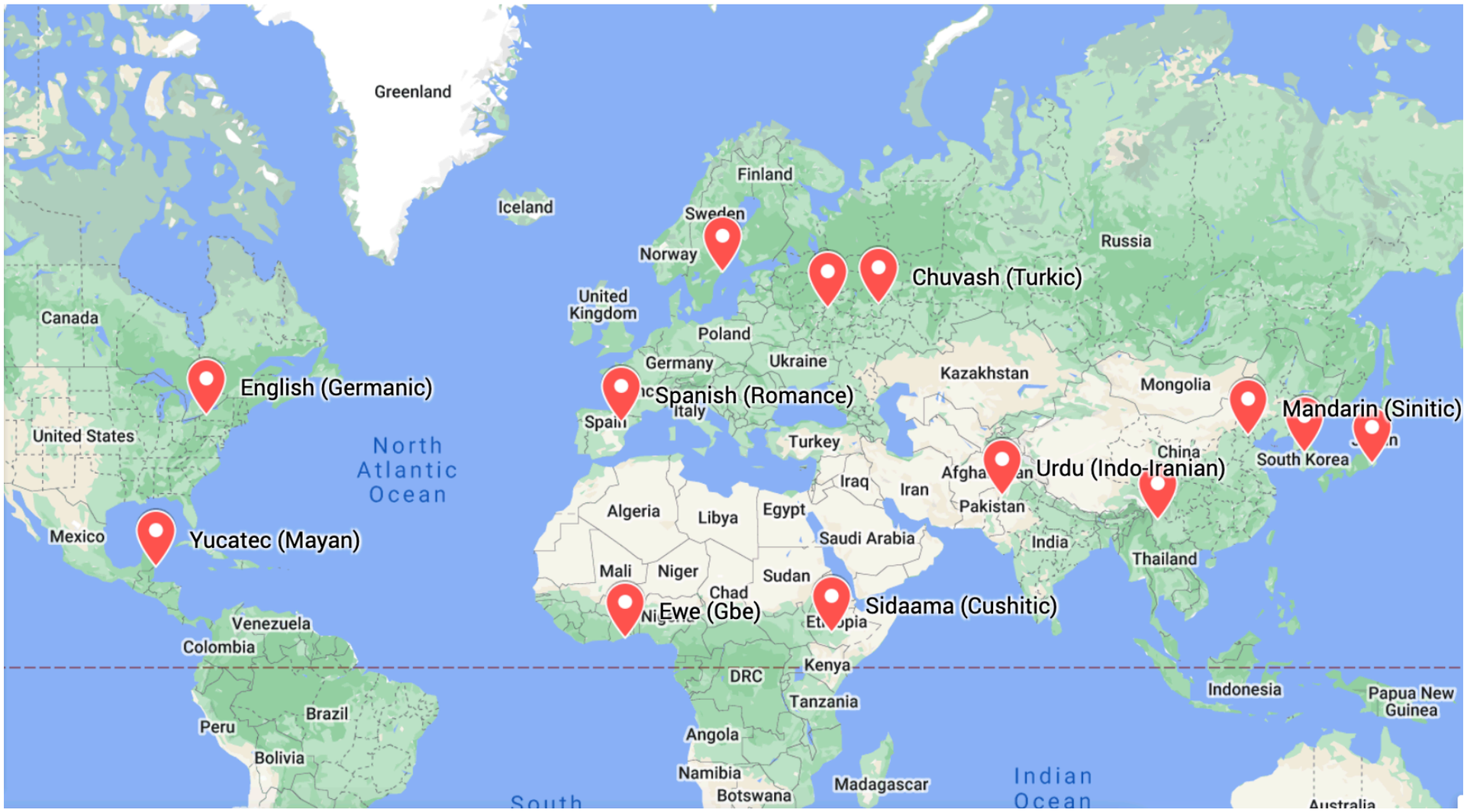
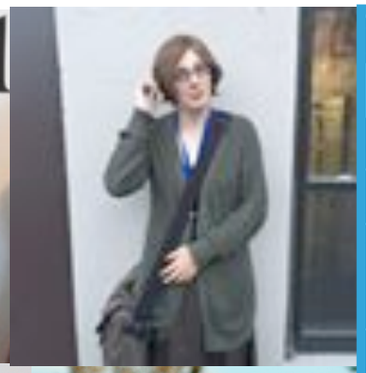



**Figure 4.1.** *The current sample of the CAL Semantic Typology subproject (widgets marking approximate field sites)*

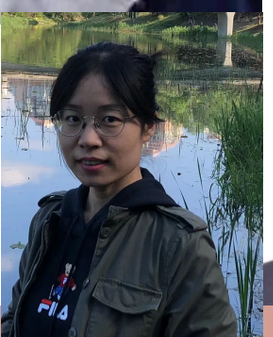
# THE LANGUAGE SAMPLE (CONT.)

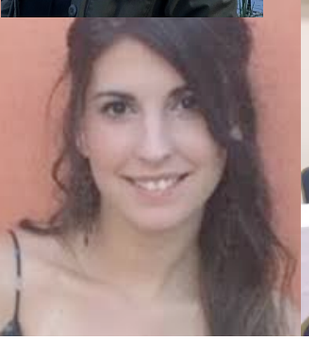
- populations and researchers

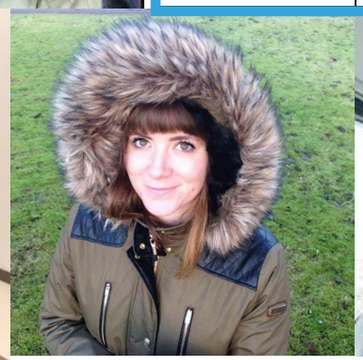
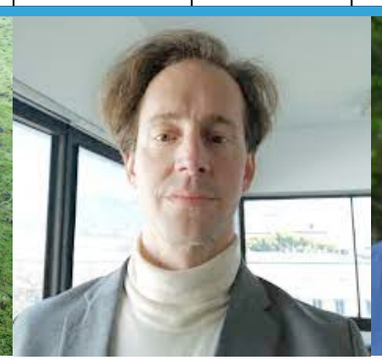




