Lessons from incremental development: 20 years of the Grammar Matrix

Emily M. Bender
University of Washington

HPSG 2023
UMass Amherst
July 7, 2023

UNIVERSITY of WASHINGTON
DEPARTMENT OF LINGUISTICS

UWNLP

DELPH-IN

Key points

- Grammar engineering allows us to handle much more complexity than pen-and-paper syntax
  - Complexity in the size of datasets, range of phenomena considered, range of languages
- Empirical & theoretical work (descriptive and documentary linguistics, typology, syntax) are necessary inputs
- The Grammar Matrix represents 20 years (and counting!) of cumulative development

This talk draws on

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

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

- Bender, Emily M., Dan Flickinger and Stephan Oepen. 2002. The Grammar Matrix: An Open-Source Starter-Kit for the Rapid Development of Cross-Linguistically Consistent Broad-Coverage Precision Grammars. Carroll, John, Nelleke Oostdijk, and Richard Sutcliffe, eds. *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*. Taipei, Taiwan. pp. 8-14.

Funding acknowledgment

- Grammar Matrix development has been supported by the US National Science Foundation under grants No BCS-0644097, BCS- 1160274, and BCS-561833.

Overview

- Grammar Engineering
- Grammar Matrix: Early history
- Growth through student projects
- Pros (and cons) of operationalizing analyses
- Reflections on hypothesis testing

Grammar Engineering

- The development of grammars-in-software: morphology, syntax, semantics
- “Precision grammars”
  - Encode linguistic analyses
  - Human- and machine-readable
  - Model grammaticality
  - Map strings to underlying representations
  - Can be used for both *parsing* and *generation*

Grammar Engineering

- A discipline and a methodology concerned with a particular empirical approach (Bierwisch 1963; Zwicky et al. 1965; Müller 1999; Butt et al. 1999; Bender 2008; Müller 2015)

- *Modeling*: Building implemented sets of grammar entities that can be used to accept or reject strings and assign structures to the accepted ones

- *Testing*: Deploying the implemented grammar over sets of examples to measure its domain of applicability

Grammar Engineering: Frameworks

- Precision grammars have been built by/in/with
  - HPSG in ALE/Controll (Götz & Meurers 1997; CoreGram: Müller 2015)
  - LFG (ParGram: Butt et al 2002)
  - F/XTAG (Doran et al 1994)
  - SFG (Bateman 1997)
  - GF (Ranta 2007)
  - OpenCCG (Baldridge et al 2007)
  - Proprietary formalisms and Microsoft and Boeing and IBM
- On implementation of MP, see e.g. Stabler 2001, Fong 2015, Herring 2016, Torr et al 2019

DELPH-IN: Deep Linguistic Processing in HPSG Initiative (www.delph-in.net)

- Informal, international consortium established in 2002
- Shared repository of open-source, interoperable resources
- Framework/formalisms:
    - Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994)
    - Minimal Recursion Semantics (MRS; Copestake et al 2005)
    - DELPH-IN joint reference formalism (Copestake 2002a)

DELPH-IN: Deep Linguistic Processing in HPSG Initiative (www.delph-in.net)

- **Grammars**: ERG (Flickinger 2000, 2011); GG (Müller & Kasper 2000, Crysmann 2003); Jacy (Siegel, Bender & Bond 2016); SRG (Marimon 2010); gCLIMB (Fokkens 2014); Indra (Moejadi 2018); ...
- **Parsing & Generation**: LKB (Copestake 2002b); PET (Callmeier 2002); ACE (http://sweaglesw.org/linguistics/ace); Agree (Slayden 2012)
- **Regression testing**: [incr tsdb()] (Oepen 2001)
- **Treebanking**: Redwoods (Oepen et al 2004), FFTB (Packard 2015)
- **Applications**: e.g., MT (Oepen et al 2007), QA from structured knowledge sources (Frank et al 2007), Textual entailment (Bergmair 2008), ontology construction (Nichols et al 2006) and grammar checking (Suppes et al 2012), robot control language (Packard 2014), sentiment analysis (Kramer & Gordon 2014), ...

English Resource Grammar (Flickinger 2000, 2011)
demo: https://delph-in.github.io/delphin-viz/demo/

- Under continuous development since 1993
- Broad-coverage: 85-95% on varied domains: newspaper text, Wikipedia, bio-medial research literature (Flickinger et al 2010, 2012; Adolphs et al 2008)
  - Robust processing techniques enable 100% coverage
- Output: derivation trees paired with meaning representations in the Minimal Recursion Semantics framework---English Resource Semantics (ERS)
  - Emerging documentation at https://delph-in.github.io/docs/erg/ErgSemantics/

English Resource Grammar
erg.delph-in.net

- Feb 2023 trunk: 295 syntactic rules, 101 lexical rules, 1268 leaf lexical types
- Generalizations captured in a type hierarchy
- Both ‘core’ (high frequency) and ‘peripheral’ constructions

```
head_subj_phrase := basic_head_subj_phrase &
  [ HD-DTR.SYNSEM.LOCAL.CAT.VAL.SUBJ < #synsem >,
    NH-DTR.SYNSEM #synsem ].
```

