



Linguistic Society of America

CENTENNIAL | 1924-2024

What we can do with what we know about how language works

Emily M. Bender
University of Washington

2024 LSA ANNUAL MEETING
#LSA2024



“I’m a linguist, so I know how language works...”



Clarion West

PRESENTS

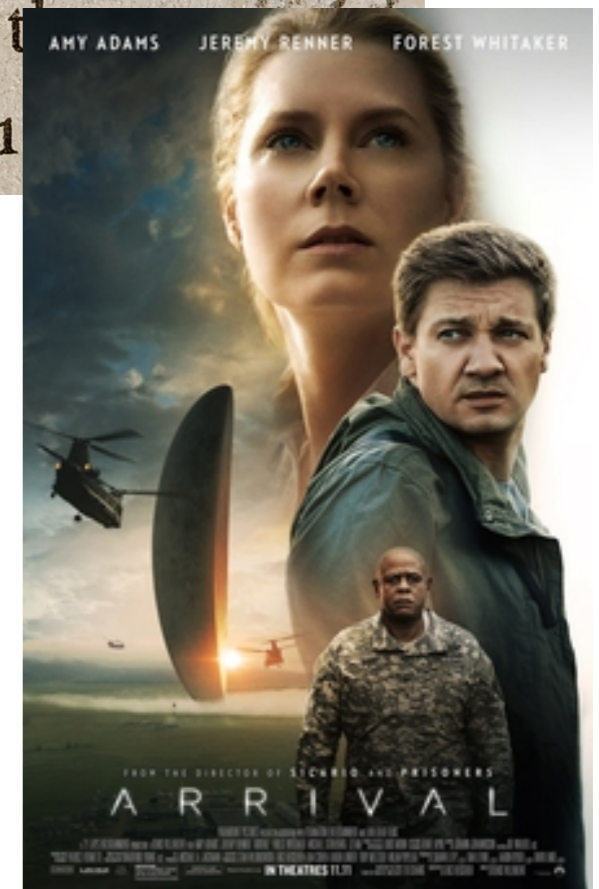
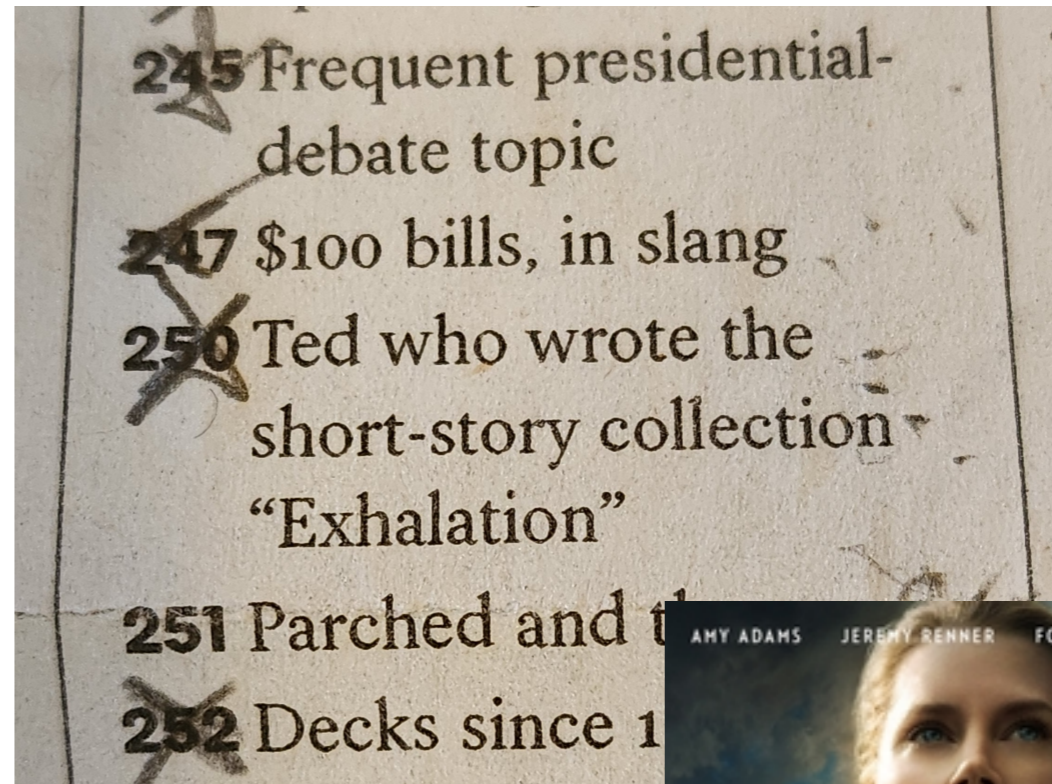
WISHFUL THINKING AND AI

*Ted Chiang & Dr. Emily M. Bender in
Conversation*



A fundraiser for Clarion West: Join us for an evening with two of Time Magazine’s 100 Most Influential People in AI as they discuss the hype, realistic expectations, and whose voices are missing in this rapidly changing field.