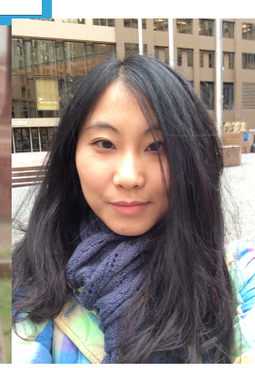








| Language | Genus | Field site | N | Researcher | Affiliation |
|---|---|---|---|---|---|
| Chuvash | Turkic | Russia | 12 | T. Nikitina | CNRS |
| English | Germanic | U.S.A. | 13 | E. Bellingham, S. Evers | U at Buffalo |
| Ewe | Kwa | Ghana/ U.S.A | 12 | J. Essegbey | U of Florida |
| Japanese | Japonic | Japan | 15 | K. Kawachi | Keio U |
| Korean | Isolate | R.O.K. | 12 | S. Park | Kyung Hee U |
| Mandarin | Chinese | China | 12 | J. Du, T. F. Li | UCAS, Beihang U |
| Russian | Slavic | Russia | 12 | A. Stepanova | U at Buffalo |
| Sidaama | Cushitic | Ethiopia | 12 | K. Kawachi | Keio U |
| Spanish | Romance | Spain | 13 | A. Ariño, I. Ibarretxe Antuñano | U of Zaragoza |
| Swedish | Germanic | Sweden | 12 | P. Järnefelt, G. Montero-Melis, E. Bylund | Stockholm U, MPI for Psycholinguistics |
| Urdu | Indic | Pakistan | 12 | S. Hafeez | U at Buffalo |
| Yucatec | Mayan | Mexico | 12 | J. Bohnemeyer | |
| Zauzou | Lolo-Burmese | China | 12 | Y. Li | Wuhan U |

**Table 4.1.** *The current sample of the CAL Semantic Typology subproject*

- causative expressions included in the analysis

**Table 4.2.** *Causative coding devices in the sample languages that were included in the analysis*

| Construction | Chu-vash | Eng-lish | Ewe | Japa-nese | Ko-rean | Man-darin | Rus-sian | Sidaa-ma | Spa-nish | Swe-dish | Urdu | Yuca-tec | Zauzou |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Lexical & not fully productive morphological causatives | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| Light verb constructions | | | | | | | | | | | ✓ | | |
| Serial verb constructions | | | ✓ | | | | | | | | | | |
| Fully productive morphological causatives | ✓ | | | ✓ | | | | | | | ✓ | | |
| Periphrastic causatives | | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| Non-sentential causer adjunct | | ✓ | | | | | | | | | ✓ | | |
| Non-sentential cause adjuncts | | | | | | ✓ | ✓ | ✓ | ✓ | | ✓ | | |
| Clause-layer serialization | | | ✓ | | | | | | | | | | |
| Causal converb constructions | ✓ | | | | ✓ | | | | | | | | ✓ |
| Causal clause constructions | | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Extent ('So X that Y') constructions | | | | | | | ✓ | | ✓ | | | | |
| Means construction | | | | | | | | ✓ | | | | | |

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# THE CLUSTER ANALYSIS

- this and the following analyses are based on data from the 43 core scenes of the CAL Clips
- for each language-specific response type (RT, i.e., causative construction type), a rating vector was calculated
  - one dimension per stimulus clip
  - coordinates represent the proportion of speakers who rated the stimulus description acceptable for the clip
    - i.e., well-formed, accurate, and appropriately informative
  - where multiple descriptions were tested for a given RT, the ratio was incremented if a least one description was rated acceptable