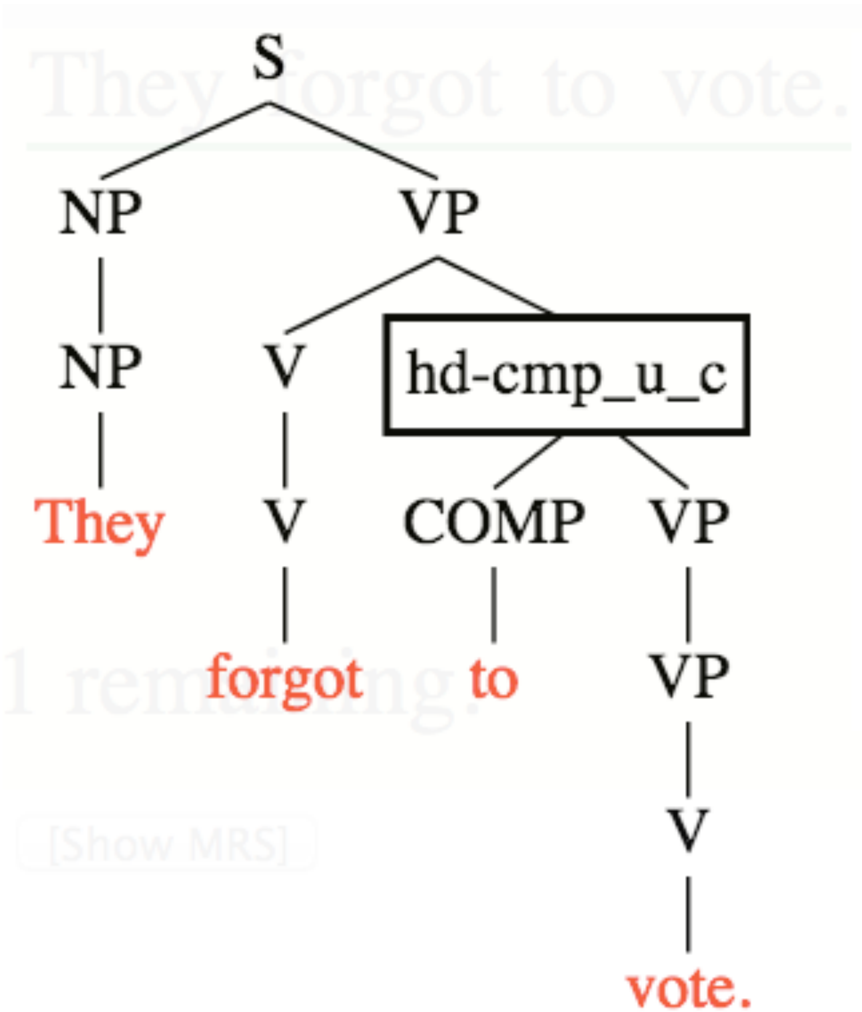
English Resource Grammar
erg.delph-in.net

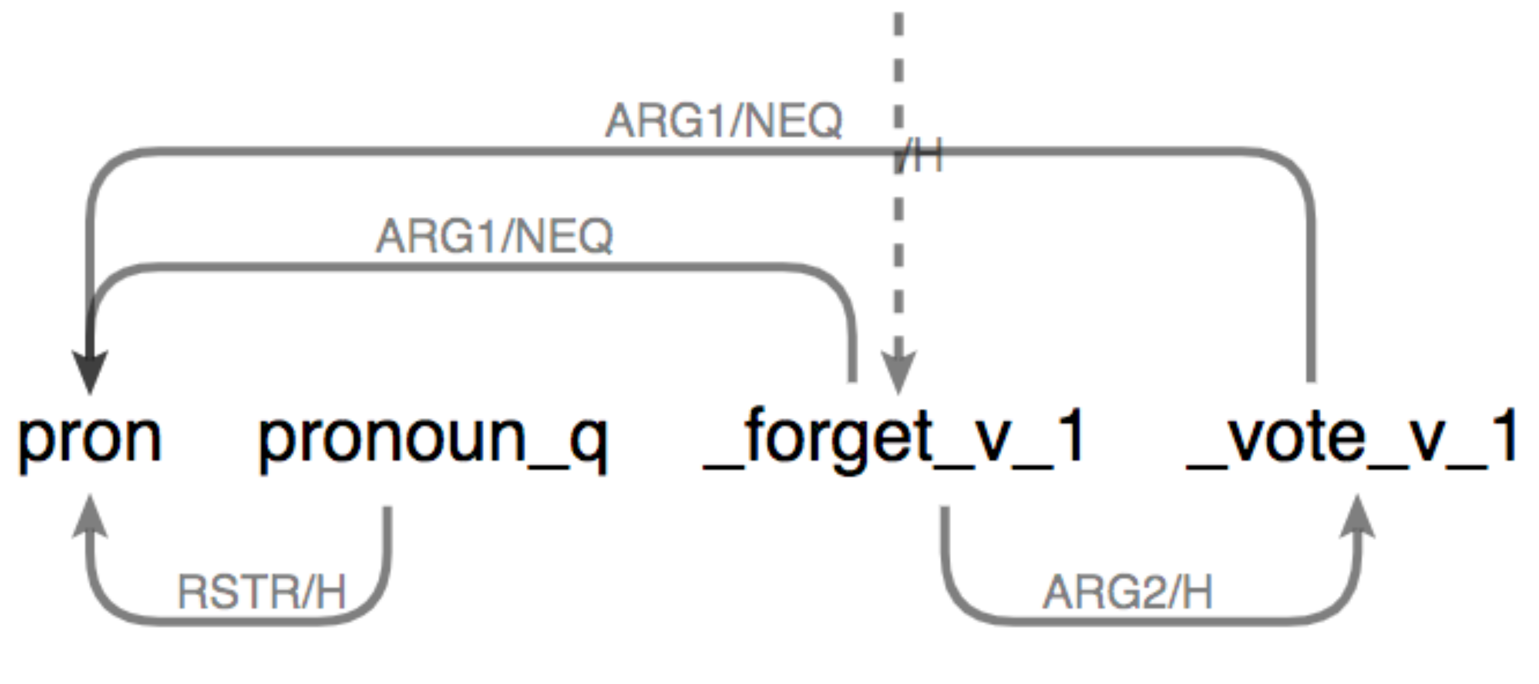
```
modgap_rel_cl := basic_non_wh_rel_cl &
  [ SYNSEM.LOCAL.CAT.HEAD.MOD < [ LOCAL.CAT.HEAD noun,
                                  --MIN modable_rel,
                                  --SIND #mind ] >,
    ARGS < [ SYNSEM
            [ LOCAL.CONT.HOOK.INDEX.SF prop,
              NONLOC.SLASH 1-dlist &
              [ LIST < mod-local &
                       [ CAT.HEAD mobile & [ MOD < synsem > ],
                         CONT.HOOK [ LTOP #sltop,
                                     INDEX #slind & [ SORT location ],
                                     XARG #xarg ] ] > ] ] ] >,
    ORTH [ FROM #from, TO #to ],
    C-CONT.RELS <! prep_relation &
                 [ LBL #sltop,
                   PRED loc_nonsp_rel,
                   ARG0 #slind & [ E [ TENSE no_tense,
                                       ASPECT no_aspect ] ],
                   ARG1 #xarg & event_or_index,
                   ARG2 #mind & [ SORT basic-entity-or-event ],
                   CFROM #from, CTO #to ] !> ].
```