Friday, November 10 at 7:30 PM
Town Hall Seattle



Linguists are in the business of knowing (and finding out) how language works

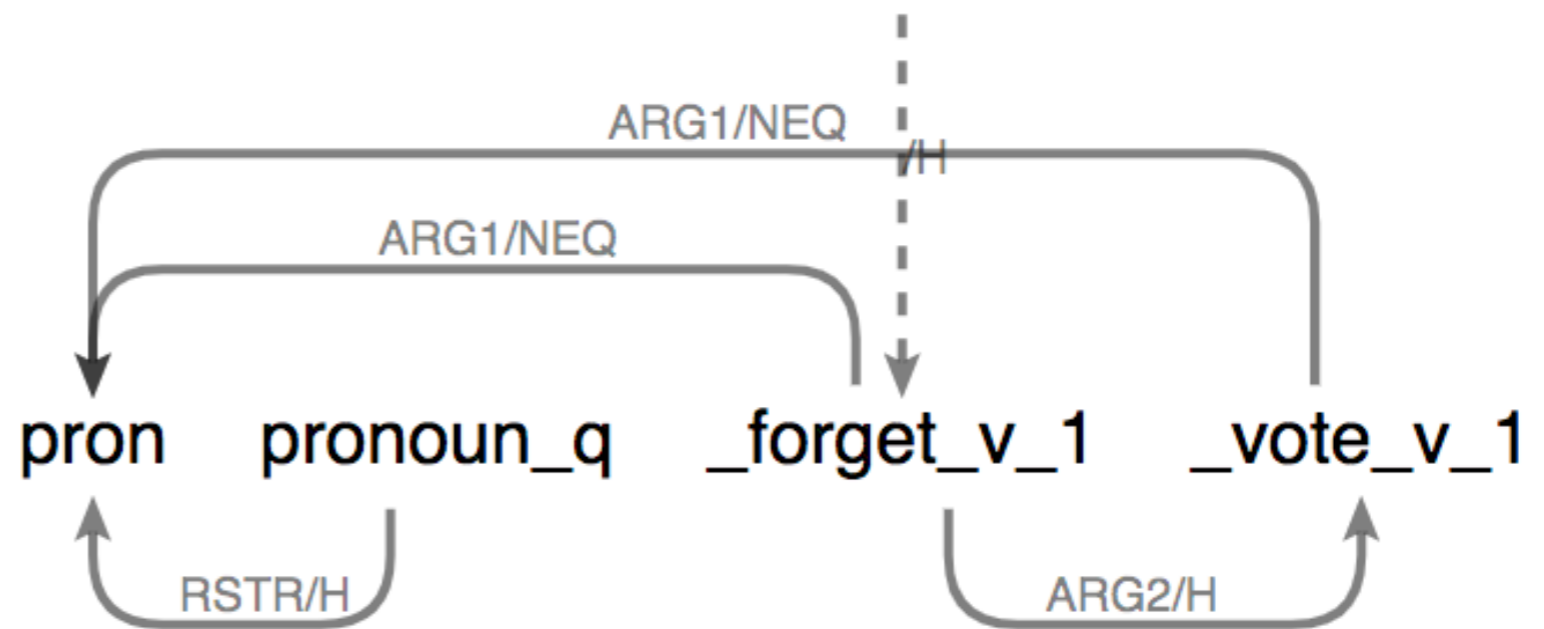
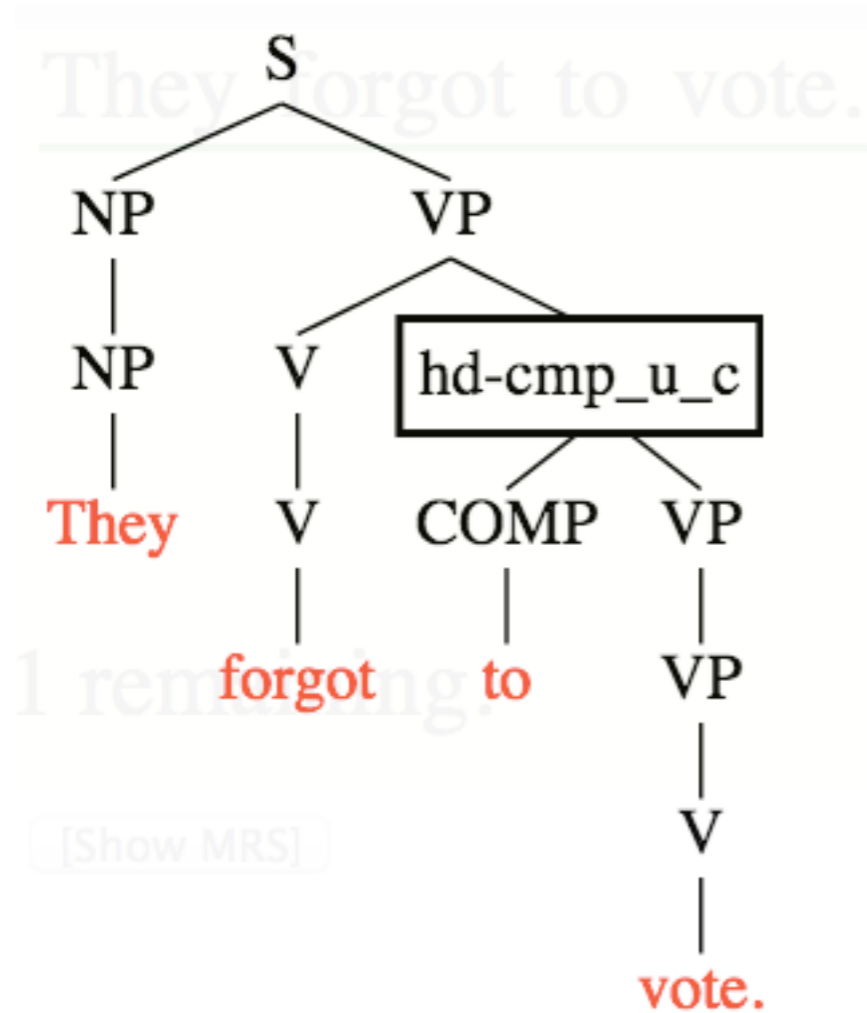
- As an abstract system
- In the process of language learning
- In face to face communication
- In interactions with societal structures
- ...

PhD: Syntax and sociolinguistics (2000)

SYNTACTIC VARIATION AND LINGUISTIC
COMPETENCE:
THE CASE OF AAVE COPULA ABSENCE

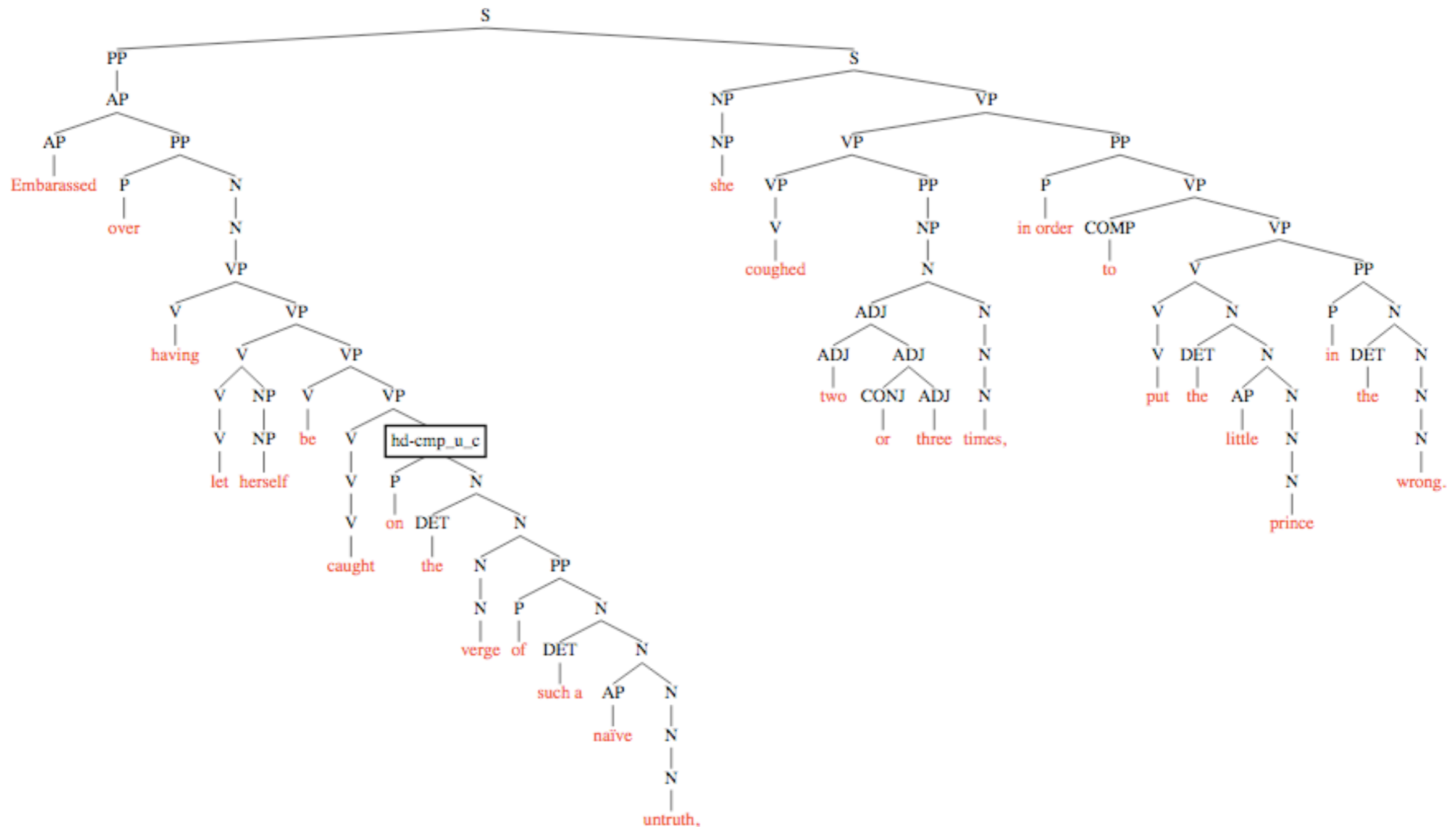
A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