- a cluster analysis was performed over all 60 RT vectors

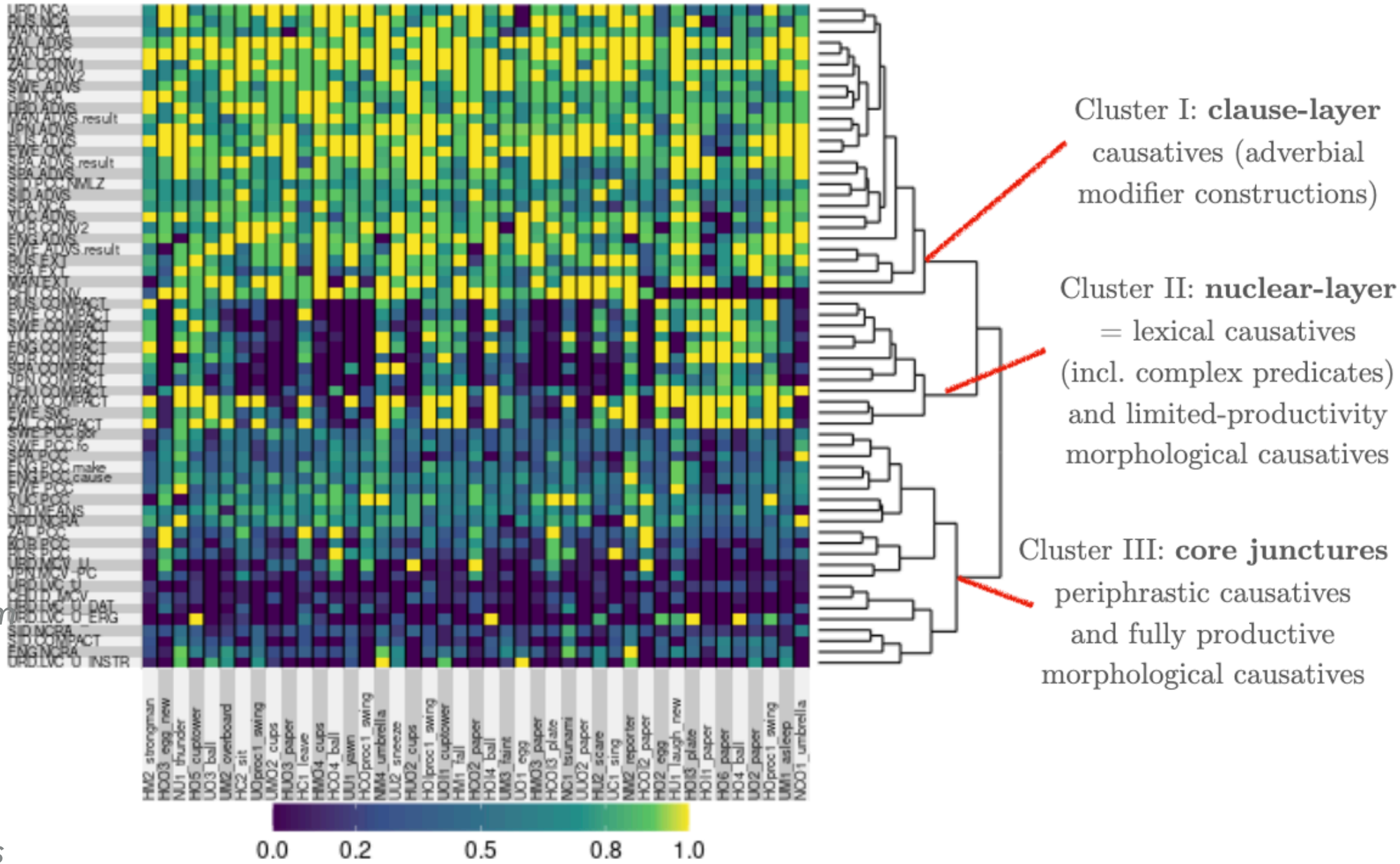



**Figure 5.1.** *Heat map and cluster dendrogram of the rating vectors associated with the 60 language-specific response types (RTs) included in the analysis (x-axis: stimulus clips; y-axis: language-specific RTs)*

- discussion
  - the rating vectors solely reflect the acceptability ratings
    - the model had no access
      to morphosyntactic information
  - remarkably, the model nevertheless was able to group
    - **nuclear-layer** = lexical and
      not fully productive morphological causatives
    - **core junctures:** periphrastic (= analytical/syntactic)
      *and fully productive morphological causatives*
    - **clause-layer** causatives: adverbial modifier constructions
      such as causal clause and converb constructions
  - suggesting that each construction type
    has a unique semantic/pragmatic profile

- discussion: a mismatch
  - fully productive morphological causatives
    such as those of Chuvash, Japanese, and Urdu
    - behave semantically and pragmatically
      like periphrastic causatives in other languages
      - confirming Shibatani (1973)
- to treat these as core junctures in RRG
  - one would have to assume
    that the causative affix introduces its own core
    - *not* merely a *nucleus*,
      as in a complex predicate analysis!
    - problem: this would be a core
      that does not admit its own operators and modifiers

- discussion: more mismatches
  - 'non-sentential causer adverbials' (English, Sidaama, Urdu) pattern with core junctures

(5.1) Non-sentential causer adjuncts: pattern w/ core junctures
*The man knocked over the cups* ***because of the woman***

(5.2) Non-sentential cause adjuncts: pattern w/ clause junctures
*The man knocked over the cups* ***because of the woman's scream***

- it seems plausible that the PP in (5.1)
  - is a core modifier
  - whereas the one in (5.2) is a clausal modifier

- discussion: more mismatches (cont.)
  - light verb constructions in Urdu and “compact” causatives in Sidaama pattern with core junctures
    - compact causatives = lexical or not-fully-productive morphological causatives
      - likely explanation:
        low acceptability across the board
- Mandarin periphrastic causatives pattern with clause-layer junctures
  - likely explanation:
    high acceptability across the board

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# FURTHER ANALYSES

- so each type of causative construction has a distinctive semantic profile depending on its juncture level
- but what exactly *are* these unique profiles?
- answering this question
  requires the application of predictive models
  - which are able to discern
    the effects of the various semantic predictor variables

- not all semantic predictor variable level combinations could be instantiated with equal frequency in the CAL Clips
- so to discover the effects of the predictor variables, we used machine learning classifiers instead of regression models
- all nuclear-layer causatives except for the fully productive morphological causatives showed a single rating maximum involving
  - absence of mediation
    (no intervening subevents or participants)
  - affectees/patients with no control over the caused event
  - intentional causers
- as predicted by the literature