English Resource Grammar
erg.delph-in.net

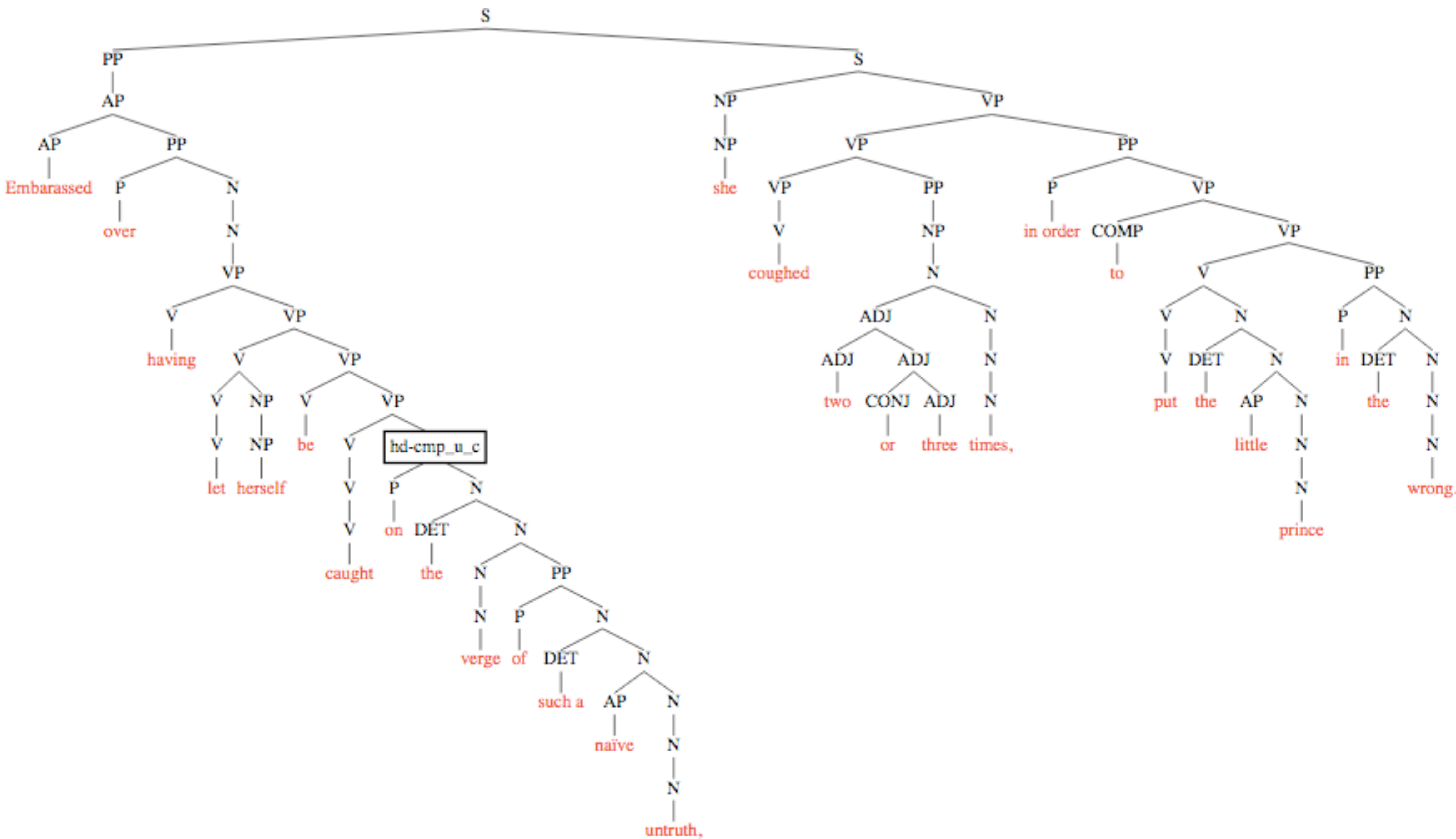
```
basic_head_subj_phrase := head_nexus_rel_phrase & head_final_infl & phrasal &
  [ SYNSEM [ LOCAL [ CAT.VAL [ COMPS < >,
                               SPR < >,
                               SUBJ *olist* & < anti_synsem_min >,
                               SPEC #spec,
                               SPCMPS < > ],
                     CONJ cnil ],
             MODIFD.RPERIPH #rperiph,
             PUNCT.PNCTPR #ppair ],
    HD-DTR.SYNSEM [ LOCAL.CAT [ VAL [ COMPS < >,
                                      SPR *olist*,
                                      SPEC #spec ],
                                MC na ],
                    MODIFD.RPERIPH #rperiph,
                    PUNCT [ LPUNCT pair_or_no_punct,
                            PNCTPR #ppair ] ],
    NH-DTR.SYNSEM canonical_synsem &
                 [ LOCAL [ CAT [ HEAD subst,
                                 VAL [ SUBJ *olist_or_prolist*,
                                       COMPS < >,
                                       SPR *olist* ] ] ],
                   NONLOC [ SLASH 0-dlist,
                            REL 0-dlist ],
                   PUNCT [ LPUNCT pair_or_no_punct,
                           RPUNCT comma_or_rbc_or_pair_or_no_punct,
                           PNCTPR ppair ] ] ].
```