RA work on the English Resource Grammar (Flickinger 2000, 2011)



[Show MRS]

RA work on the English Resource Grammar (Flickinger 2000, 2011)



The Grammar Matrix:

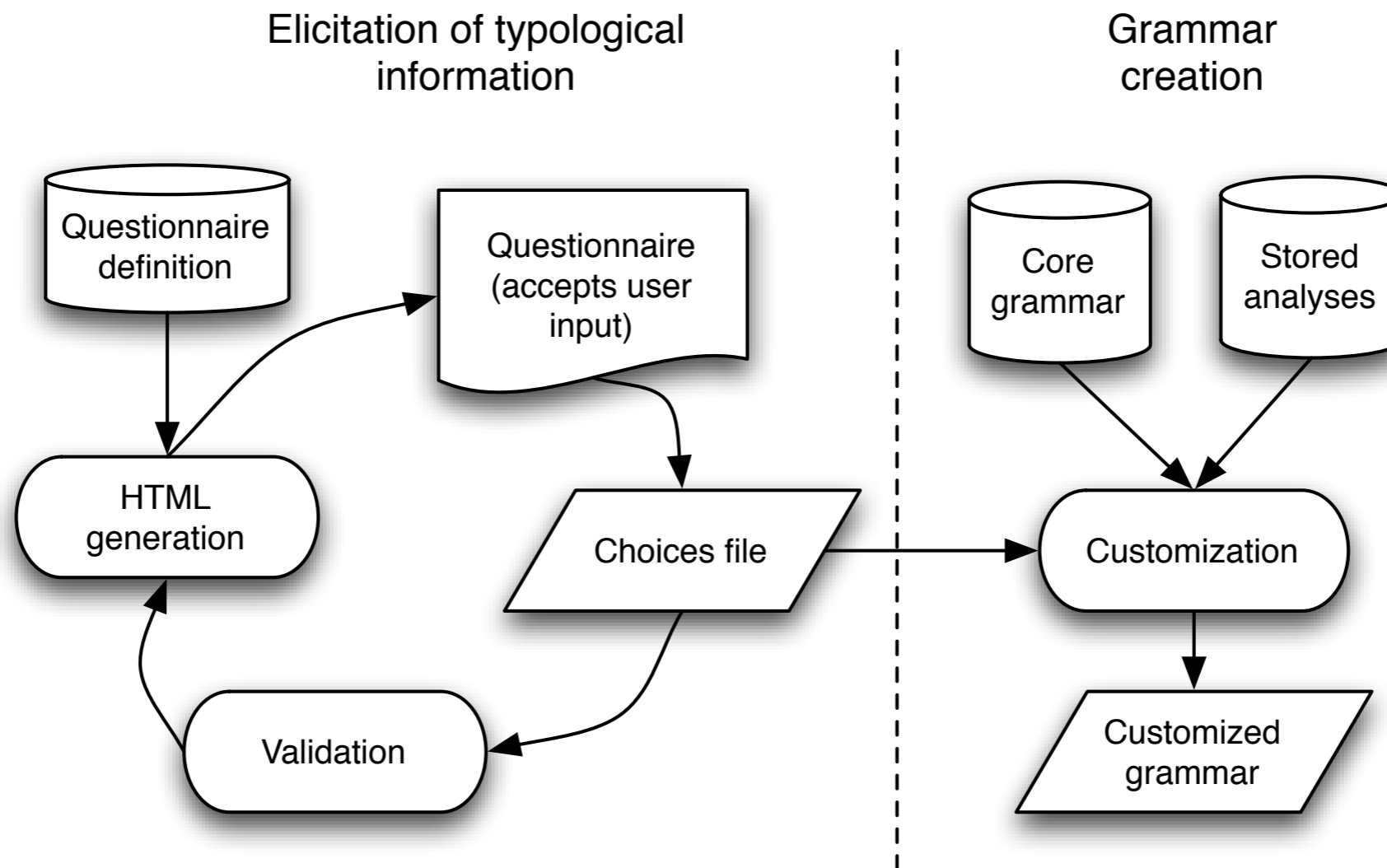
Computational methods for combining typological breadth & syntactic depth
(Bender et al 2002, 2010; Zamaraeva et al 2022)

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

The Grammar Matrix:

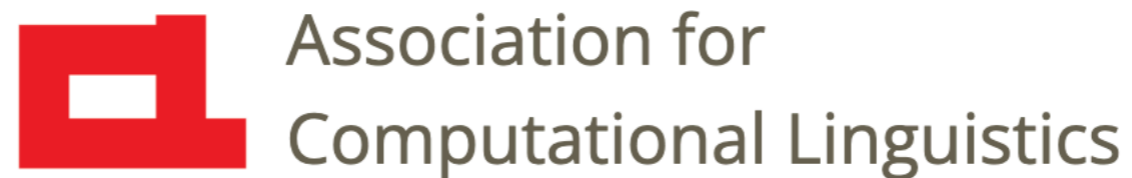
Computational methods for combining typological breadth & syntactic depth
(Bender et al 2002, 2010; Zamaraeva et al 2022)



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

Computational linguistics

- Faculty director, UW CLMS (since program inception in 2005)
 - <https://www.compling.uw.edu/>
- NAACL Board (2012-2019; Chair 2016-2017)
- ACL Executive Board (2016-2017 & 2022-2025; President 2024)



POV: A linguist living in machine-learning land

- Emphasis is on getting the computer to ‘learn’ directly from data
 - ‘Better’ if the researcher never looks at the data at all
 - Devaluing of domain expertise (the ‘expensive’ people the machine is supposed to replace)
 - => Hierarchies of knowledge (Raji et al 2021, Gebru & Torres 2023)
- Frequent thought while reviewing papers: These people don’t know the first thing about how language works!