- example: English

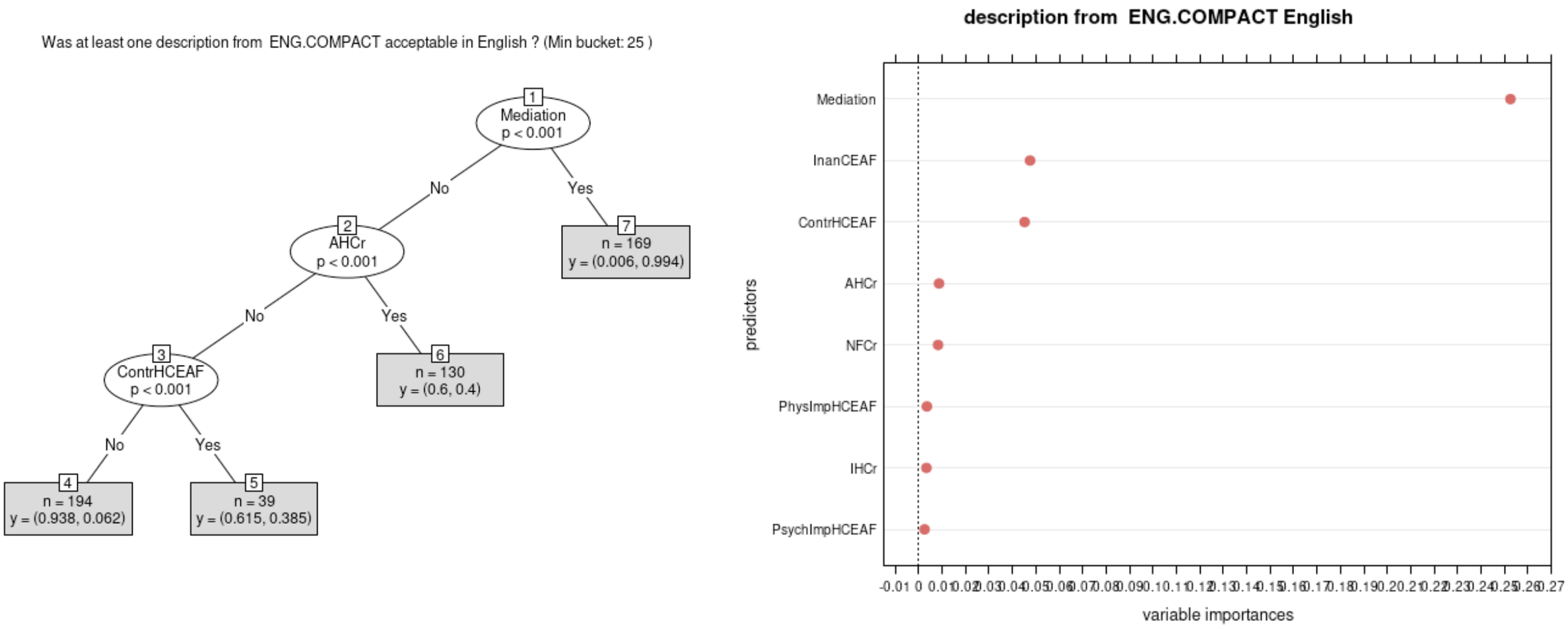



**Figure 6.1.** *Conditional inference tree and variable importance plot based on a random forest model of the English 'compact' causative construction (i.e., base-transitive causative verbs). (AHCr - Accidental human causer; ContrHCEAF - Causee/affectee with control over the caused event; InanCEAF - Inanimate causee/affectee; NFCr - Natural force causer; PhysImpHCEAF - Physically impacted causee/affectee; IHCr - Intentional human causer; PsychImpHCEAF - Psychologically impacted causee/affectee)*

- mediation proved generally the top variable for nuclear-layer causatives
- one exception: ergative-marked causer NPs entail or implicate intentionality with nuclear Urdu causatives

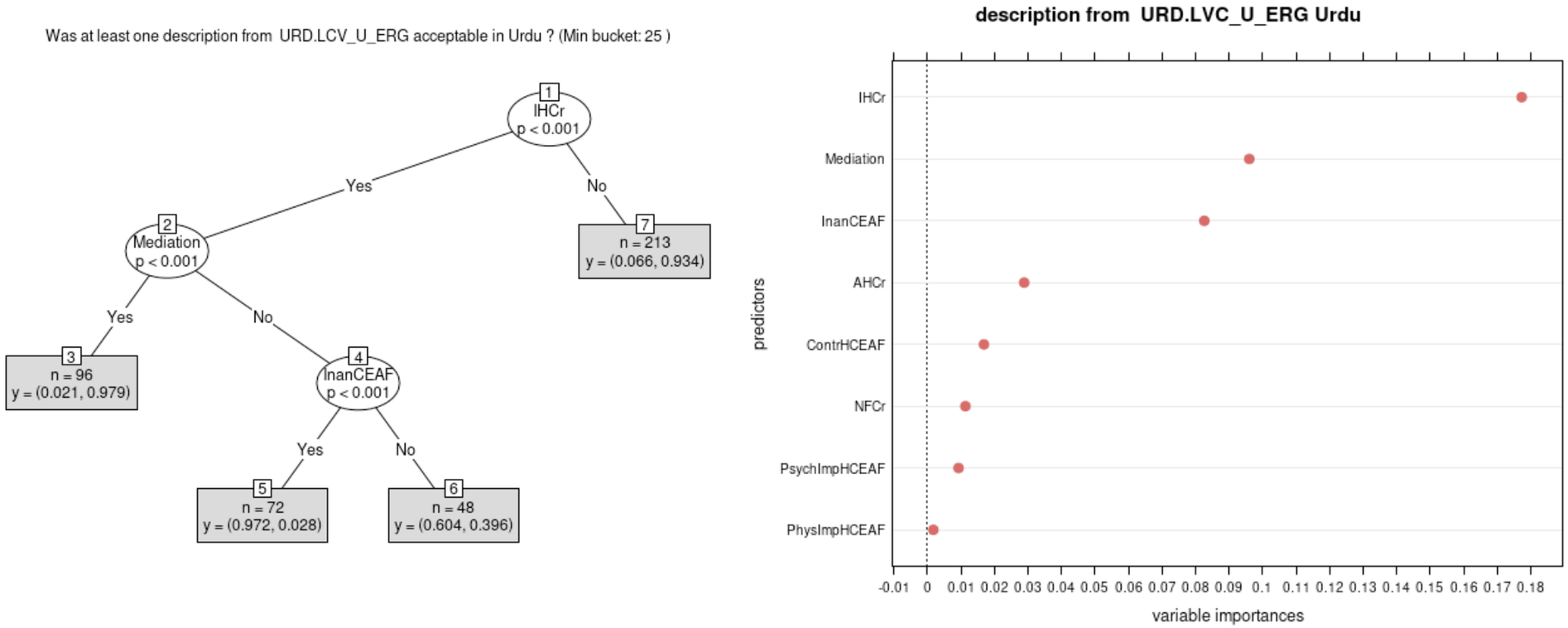



**Figure 6.2.** *Conditional inference tree and variable importance plot based on a random forest model of the Urdu light verb construction with ergative causer NP. (IHCr - Intentional human causer; InanCEAF - Inanimate causee/affectee; AHCr - Accidental human causer; ContrHCEAF - Causee/affectee with control over the caused event; NFCr - Natural force causer; PhysImpHCEAF - Physically impacted causee/affectee; PsychImpHCEAF - Psychologically impacted causee/affectee)*