ERG: Examples





ERG: Examples



ERG: Examples

INDEX: e2
RELS:
h1:subord_rel(ARG0: e4,ARG1: h5,ARG2: h6)
h7:"_embarassed/JJ_u_unknown_rel"(ARG0: e8,ARG1: i9)
h7:_over_p_rel(ARG0: e10,ARG1: e8,ARG2: x11)
h12:udef_q_rel(ARG0: x11,RSTR: h13,BODY: h14)
h15:nominalization_rel(ARG0: x11,ARG1: h16)
h16:"_let_v_1_rel"(ARG0: e17,ARG1: i18,ARG2: h19)
h20:pron_rel(ARG0: x21)
h22:pronoun_q_rel(ARG0: x21,RSTR: h23,BODY: h24)
h25:"_catch_v_1_rel"(ARG0: e26,ARG1: i27,ARG2: x21,ARG3: h28)
h25:parg_d_rel(ARG0: e29,ARG1: e26,ARG2: x21)
h30:_on_p_rel(ARG0: e31,ARG1: x21,ARG2: x32)
h33:_the_q_rel(ARG0: x32,RSTR: h34,BODY: h35)
h36:"_verge_n_1_rel"(ARG0: x32)
h36:_of_p_rel(ARG0: e37,ARG1: x32,ARG2: x38)
h39:_such+a_q_rel(ARG0: x38,RSTR: h40,BODY: h41)
h42:"_naïve/JJ_u_unknown_rel"(ARG0: e43,ARG1: x38)
h42:"_untruth_n_1_rel"(ARG0: x38)
h44:pron_rel(ARG0: x3)
h45:pronoun_q_rel(ARG0: x3,RSTR: h46,BODY: h47)
h48:"_cough_v_1_rel"(ARG0: e2,ARG1: x3)
h48:loc_nonsp_rel(ARG0: e49,ARG1: e2,ARG2: x50)
h51:udef_q_rel(ARG0: x50,RSTR: h52,BODY: h53)
h54:card_rel(CARG: "2",ARG0: e56,ARG1: x50)
h57:_or_c_rel(ARG0: e58,L-INDEX: e56,R-INDEX: e59,L-HNDL: h54,R-HNDL: h60)
h60:card_rel(CARG: "3",ARG0: e59,ARG1: x50)
h57:"_times_n_1_rel"(ARG0: x50)
h62:"_in+order+to_x_rel"(ARG0: e63,ARG1: h64,ARG2: h65)
h66:"_put_v_1_rel"(ARG0: e67,ARG1: x3,ARG2: x68,ARG3: h69)
h70:_the_q_rel(ARG0: x68,RSTR: h71,BODY: h72)
h73:"_little_a_1_rel"(ARG0: e74,ARG1: x68)
h73:"_prince_n_of_rel"(ARG0: x68,ARG1: i75)
h76:_in_p_rel(ARG0: e77,ARG1: x68,ARG2: x78)

LinGO Grammar Matrix: Motivations and early history

- Speed up grammar development
  - Initial context: Project DeepThought
  - Leverage resources from resource-rich language to enhance NLP for resource-poor languages
  - Claim: Some of what was learned in ERG development is not English-specific
- Interoperability: a family of grammars compatible with the same downstream processing tools

Grammar Matrix:
Motivations and early history

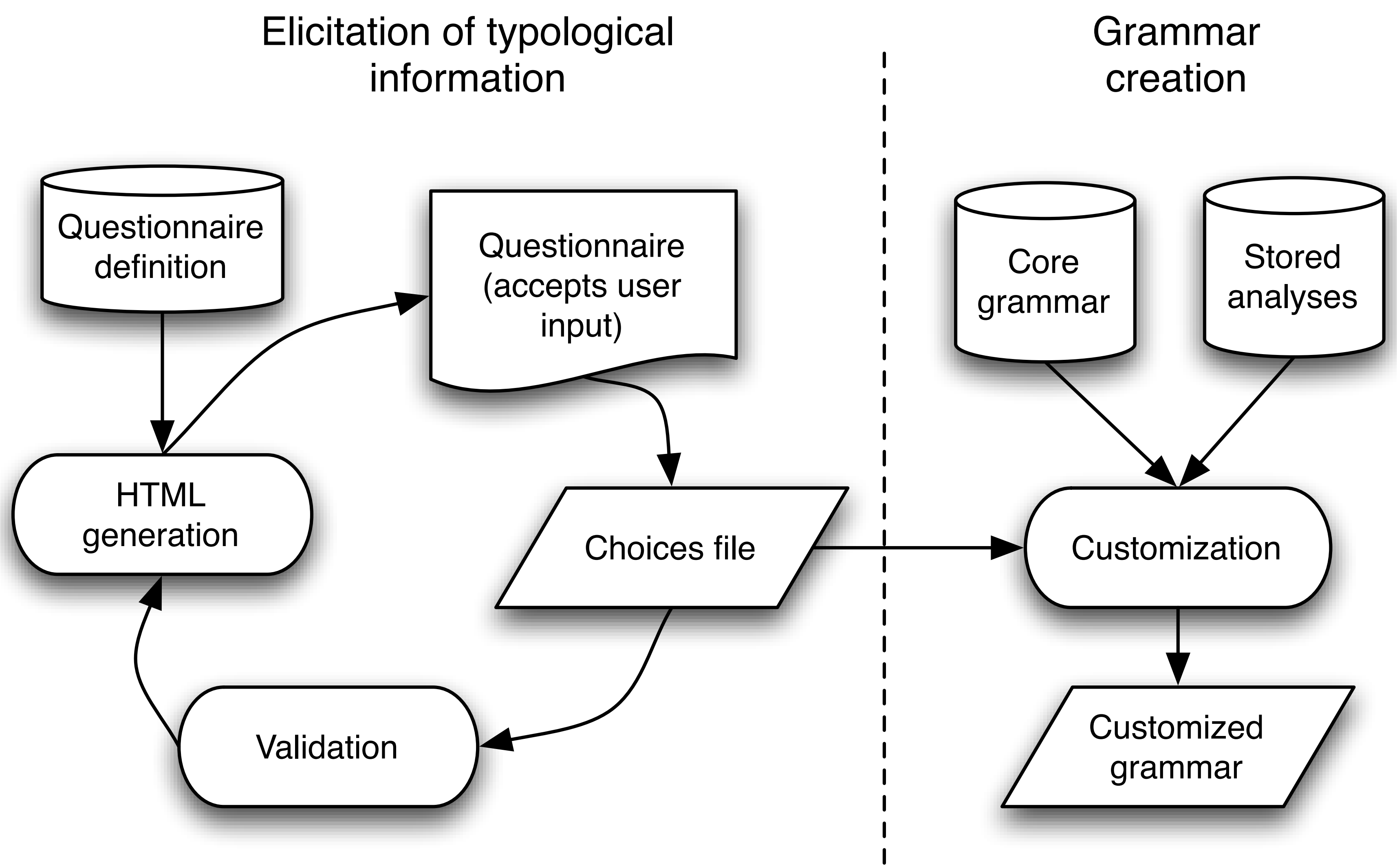
- With reference to Jacy (Siegel et al 2016), strip everything from ERG (Flickinger 2000, 2011) which looks English-specific
- Resulting “core grammar” doesn’t parse or generate anything, but supports quick start-up for scaleable resources (Bender et al 2002)
- Used in the development of grammars for Norwegian (Hellan & Haugereid 2003), Modern Greek (Kordoni & Neu 2005), Spanish (Marimon 2010) and Italian
- Used as the basis of multilingual grammar engineering course at UW (Ling 567): 133 languages since 2004