But compiling is an interdisciplinary field



The 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Held at the Le Centre Sheraton Montréal
1201, boul. René-Lévesque ouest, Montréal, (Québec),
Canada, H3B-2L7

June 3-8, 2012

Email: acl-AT-aclweb.org

T1: 100 Things You Always Wanted to Know about Linguistics But Were Afraid to Ask*

***... for fear of being told 1000 more**

Emily M. Bender
Morning session, 9am-12:30pm

But compiling is an interdisciplinary field



Tutorials

Tutorials will be held on July 15th, 2018. All tutorials will run for a half-day at the times noted below

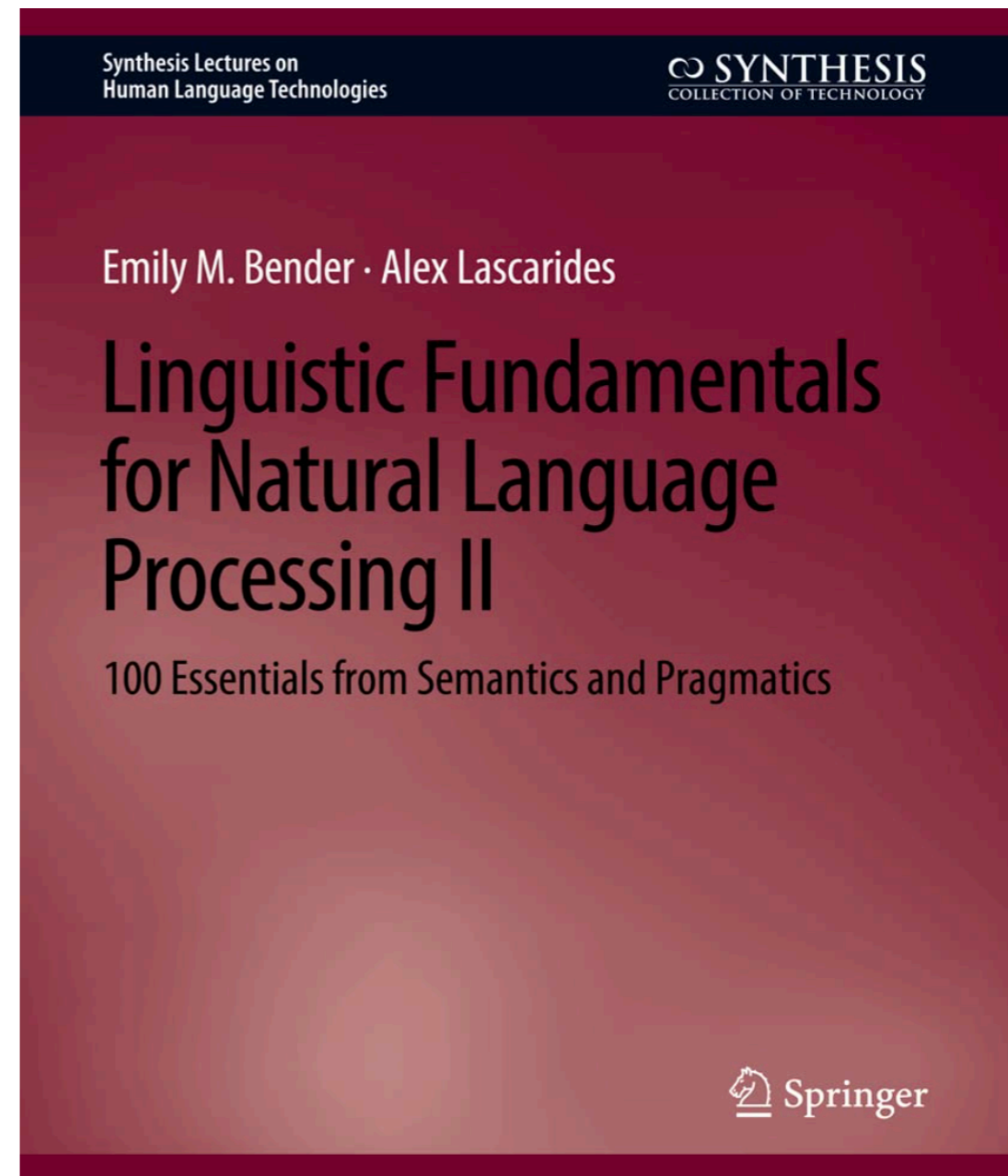
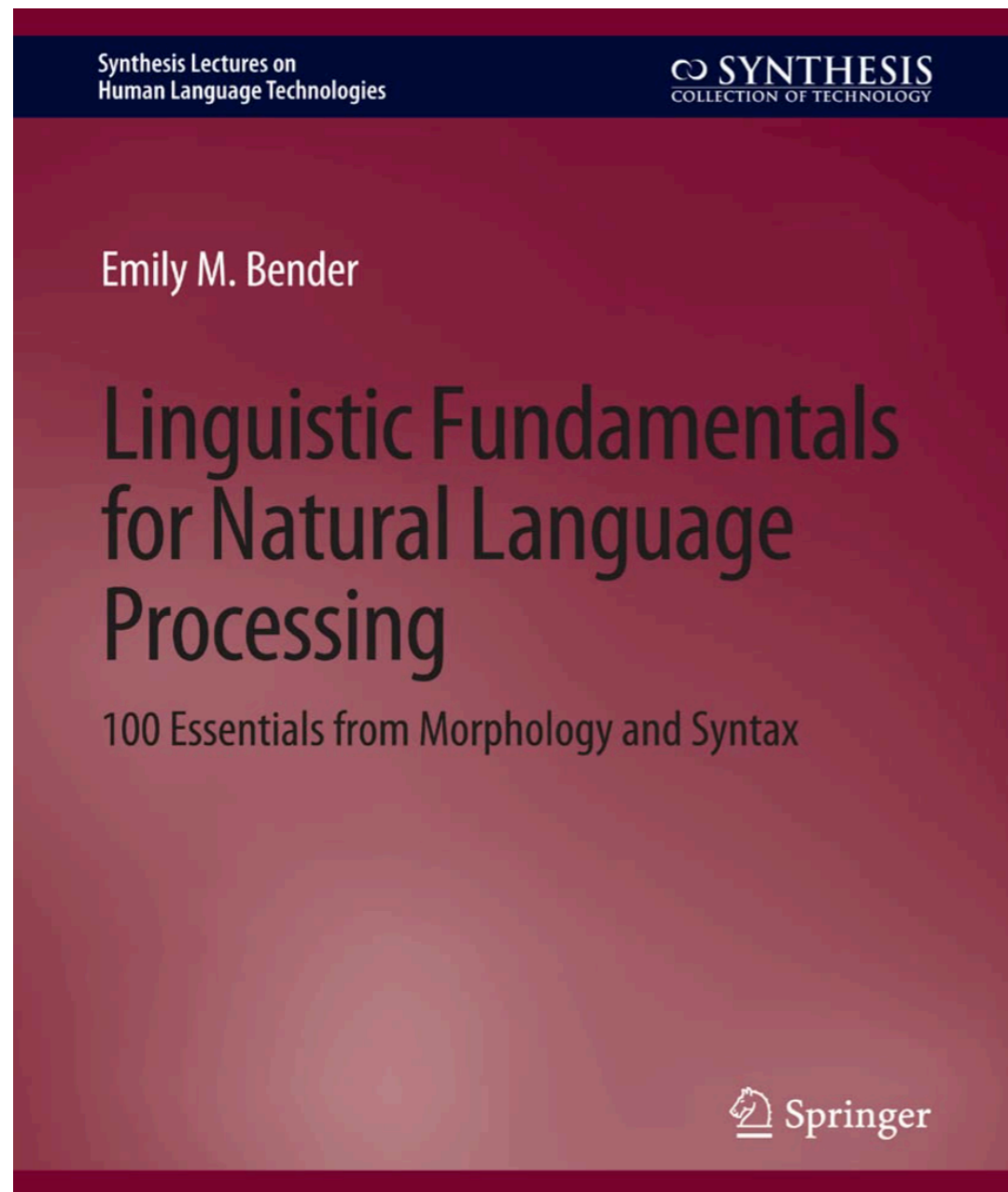
T1: 100 Things You Always Wanted to Know about Semantics & Pragmatics But Were Afraid to Ask