## FURTHER ANALYSES (CONT.)

- the surprise: the semantic prototypes of complex causatives aren't simply complementary to those of compact causatives
  - core-layer causatives in particular often show multiple discrete prototypes, one of which involves natural forces
    - example: Zauzou (Loloish, Yunan Province, PRC)

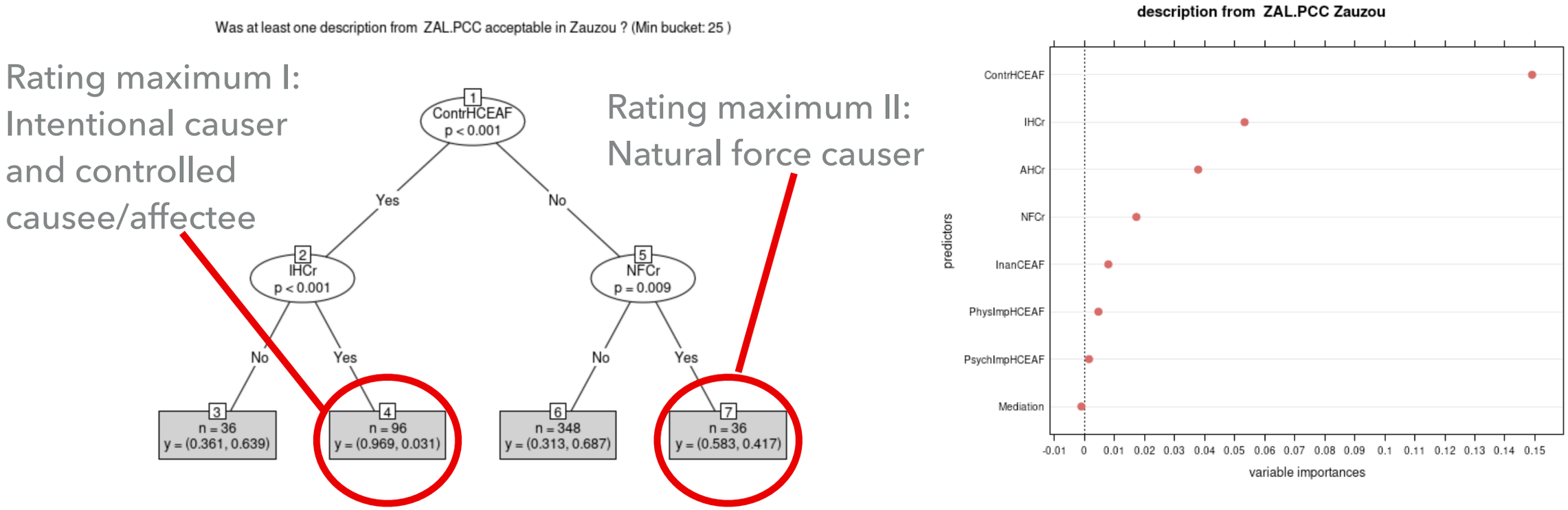



**Figure 6.3.** *Conditional inference tree and variable importance plot based on a random forest model of the Zauzou periphrastic causative construction. (ContrHCEAF - Causee/affectee with control over the caused event; IHCr - Intentional human causer; NFCr - Natural force causer; AHCr - Accidental human causer; InanCEAF - Inanimate causee/affectee; PhysImpHCEAF - Physically impacted causee/affectee; PsychImpHCEAF - Psychologically impacted causee/affectee)*

- overall, of 11 periphrastic causative constructions
  - 6 show evidence of multiple prototypes
  - 7 show evidence of natural force causer prototypes
  - in contrast, the fully productive morphological causatives of Japanese and Urdu show a single prototype
    - involving mediation and intentional causers
- the fully-productive morphological causative of Chuvash elicited low acceptability across the board

# FURTHER ANALYSES (CONT.)

- to assess inter-speaker variation, we computed separate rating vectors for each participant and response type
  - and generated multi-dimensional scaling plots of their Hamming distances by language

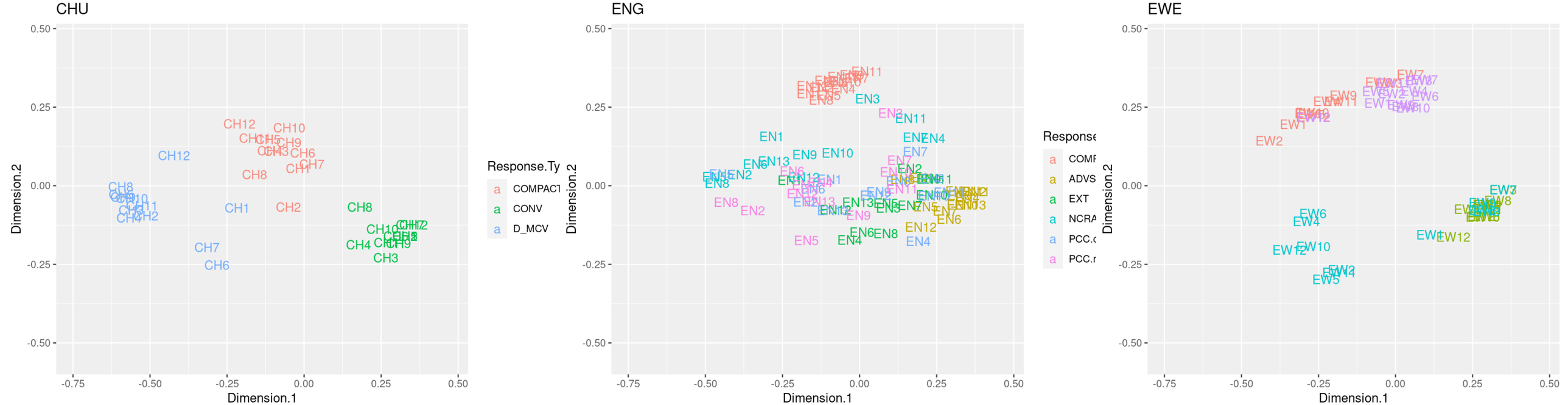



**Figure 6.4.** *Plotting the first two dimensions of a multi-dimensional scaling model of the rating vectors by participant and response type for Chuvash, English, and Ewe*