Grammar customization: Motivations

- The Grammar Matrix core grammar is not itself a functioning grammar fragment
  - can't be directly tested
- Human languages vary along many dimensions, but not infinitely
- Can be seen as solving many of the same problems in different ways
- Many phenomena are "widespread, but not universal" (Drellishak, 2009)
  - we can do more than refining the core
- Also, grammar engineering lab instructions started getting mechanistic

LinGO Grammar Matrix Customization System

(Bender & Flickinger 2005, Drellishak 2009, Bender et al 2010, Zamareva et al 2022)



http://www.delph-in.net/matrix/customize/matrix.cgi

- ? General Information
- Word Order
- Number
- Person
- Gender
- Case
- Adnominal Possession
- Direct-inverse
- Tense, Aspect and Mood
- Evidentials
- Other Features
- Sentential Negation
- Coordination
- Matrix Yes/No Questions
- Constituent (wh-) Questions
- Information Structure
- Argument Optionality
- Nominalized Clauses
- Clausal Complements
- Clausal Modifiers
- Lexicon
- Morphology
- Import Toolbox Lexicon
- Test Sentences
- Test by Generation Options

Archive type: (•) .tar.gz ( ) .zip

[Create Grammar] [Test by Generation]

▼ neg-prefix (verb-pc1)

× **Verb Position Class 1:**

Position Class Name: neg-prefix

Obligatorily occurs: ☐

Appears as a prefix or suffix: prefix

Possible inputs: endoclitics (verb-pc44) ▼

Morphotactic Constraints:

Add a Require constraint

Add a Forbid constraint

Lexical Rule Types that appear in this Position Class:

▶ neg (verb-pc1_lrt1)

▼ finite-neg (verb-pc1_lrt2)

× **Lexical Rule Type 2:**

Name: finite-neg

Supertypes: neg (verb-pc1_lrt1) ▼

Features:

× Name: form Value: finite ▼ Specified on: the verb

Add a Feature

Morphotactic Constraints:

× Lexical Rule Type 2 requires one of the following: verb-pc5_lrt1, verb-pc5_lrt ▼

Add a Require constraint

Current libraries (1/2)

- Word order (Bender & Flickinger 2005, Fokkens 2010)
- Coordination (Drellishak & Bender 2005)
  - Agreement in coordination (Dermer ms)
- Matrix yes-no questions* (Bender & Flickinger 2005)
- Morphotactics (O'Hara 2008, Goodman 2013)
- Case (+ direct-inverse marking) (Drellishak 2009)
- Agreement (person, number, gender) (Drellishak 2009)
- Argument optionality (pro-drop) (Saleem & Bender 2010)
- Tense and aspect (Poulson 2011)
- Sentential negation (Bender & Flickinger 2005, Crowgey 2012)

Current libraries (2/2)

- Information structure (Song 2014)
- Adjectives (attributive, predicative, incorporated) (Trimble 2014)
- Evidentials (Haeger7)
- Valence alternations (Curtis 2018)
- Adnominal possessives (Nielsen 2018)
- Nominalization (Howell et al 2018)
- Adverbial clauses (Howell & Zamaraeva 2018)
- Clausal complements (Zamaraeva et al 2019)
- *Wh*- questions (Zamaraeva 2021)

Grammar Matrix demo

- matrix.ling.washington.edu/customize/matrix.cgi

Creating a library for the customization system

- Choose phenomenon
- Review typological literature on phenomenon
- Refine definition of phenomenon
- Conceptualize range of variation within phenomenon
- Review HPSG (& broader syntactic) literature on phenomenon
- Pin down target MRSs
- Develop HPSG analyses for each variant
- Implement analyses in tdl
- Develop test suites
- Develop questionnaire & extend python backend
- Run regression tests
- Test with pseudo-languages
- Test with illustrative languages
- Test with held-out languages
- Add tests to regression tests
- Add to MatrixDoc pages

Languages in the Grammar Matrix regression testing system (Zamaraeva et al 2022:81)

Figure 10: Languages in the Grammar Matrix regression testing system. Map: Google Maps; map data: SIO, NOAA, U.S. Navy, NGA, GEBCO Image Landsat/Copernicus. Locations/spellings as in WALS (Dryer and Haspelmath 2013) or Glottolog (Hammarström *et al.* 2021). Starred* numbers are not fully visible. 1. Cree [crk]; 2. Makah [myh]; 3. Yakima Sahaptin [yak]; 4*. Umatila Sahaptin [uma]; 5. Eastern Pomo [peb]; 6*. Washo [was]; 7. Western Apache [apw]; 8. Lakota [lkt]; 9. Wichita [wic]; 10. Penobscot [pen]; 11. Zapotec [zpt]; 12. Awa Pit [kwi]; 13. Uranina [ura]; 14*. Shipibo-Konibo [shp]; 15. Moseten [cas]; 16. Paresi [pab]; 17. Apinaje [apn]; 18. West Greenlandic [kal]; 19. North Frisian [frr]; 20. English [eng]; 21. German [deu]; 22. French [fra]; 23. Basque [eus]; 24. Finnish [fin]; 25. Russian [rus]; 26. Ancient Greek [grc]; 27. Turkish [tur]; 28. Georgian [kat]; 29*. Tsez [ddo]; 30. Hebrew [heb]; 31. Gulf Arabic [afb]; 32. Jalkunan [bxl]; 33. Yoruba [yor]; 34. Madi [mhi]; 35*. Lango [laj]; 36. Zulu [zul]; 37*. Luo [luo]; 38. Kazakh [kaz]; 39. Ladakhi [ljb]; 40. Hindi [hin]; 41. Rawang [raw]; 42. Blang [blr]; 43. Quiang [cng]; 44. Mandarin [cmn]; 45. Pacoh [pac]; 46. Javanese [jav]; 47. Rukai [dru]; 48. Tagalog [tgl]; 49. Mongolian [mon]; 50. Japanese [jpn]; 51. Fore [for]; 52. Wambaya [wmb]; 53. Wangkangurru [wgg]; 54. Dyiribal [dbl]; 55. Lavukaleve [lvk]; 56. Fijian [fij]; 57. Maori [mri]; 58. Yukaghir [yux]; 59. Chukchi [ckt]; 60. Yup'ik [esu]