Emily M. Bender

09:00 – 12:30

Location: 216, MCEC

But compiling is an interdisciplinary field



POV: A linguist living in machine-learning land



@emilybender@dair-community.social on Mastodon
@emilybender

Dear Computer Scientists,

"Natural Language" is **not** a synonym for "English".

That is all.

-Emily

9:32 AM · Nov 26, 2018

The #BenderRule origins, part 1

4 Prescriptions for Typologically-Informed NLP

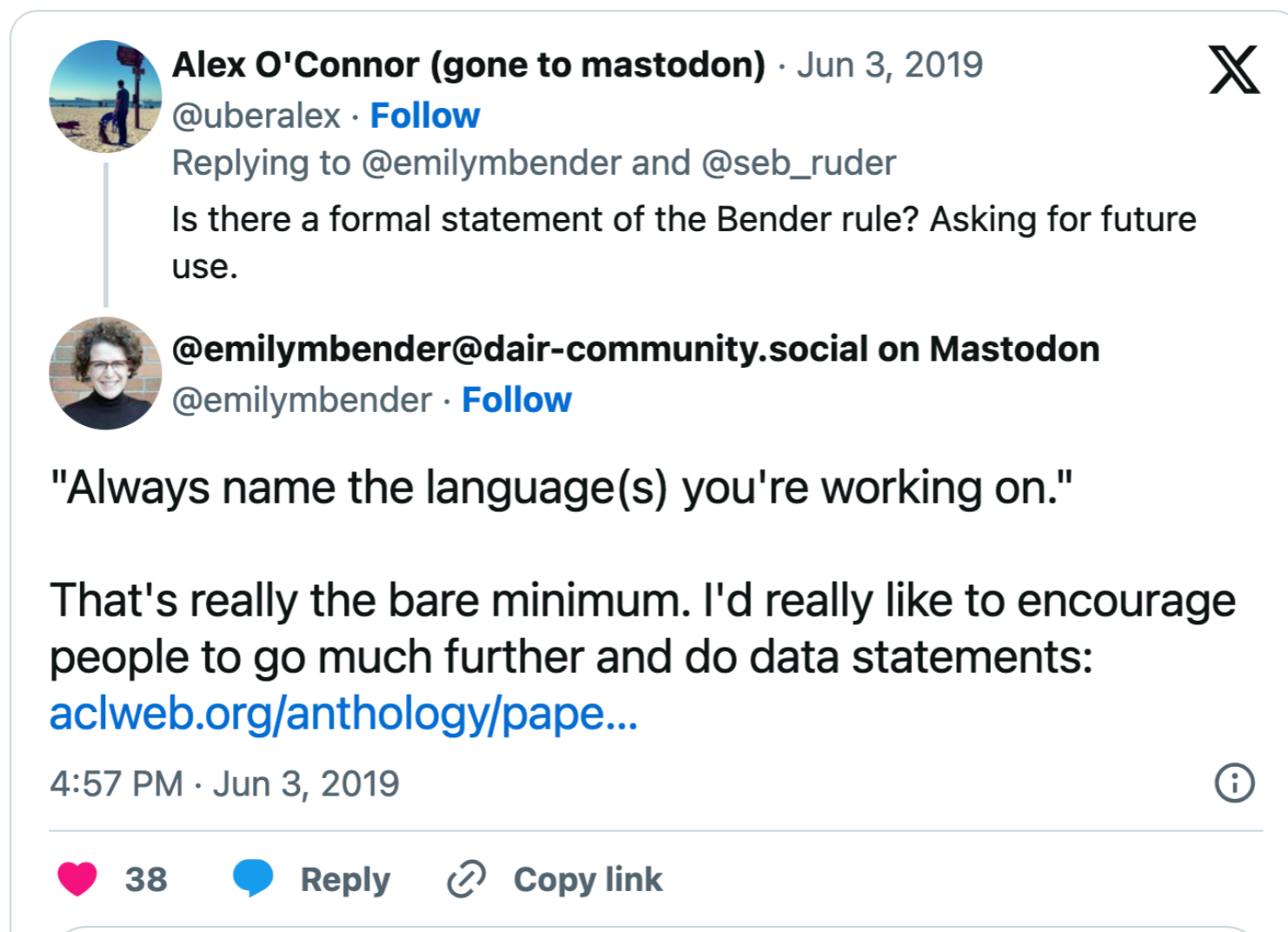
This section provides a brief list, in etiquette-book format, of suggestions for producing more typologically-informed (and thus hopefully more language-independent) NLP research. These suggestions are aimed at both authors and reviewers, or rather, at members of the NLP community in both their author and reviewer guises.

