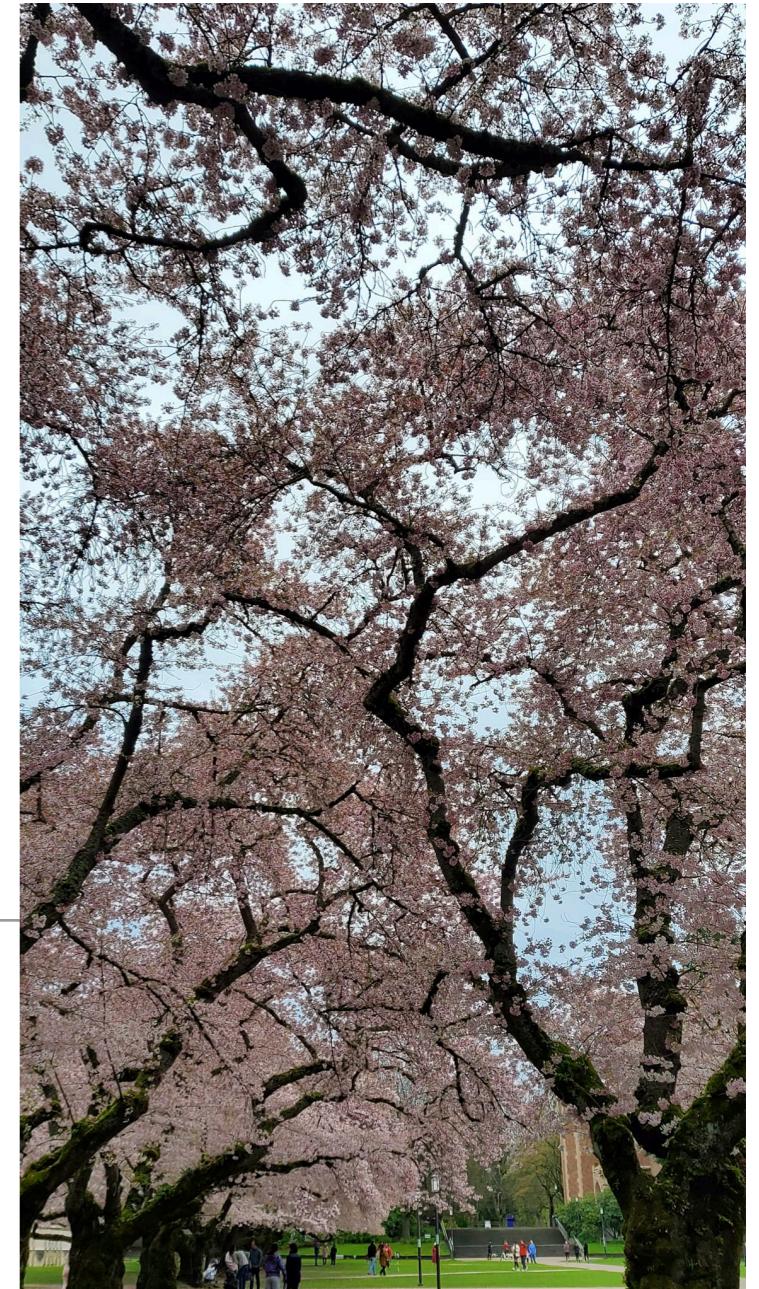


University of Washington site update

DELPH-IN Summit
July 18, 2022
Fairhaven, WA

Emily M. Bender, Liz Conrad, Allison Dods, Yi-Chien Lin,
Rosetta Pendleton, Tom Liu, Tara Wueger, Andrew Briand
+ special guest star Rafe Menon



567 Languages (WI 2022)

arp	Arapaho	Algic	Kodama & Pendleton
bbl	Tsova-Tush	Nakh-Daghestanian	Briand & Brown
bcj	Bardi	Nyulnyulan	Nezakati & Wueger
erk	South Efate	Austronesian	Ban & Bansal
hix	Hixkaryana	Cariban	Lin & Soni
mni	Meitei	Sino-Tibetan	Liu & Zeng
ttv	Titan	Austronesian	Guo & Wang
wbl	Wakhi	Indo-European	Langford & Okada
yaq	Hiaki	Uto-Aztecán	Iyer & O'Connell
eng	Indo-European		
sje	Pite Saami		Nielsen & Spivey