Table 1: The Grammar Matrix libraries with selected typological sources

Library	Citation(s)	Selected typological sources
Coordination	Drellishak and Bender 2005	Payne 1985; Stassen 2000; Drellishak 2004
Polar Questions	Bender and Flickinger 2005	–
Person, Number, Gender	Drellishak 2009	Cysouw 2003; Siewierska 2004; Corbett 2000
Agreement	Drellishak 2009	Corbett 2006
Case; Direct-Inverse	Drellishak 2008, 2009	Givón 1994
Argument Optionality	Saleem and Bender 2010; Saleem 2010	Ackema *et al.* 2006; Dryer 2013a
Tense	Poulson 2011	Comrie 1985; Dahl 1985
Aspect	Poulson 2011	Comrie 1976; Bybee *et al.* 1994
Lexicon	Drellishak and Bender 2005; Trimble 2014	Dixon 2004
Morphotactics	O'Hara 2008; Goodman 2013	–
Sentential Negation	Crowgey 2012, 2013	Dahl 1979; Dryer 2013b
Information Structure	Song 2014	Féry and Krifka 2008; Büring 2009
Adjectives; Copulas	Trimble 2014	Dixon 2004; Stassen 1997, 2013
Evidentials	Haeger 2017	Aikhenvald 2004; Murray 2017
Nominalized Clauses	Howell *et al.* 2018	Noonan 2007
Clausal Modifiers	Howell and Zamaraeva 2018	Thompson *et al.* 1985
Valence Change	Curtis 2018b,a	Haspelmath and Müller-Bardey 2001
Adnominal Possession	Nielsen and Bender 2018; Nielsen 2018	Payne and Barshi 1999; Heine 1997
Clausal Complements	Zamaraeva *et al.* 2019b	Noonan 2007
Constituent Questions	Zamaraeva 2021a	Haspelmath *et al.* 2013; Hagège 2008

(Zamaraeva et al 2022:64)

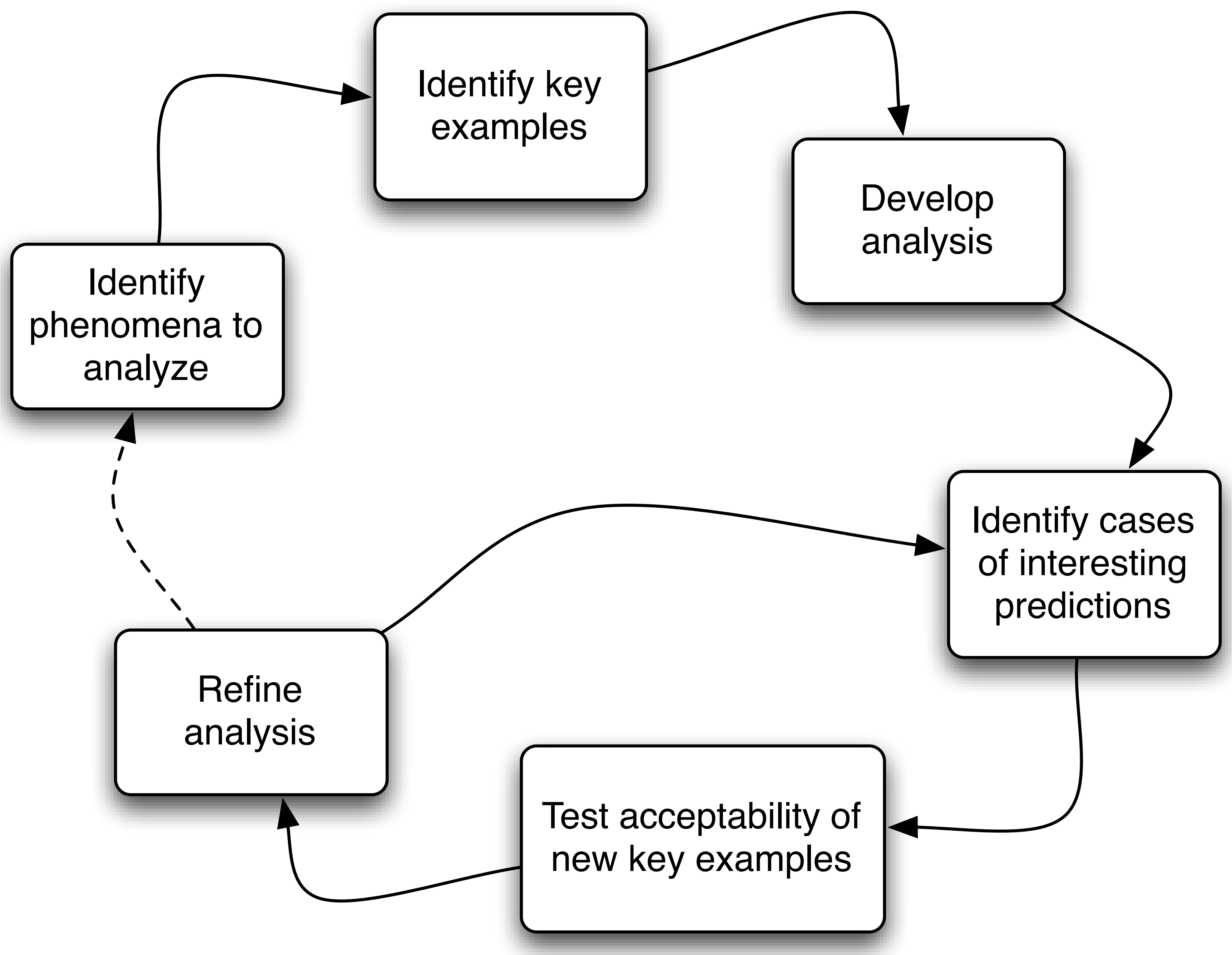
Typology and the Grammar Matrix

- Typological surveys provide critical knowledge about the range of variation for specific linguistic phenomena
- Implementation in the Grammar Matrix puts analyses of all of those variants into a *system où tout se tient* with all of the other implemented phenomena
- Implementation in the Grammar Matrix allows for evaluation on **held out** languages