Do state the name of the language that is being studied, even if it's English. Acknowledging that we are working on a particular language foregrounds the possibility that the techniques may in fact be language-specific. Conversely, neglecting to state that the particular data used were in, say, English, gives false veneer of language-independence to the work.

(Bender 2011: On Achieving and Evaluating
Language-Independence in NLP)

The #BenderRule origins, part 2

- At NAACL 2019 (and elsewhere) made a pest of myself: “What language was this on?”



A screenshot of a Mastodon conversation. The top post is from Alex O'Connor (@uberalex), dated June 3, 2019, replying to @emilybender and @seb_ruder. The text of the post asks: "Is there a formal statement of the Bender rule? Asking for future use." The bottom post is from @emilybender@dair-community.social, dated June 3, 2019, at 4:57 PM. The text of the reply reads: "Always name the language(s) you're working on." followed by "That's really the bare minimum. I'd really like to encourage people to go much further and do data statements: [aclweb.org/anthology/pape...](\"http://aclweb.org/anthology/pape...\")". The interface shows 38 likes, a reply button, and a copy link button.

Alex O'Connor (gone to mastodon) · Jun 3, 2019
@uberalex · **Follow**
Replying to @emilybender and @seb_ruder
Is there a formal statement of the Bender rule? Asking for future use.

@emilybender@dair-community.social on Mastodon
@emilybender · **Follow**
"Always name the language(s) you're working on."
That's really the bare minimum. I'd really like to encourage people to go much further and do data statements:
aclweb.org/anthology/pape...
4:57 PM · Jun 3, 2019

38 Reply Copy link

(Bender 2019)

Similar precepts hold in other areas of linguistics

- Case in point, the Holliday Rule:



Nicole Holliday (mixedlinguist@bsky.social)

@mixedlinguist



Introducing the Holliday rule for ling papers, summarized as "It's alright to say they're white" 😂. If you got participants, you should ask their race(s), tell us what they said, even if they're white! h/t [@kirbyconrod](#) for the idea, & [@emilymbender](#) for pioneering this kinda rule!

11:04 AM · Dec 2, 2020

22 Reposts **7** Quotes **153** Likes **5** Bookmarks

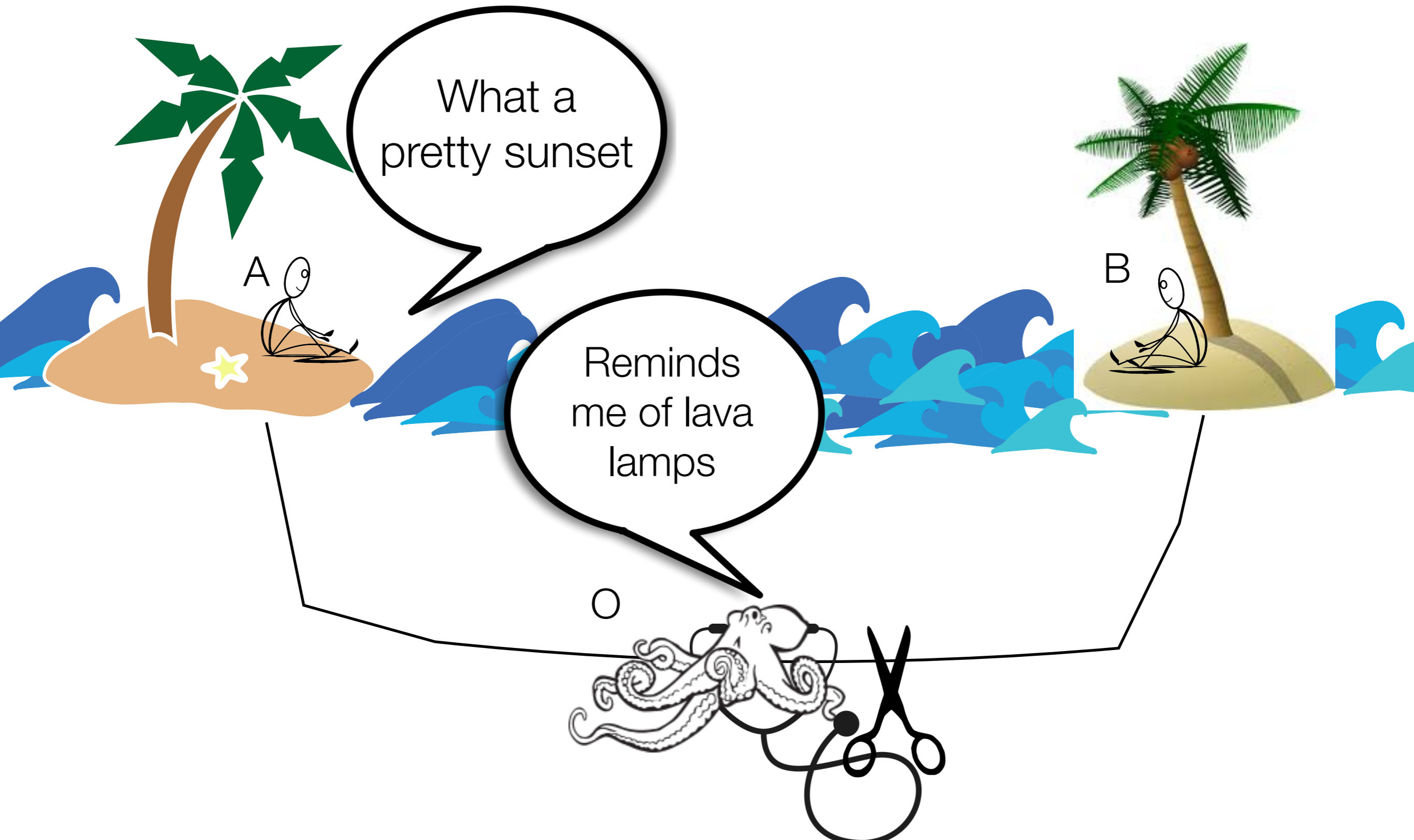
POV: A linguist living in machine-learning land

- With the release of BERT (Devlin et al 2019) by Google in 2018, started to see repeated claims of BERT and similar language models “understanding”
- But a language model (better: corpus model, Veres 2022) models the distribution of word forms in text
 - Distributional similarity gives a clue to semantic similarity (Harris 1954, Firth 1957)
 - ... but also lots of other information like biases (Caliskan 2017, Bolukbasi et al 2016, Herbelot et al 2012, Garg et al 2018)
 - ... and nothing like understanding
- => Interminable Twitter arguments and then finally the octopus paper