- in every language,
  inter-speaker variation is minimal with nuclear causatives
  - except for the fully productive morphological ones
  - and maximal with core-layer
    and fully productive morphological causatives


**Figure 6.5.** *Plotting the first two dimensions of a multi-dimensional scaling model of the rating vectors by participant and response type for Japanese, Korean, and Mandarin*

- discussion
  - inter-speaker agreement with nuclear-layer causatives is consistent with them having unique prototypes

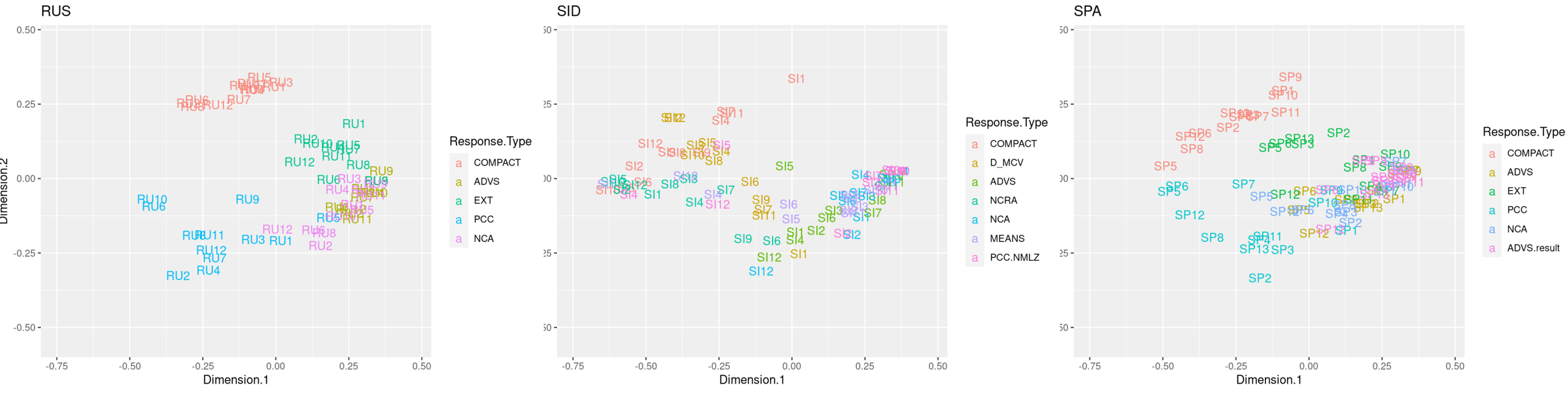



**Figure 6.6.** *Plotting the first two dimensions of a multi-dimensional scaling model of the rating vectors by participant and response type for Russian, Sidaama, and Spanish*

# FURTHER ANALYSES (CONT.)

- discussion (cont.)
  - clause-layer causatives show relatively high acceptability across the board

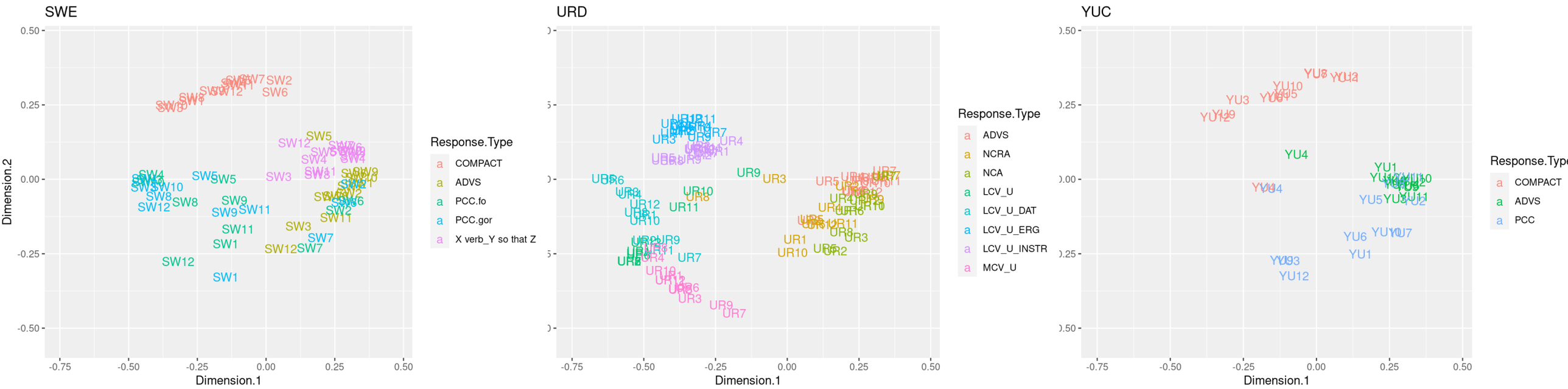



**Figure 6.7.** *Plotting the first two dimensions of a multi-dimensional scaling model of the rating vectors by participant and response type for Swedish, Urdu, and Yucatec*

- discussion (cont.)
  - intermediate-complexity constructions are "caught in the middle"
    - lacking both unique semantic prototypes and across-the-board acceptability