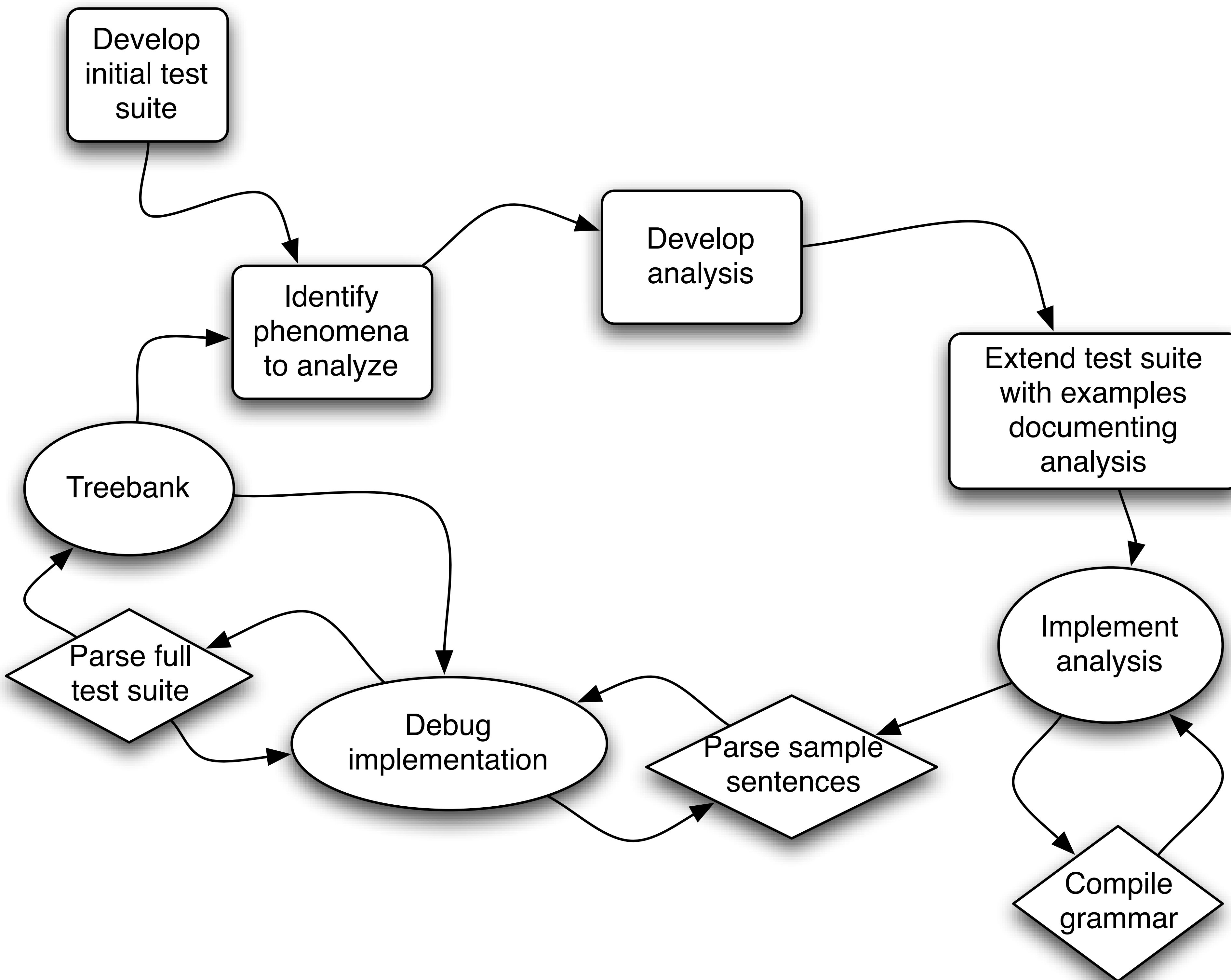
The value of commitment

- Grammar engineering means building something
- Building something means choices
- Choices mean specific commitments
- … But this is a feature, and not a bug!

Pen and paper syntax work-flow



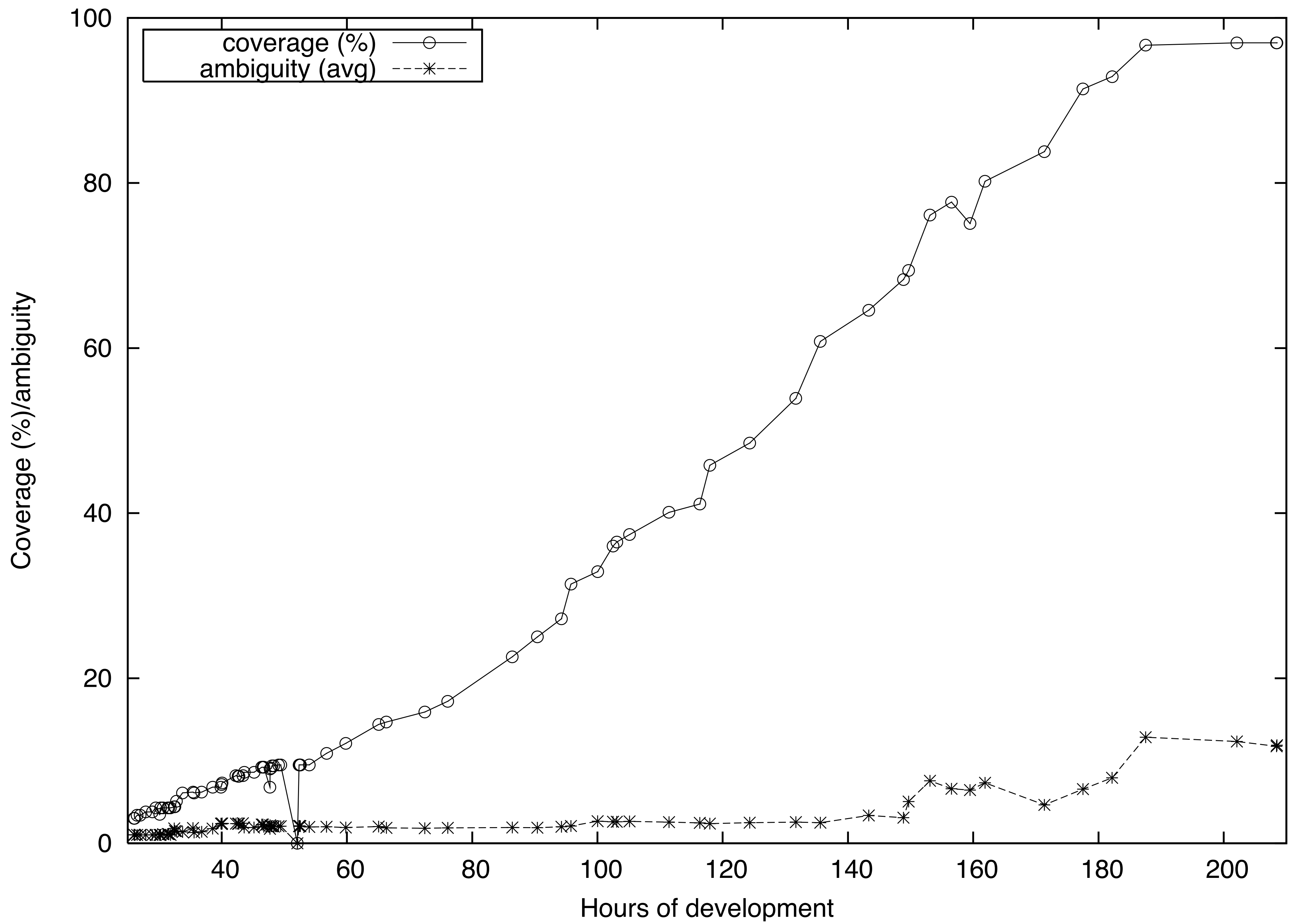
Grammar engineering work flow (Bender et al 2011)