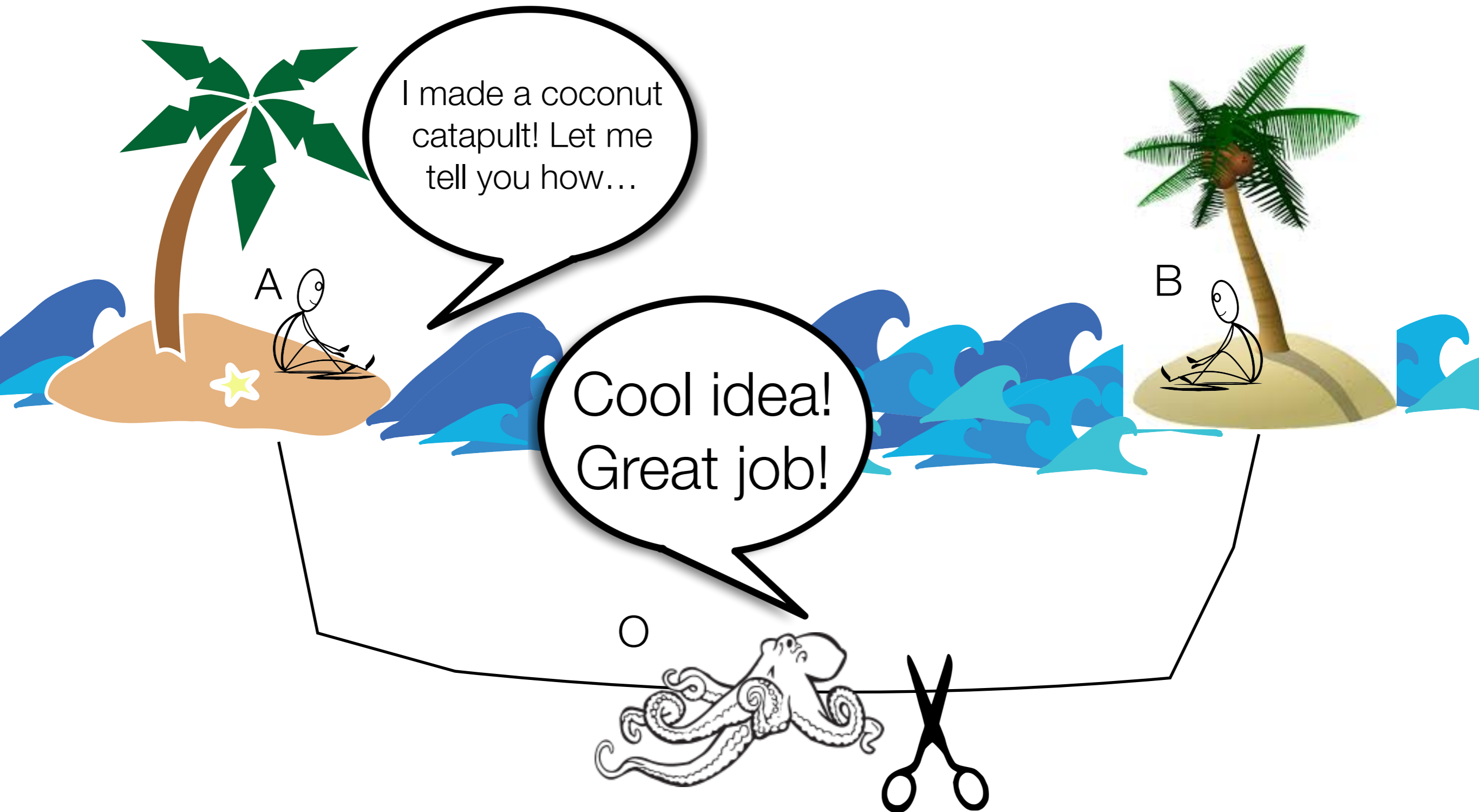
Bender & Koller 2020: Working definitions

- **Form** : marks on a page, pixels or bytes, movements of the articulators
- **Meaning** : relationship between linguistic form and something external to language
 - $\underline{M} \subseteq E \times I$: pairs of expressions and communicative intents
 - $\underline{C} \subseteq E \times S$: pairs of expressions and their standing meanings
- **Understanding** : given an expression e , in a context, recover the communicative intent i

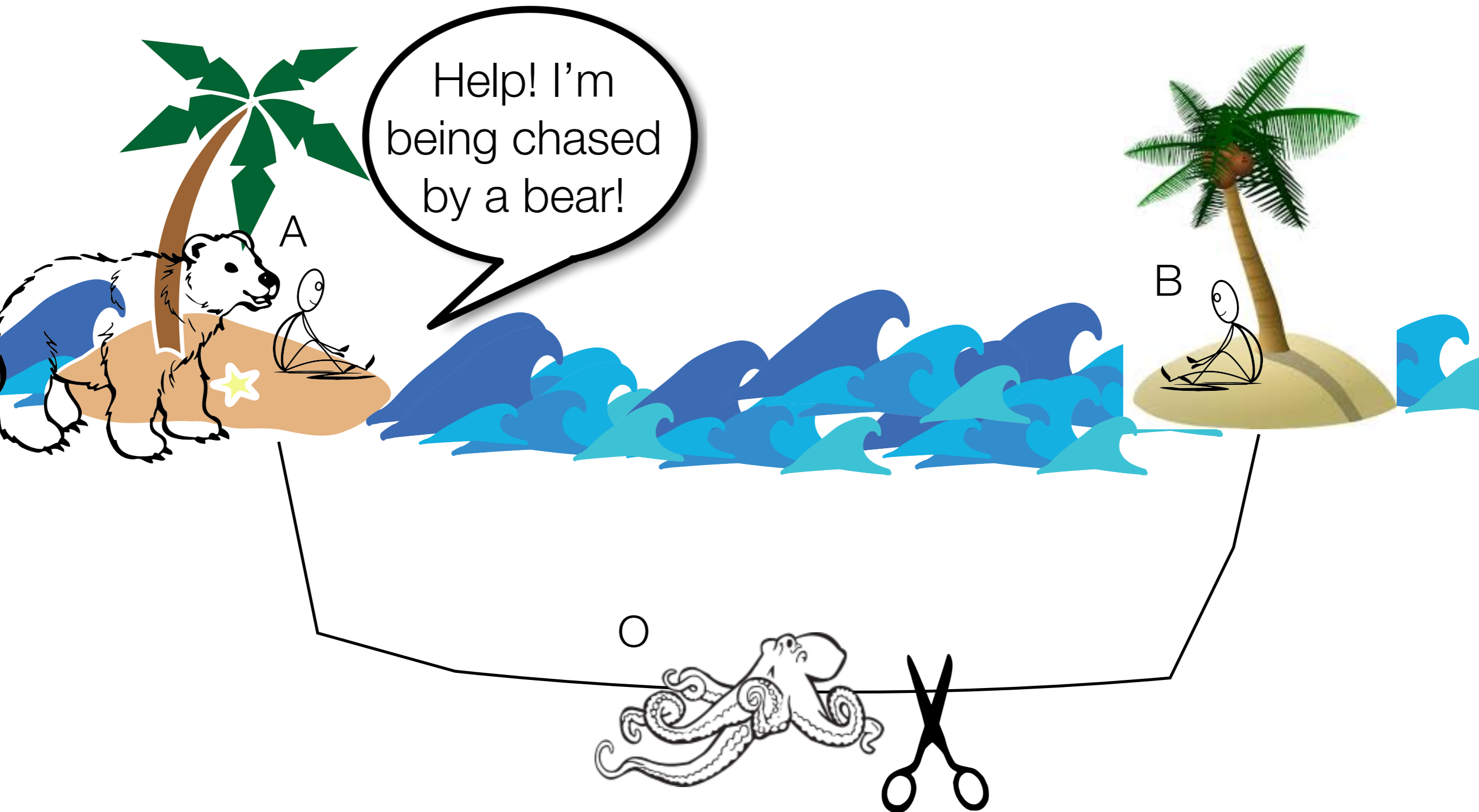
Thought experiment: Meaning from form alone