Languages - mapped



lat/long from WALS & Glottolog, map from Google

Items with end-to-end output: Final (transfer rule propagation)

	arp	bbl	bcj	eng	erk	hix	mni	sje	ttv	wbl	yaq
arp	14	12	8	11	11	9	9	12	12	12	7
bbl	9	14	8	13	14	11	8	11	9	9	6
bcj	6	7	9	7	5	7	6	7	7	8	6
eng	13	22	12	26	20	14	12	21	18	17	12
erk	10	12	8	14	17	10	10	15	13	13	8
hix	7	10	6	11	8	11	8	9	10	10	7
mni	5	5	7	5	3	5	6	5	4	5	3
sje	10	15	8	16	15	9	9	18	14	14	10
ttv	1	1	8	14	14	11	9	14	16	14	1
wbl	10	12	7	9	10	10	5	12	14	16	10
yaq	5	7	7	6	7	7	3	7	9	9	10

Projects & Theses

- **Wueger** senior honors thesis: Isolating the source of poor morphological output from MOM/AGGREGATION for Bardi & developing an initial solution
- **Liu** senior honors thesis: Grammar Matrix debugging, focused on adnominal possession and its interactions; towards a model of a debugging process
- **Dods** MS thesis: AGG inference for adnominal possession
- **Lin** MS thesis: AGG inference for valence changing morphology
- **Pendleton** MS project: reviving and extending test by generation
- **Briand** MS project: isolating cause of ERG slow-down since 1214

Additional efforts

- **Conrad:** Making AGG data handling, set-up and testing harness robust; towards investigation of information structure x free-word order interactions
- **Liu:** Improving the responsiveness of the Grammar Matrix questionnaire
- **Wueger:** More robust handling of root-finding in morphological inference
- **Bansal:** Serial verb constructions, likely in Thai, possibly working towards a Grammar Matrix library for SVCs

Journal publication!

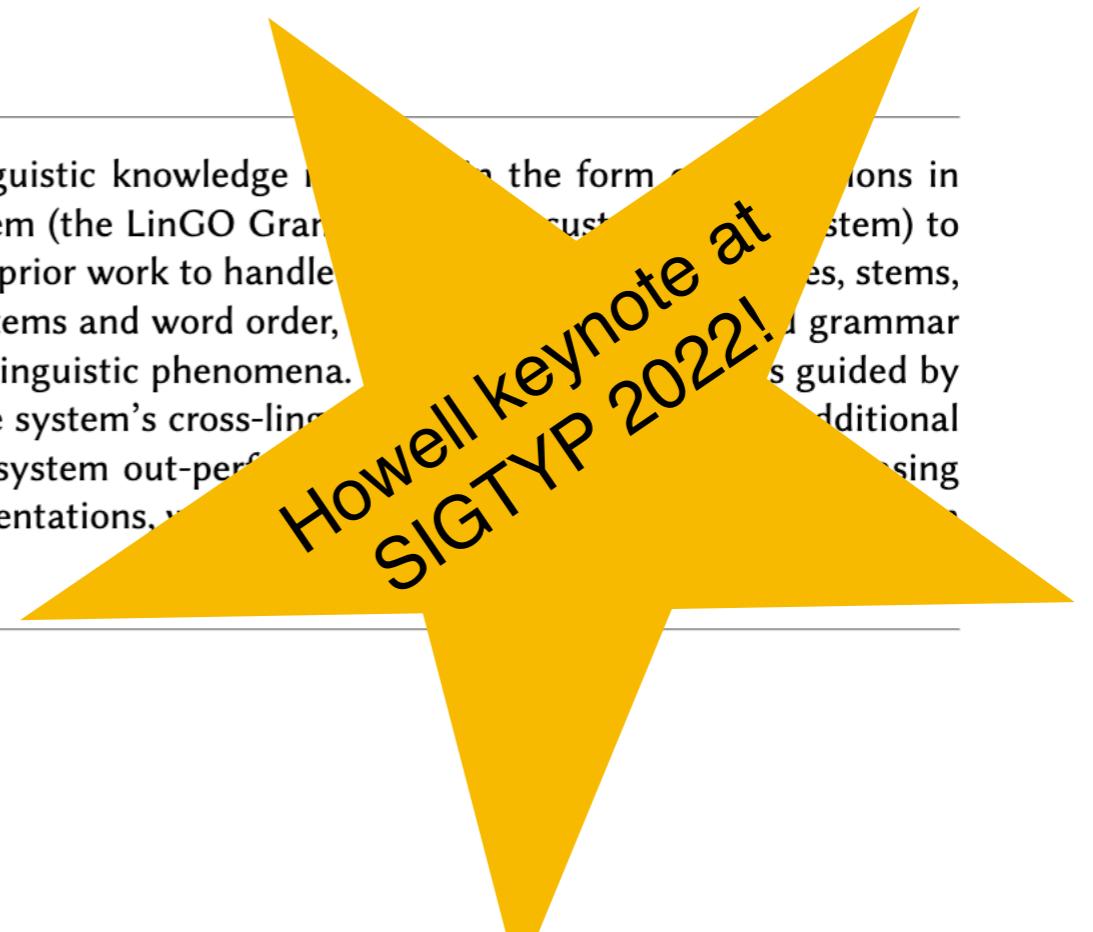
Howell & Bender 2022 - NEJLT 8(1)