ZAL

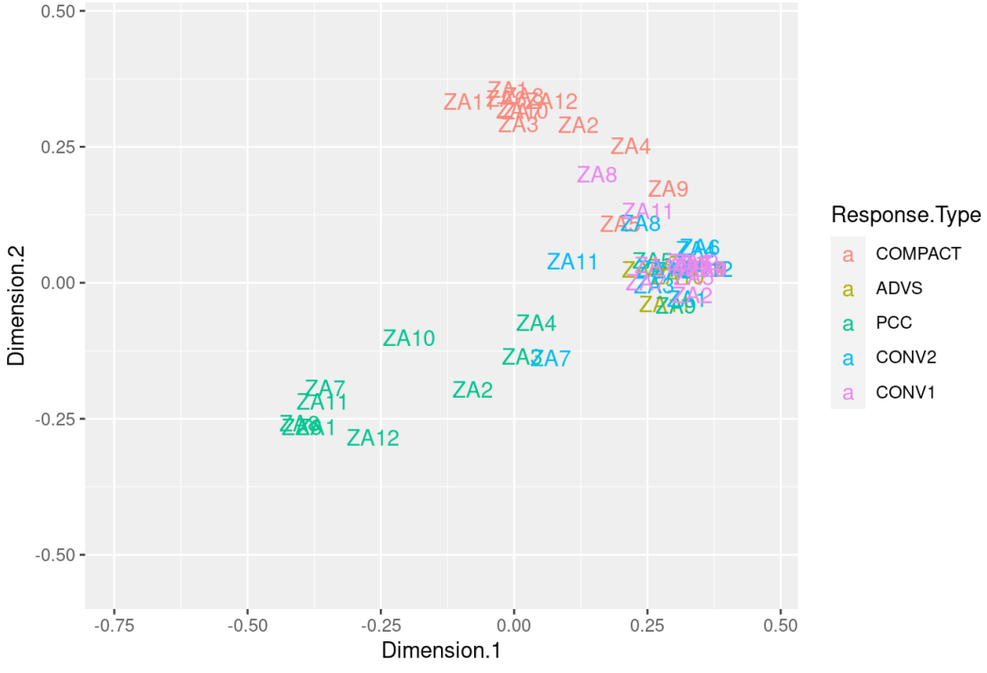



**Figure 6.8.** *Plotting the first two dimensions of a multi-dimensional scaling model of the rating vectors by participant and response type for Zauzou*

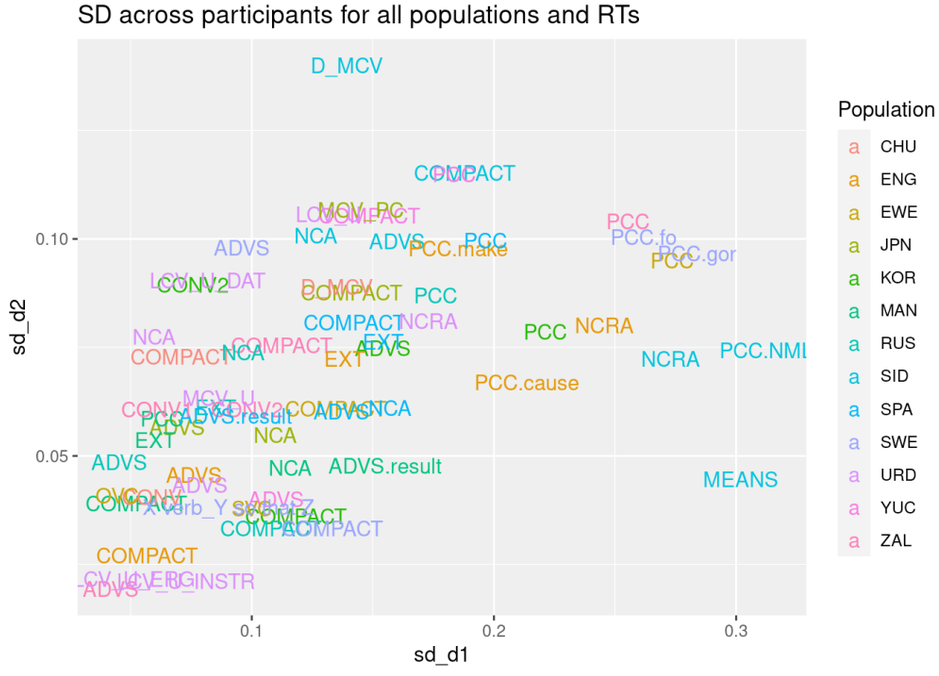



**Figure 6.9.** *Plotting the standard deviation of the first and second dimension of a multi-dimensional scaling model of the rating vectors by response type (labels) and language (colors)*

# SYNOPSIS

- Testing assumptions about the syntax-semantics interface
- A new study design for semantic typology
- Variables and stimuli: the CAL Clips
- The language sample
- The cluster analysis
- Further analyses
- Implications for RRG

# IMPLICATIONS FOR RRG

- nuclear-layer causatives, core-layer causatives, and clause-layer causatives
  - are each associated with a unique semantic profile across languages
- this supports a core tenet of RRG:
  - the view that the LSC is part of the syntax-semantics interface and its layers are semantically "bootstrapped"
    - i.e. defined in terms of the communicative functions they (prototypically?) serve to perform
      - cf. also Bohnemeyer & Van Valin (2017)