Verifying predictions

- Application of test-driven development (Beck 2003) to linguistic research
- Create test suites of positive and negative examples
- Refine implemented grammar until performance on test suite is satisfactory

Wambaya grammar development (Bender 2008)



Verifying predictions

- Application of test-driven development (Beck 2003) to linguistic research
- Create test suites of positive and negative examples
- Refine implemented grammar until performance on test suite is satisfactory
- For Grammar Matrix libraries: Test with held-out languages.
    - Freeze system development
    - Create test suite & grammar specification
    - Measure

CLIMB: Virtual time travel in grammar engineering

- Fokkens (2011, 2014) observes that not only do linguistic phenomena interact, but analyses of linguistic phenomena interact
- Analytic choices made early in grammar development constrain the range of possible analyses later
- CLIMB (Fokkens 2011, 2014) extends metagrammar engineering to the full grammar development process
  - Builds directly on Grammar Matrix code base
  - Allows grammar writers to keep multiple hypotheses in play

Testing interactions

- Coordination breaks everything, so put coordination into the grammar (or metagrammar early)
- Run regression tests, so as to verify that new additions don't break previous functionality
    - No loss of coverage
    - No increase in (spurious) ambiguity
- Create specific tests that combine phenomena in a single sentence

Language-specific hypothesis testing

- Paresi-Haliti [pab] (Arawakan): mostly V-final (Brandão 2014) or SVO+OSV dominant (da Silva 2013)?
- Zamaraeva (2021) presents to Matrix-derived grammars (SOV & free WO), tested over 67 items

Table 3: Two word order hypotheses for Paresi-Haliti

Hypothesis	Raw coverage (%)	Validated coverage (%)	Overgeneration (%)	Ambiguity
SOV	40/64 (62.5)	25/64 (39.0)	2/3 (66.7)	43.95
free	53/64 (82.8)	36/64 (56.0)	2/3 (66.7)	36.17

(Zamaraeva et al 2022:90)

Language-specific hypothesis testing

- The ambiguity produced by these grammars revealed a bug in the interaction between the adnominal possession & *wh* questions libraries (now fixed) and an interaction between the nominalization & *wh* question libraries.

- Results after hand improvement of the grammars:

Hypothesis	Raw coverage (%)	Validated coverage (%)	Overgeneration (%)	Ambiguity
SOV	41/64 (64.1)	36/64 (56.2)	2/3 (66.7)	4.02
free	51/64 (79.7)	42/64 (65.6)	2/3 (66.7)	3.98

Table 4: Improved grammars for Paresi-Haliti

(Zamaraeva et al 2022:91)

Accumulating evidence for analyses robustness

- Copula introduced in Trimble's (2014) library for adjectives, without reference to polar questions + subject-auxiliary inversion rules from Bender & Flickinger 2005 => minimal spurious ambiguity, easy to eliminate

- Basic lexical type hierarchy has served as a good foundation for all libraries, including most recently the addition of *wh* question words (Zamaraeva 2021)

- Nielsen's (2018) adnominal possession library, developed before Zamaraeva's (2021) *wh* question library, provided for Jalkunan possessive pronouns, even though the *wh* question library development hadn't covered them.

Core (abstract) cross-linguistic hypotheses

- Words and phrases combine to make larger phrases.
- The semantics of a phrase is determined by the words in the phrase and how they are put together.
- Some rules for phrases add semantics (but some don't).
- Most phrases have an identifiable head daughter.
- Heads determine which arguments they require and how they combine semantically with those arguments.
- Modifiers determine which kinds of heads they can modify, and how they combine semantically with those heads.
- No lexical or syntactic rule can remove semantic information

Getting started with the Grammar Matrix

- Getting started page: https://delph-in.github.io/docs/matrix/MatrixGettingStarted/
- Customization system: https://matrix.ling.washington.edu/customize/matrix.cgi
- Ling 567 @ UW: https://courses.washington.edu/ling567/
- Grammar engineering FAQ: https://delph-in.github.io/docs/matrix/GrammarEngineeringFAQ/
- Matrix Docs: https://delph-in.github.io/docs/matrix/MatrixDocTop
- DELPH-IN Discourse site: https://delphinqa.ling.washington.edu/

Conclusion

- A flexible framework for building up, over time and in a data-driven fashion, a set of analyses which are demonstrably useful for describing the repertoire of grammatical variation in the world's languages
- This follows from three properties of the framework:
    - the Grammar Matrix design is informed by typological literature + HPSG theory
    - the development methodology prioritizes cross-linguistic applicability of the analyses
    - the regression testing system affords automatic verification of compatibility of new + existing analyses

Thank you!