Building Analyses from Syntactic Inference in Local Languages: An HPSG Grammar Inference System

Kristen Howell, University of Washington and LivePerson Inc., USA kphowell@uw.edu

Emily M. Bender, University of Washington, USA ebender@uw.edu

Abstract We present a grammar inference system that leverages linguistic knowledge in interlinear glossed text (IGT) and in a meta-grammar engineering system (the LinGO Grammar Engineering System) to automatically produce machine-readable HPSG grammars. Building on prior work to handle affixes and position classes, and preliminary work on inferring case systems and word order, we present a new inference system called **BASIL** that covers a wide range of fundamental linguistic phenomena. We test **BASIL** on 27 genealogically and geographically diverse languages, and we test the system's cross-linguistic generalization on 5 held-out languages, using datasets provided by field linguists. Our system out-performs previous work in grammar inference while limiting ambiguity and producing richer semantic representations, without requiring any manual intervention or domain-specific knowledge.



Howell keynote at
SIGTYP 2022!

Journal publication!

Zamaraeva et al 2022 (to appear) - JLM

20 years of the Grammar Matrix:
cross-linguistic hypothesis testing of
increasingly complex interactions

*Olga Zamaraeva¹, Chris Curtis², Guy Emerson³, Antske Fokkens⁴⁵,
Michael Wayne Goodman⁶, Kristen Howell⁶, T.J. Trimble⁷, and Emily
M. Bender⁸*

¹ Universidade da Coruña

² Firemuse Research

³ University of Cambridge

⁴ Vrije Universiteit Amsterdam

⁵ Eindhoven University of Technology

⁶ LivePerson Inc.

⁷ Independent Scholar

⁸ University of Washington

The Grammar Matrix at 20

- “a system of analyses for which there is a demonstrated area of applicability – which also grows over time”
- Summary of additions to the Grammar Matrix since 2010
- Case studies of using the Grammar Matrix to identify tensions between analyses
- Examples of analyses that have proven particularly robust over the years
- How the Grammar Matrix supports other projects

Additions to the Grammar Matrix: Libraries

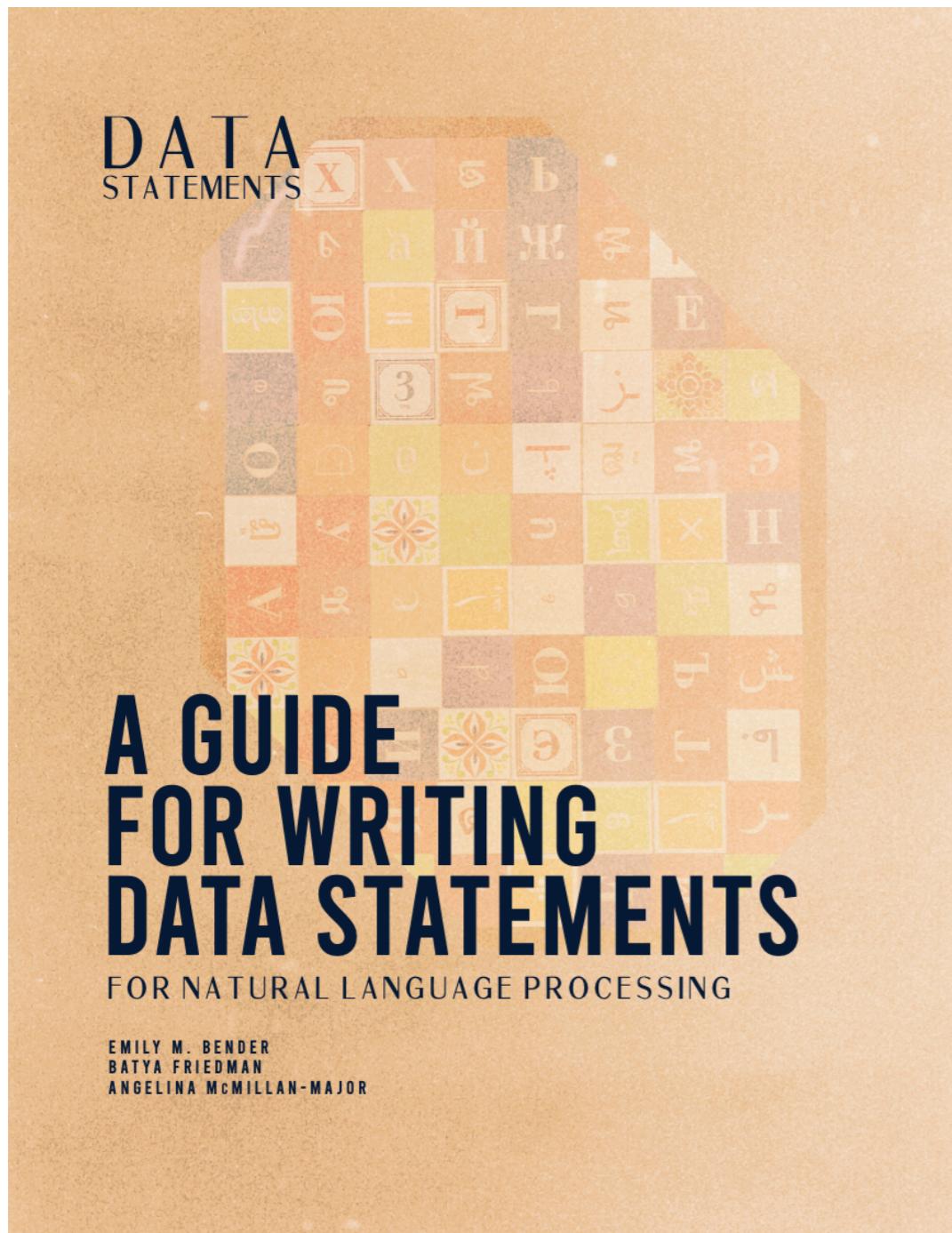
Table 1: The Grammar Matrix libraries with selected typological sources