- the profiles

**Table 7.1.** *Summary of the semantic profiles associated with causatives at different LSC layers*

| LSC layer/ juncture level | Construction types | Semantic/pragmatic profile | Mismatches |
|---|---|---|---|
| Nuclear | Lexical and morphological causatives incl. complex predicates (e.g., LVCs) | Unique semantic prototype: direct, unmediated causation (Urdu: intentional causation) | Fully productive morphological causatives pattern with core junctures; same for low-frequency LVCs (Urdu) and low-acceptability lexical/ morphological causatives (Sidaama) |
| Core | Periphrastic causatives; non-sentential causer adjuncts | Diffuse – multiple prototypes; common denominator: atypical causation (Lakoff & Johnson 1980: 69-76; Haspelmath 2008) | Mandarin periphrastic causatives pattern with clause-layer junctures |
| Clause | Clause-layer modifiers; extent and means constructions | Across-the-board acceptability due to the flexibility of expressing causality between arbitrary pairs of event descriptions | N/A |

- mismatches
  - fully productive morphological causatives pattern with periphrastic causatives
    - in that they can express indirect causation
  - one could accommodate this finding by positing a separate core introduced by the causative morpheme
- perhaps more plausibly, one might attribute the mismatch to a limitation of the test
  - it does not seem "reasonable" to expect the layers to be semantically uniform *in every respect*
    - but this in turn does perhaps support a prototype view of their semantics?

- further mismatches
  - a few individual language-specific constructions pattern semantically against their morphosyntactic analysis
    - this kind of variation can be attributed to semantic/pragmatic "ecology" of the causative domain
      - in each individual language
        - e.g., to frequency effects
- all of these mismatches can be viewed as limitations of the method we used
  - or alternatively as support for a prototype view of the semantics of the LSC layers

# ACKNOWLEDGMENTS


- epic thanks to
  - the study participants
  - colleagues who have provided advice:
    Dare Baldwin; Dedre Gentner; Beth Levin; Gail Mauner;
    Eric Pederson; Robert D. Van Valin, Jr., Phillip Wolff
    - all of whom shall be held blameless for any foolish and harebrained claims in this presentation
- our sponsor
  - the material presented here is based upon work supported by the National Science Foundation
    under Grant No. BCS153846 and BCS-1644657,
    'Causality Across Languages'; PI J. Bohnemeyer.

# Thanks!

# REFERENCES


Bohnemeyer, J., N.J. Enfield, J. Essegbey, & S. Kita. (2010). The Macro-Event Property: The segmentation of causal chains. In *Event representation in language: Encoding events at the language-cognition interface*, eds. Jürgen Bohnemeyer and Eric Pederson, 43–67. Cambridge: Cambridge University Press.

Bohnemeyer, J. & R. D. Van Valin Jr. (2017). The Macro-Event Property and the Layered Structure of the Clause. *Studies in Language* 41(1): 142-197.

Breiman, L. (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.

Breiman, L. (2001). Random forests. *Machine Learning* 45(1): 5-32.

Comrie, B. (1981). Language universals and linguistic typology: Syntax and morphology. Chicago: University of Chicago Press.

Dixon, R.M. (2000). A typology of causatives: form, syntax and meaning. In *Changing valency: Case studies in transitivity*, eds. Robert M. W. Dixon and Alexandra Y. Aikhenvald, 30--83. Cambridge: Cambridge University Press.

Escamilla Jr, R.M. (2012). *An updated typology of causative constructions: Form-function mappings in Hupa (Californian Athabaskan), Chungli Ao (Tibeto-Burman) and Beyond*. PhD Dissertation, University of California, Berkeley.

Haiman, J. (1983). Iconic and economic motivation. *Language* 59(4):781–819.

Haspelmath, M. (2008). Frequency vs. iconicity in explaining grammatical asymmetries. *Cognitive Linguistics* 19(1): 1-33.

Kemmer, S. & A. Verhagen. (1994). The grammar of causatives and the conceptual structure of events. *Cognitive Linguistics* 5(2):115–156.

Lakoff, G. & M. Johnson. (2033 [1980]). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

Levshina, N. (2015). European analytic causatives as a comparative concept: Evidence from a parallel corpus of film subtitles. *Folia Linguistica* 49(2): 487–520.

Levshina, N. (2016). Why we need a token-based typology: A case study of analytic and lexical causatives in fifteen European languages. *Folia Linguistica* 50(2): 507–542.

Levshina, N. (2017). Measuring iconicity: A quantitative study of lexical and analytic causatives in British English. *Functions of Language* 24(3): 319–347.

Levshina, N. (2022). Semantic maps of causation: New hybrid approaches based on corpora and grammar descriptions. *Zeitschrift für Sprachwissenschaft* 41(1): 179-205.

## REFERENCES (CONT.)


McCawley, J. (1976). Remarks on what can cause what. In *Syntax and Semantics VI: The grammar of causative constructions*, ed. Masayoshi Shibatani, 117–129. New York, NY: Academic Press.

McCawley, J. (1978). Conversational implicature and the lexicon. In *Syntax and semantics IX: Pragmatics*, ed. Peter Cole, 245-258. New York, NY: Academic Press.

Shibatani, M. (ed.) (1976). *The grammar of causative constructions*. New York: Academic Press (Syntax and Semantics; 6).

Shibatani, M. & P. Pardeshi. (2002). The causative continuum. In *The grammar of causation and interpersonal manipulation*, ed. Masayoshi Shibatani, 85–126. Amsterdam: Benjamins.

Talmy, L. (1976). Semantic causative types. In Masayoshi Shibatani (ed.), *Syntax and semantics, vol. 6: The grammar of causative constructions*, 43-116. New York: Academic Press.

Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science* 12:49-100.

Verhagen, A. & S. Kemmer. (1997). Interaction and causation: Causative constructions in modern standard Dutch. *Journal of Pragmatics* 27:61–82.

Van Valin, R. D. Jr. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Van Valin, R. D. Jr. (2023). Principles of Role and Reference Grammar. In D. Bentley, R. Meiral Usón, W. Nakamura, & R. D. Van Valin Jr. (eds.), *The Cambridge handbook of Role and Reference Grammar*. Cambridge: Cambridge University Press. 17-177.

Wolff, P. 2003. Direct causation in the linguistic coding and individuation of causal events. *Cognition* 88(1): 1–48.