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 <i>et al.</i> 2006; Dryer 2013a
Tense	Poulson 2011	Comrie 1985; Dahl 1985
Aspect	Poulson 2011	Comrie 1976; Bybee <i>et al.</i> 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 <i>et al.</i> 2018	Noonan 2007
Clausal Modifiers	Howell and Zamaraeva 2018	Thompson <i>et al.</i> 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 <i>et al.</i> 2019b	Noonan 2007
Constituent Questions	Zamaraeva 2021a	Haspelmath <i>et al.</i> 2013; Hagège 2008

Additions to the Grammar Matrix: Methodological and Formal Innovations

- CLIMB (Fokkens 2014)
- Spring Cleaning (Fokkens et al 2011)
- Test-driven development (Zamaraeva 2021a)
- Evaluation with held-out language families (from Haeger 2017 onwards)
- Emerson-Turing types (Emerson 2017, 2019, 2021, forthcoming)

Data Statements – v2 schema + writing guide with Angie McMillan-Major & Batya Friedman



SCHEMA ELEMENTS VERSION 2

- 1 HEADER
 - 2 EXECUTIVE SUMMARY
 - 3 CURATION RATIONALE
 - 4 DOCUMENTATION FOR SOURCE DATASETS
 - 5 LANGUAGE VARIETIES
 - 6 SPEAKER DEMOGRAPHIC
 - 7 ANNOTATOR DEMOGRAPHIC
 - 8 SPEECH SITUATION AND TEXT CHARACTERISTICS
 - 9 PREPROCESSING AND DATA FORMATTING
 - 10 CAPTURE QUALITY
 - 11 LIMITATIONS
 - 12 METADATA
 - 13 DISCLOSURE AND ETHICAL REVIEW
 - 14 OTHER
 - 15 GLOSSARY

<http://techpolicylab.uw.edu/data-statements/>

AI hype & the media

Home // Radio // Here & Now

A metal head made of motor parts symbolizes artificial intelligence, or AI, at the Essen Motor Show for tuning and motorsports in Essen, Germany. (Martin Meissner/AP)

Don't worry about the robot revolution: One expert explains why AI is nowhere near sentience

June 27, 2022

06:16

SHOWS PEOPLE FORUMS STITCHER



 SHOWS PEOPLE FORUMS STITCHER

FACTUALLY! The Real Problem with A.I. with Emily Bender

FACTUALLY! WITH ADAM CONOVER #162 JUNE 21, 2022

Is artificial intelligence a problem, or is the real problem how we're using the term in the first place? Linguistics professor Emily Bender joins Adam to discuss why we should resist the urge to be impressed when it comes to big tech's AI promises, and how our belief in the fantasy of A.I. could be worse than the reality. You can follow Emily on Twitter at @emilymbender.

INNOVATIONS

This AI model tries to re-create the mind of Ruth Bader Ginsburg

Its creators say the AI app helps ordinary people understand how artificial intelligence is progressing. Critics contend there's much work to be done.

By Pranshu Verma
June 14, 2022 at 7:00 a.m. EDT