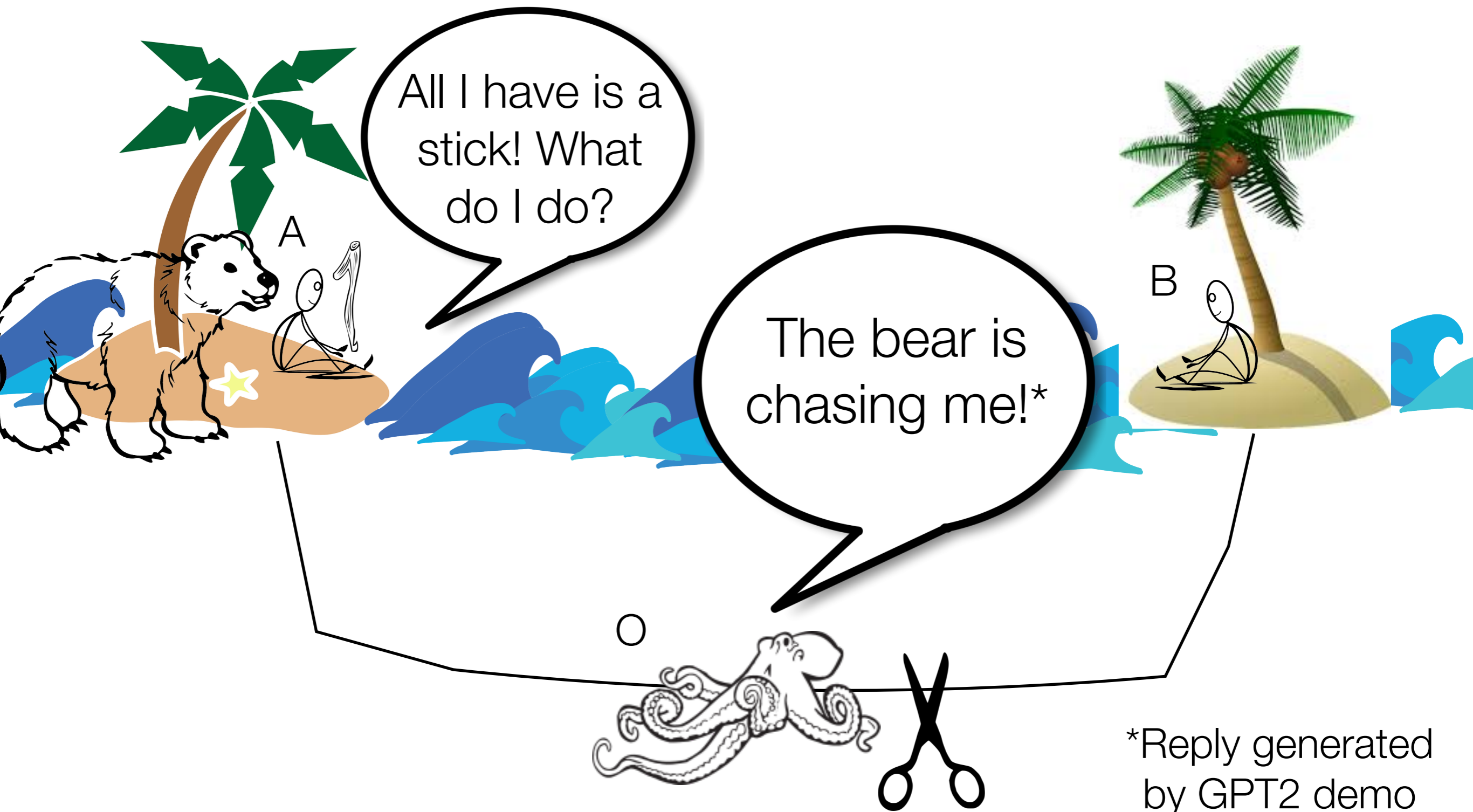
Thought experiment: Meaning from form alone



Thought experiment: Meaning from form alone



Thought experiment: Meaning from form alone



Thought experiment: Meaning from form alone



So how do babies learn language?



- Interaction is key: Exposure to a language via TV or radio alone is not sufficient (Snow et al 1976, Kuhl 2007)
- Interaction allows for joint attention: where child and caregiver are attending to the same thing and mutually aware of this fact (Baldwin 1995)
- Experimental evidence shows that more successful joint attention leads to faster vocabulary acquisition (Tomasello & Farrar 1986, Baldwin 1995, Brooks & Meltzoff 2005)
- Meaning isn't in form; rather, languages are rich, dense ways of providing cues to communicative intent (Reddy 1979). Once we learn the systems, we can use them in the absence of co-situatedness.

Meanwhile...

- In 2016, CLMS Advisory Board member Lesley Carmichael suggested that the curriculum should include something on ethics
- First taught “Ethics and NLP” (now “Societal Impacts of Language Technology”) seminar in Winter 2017, asking:
 - What ethical considerations arise in the design and deployment of NLP technologies?
 - Which of these are specific to NLP (as opposed to AI or technology more generally?)
 - What best practices can/should NLP developers deploy in light of the ethical concerns identified?

What can go wrong? Just about everything...

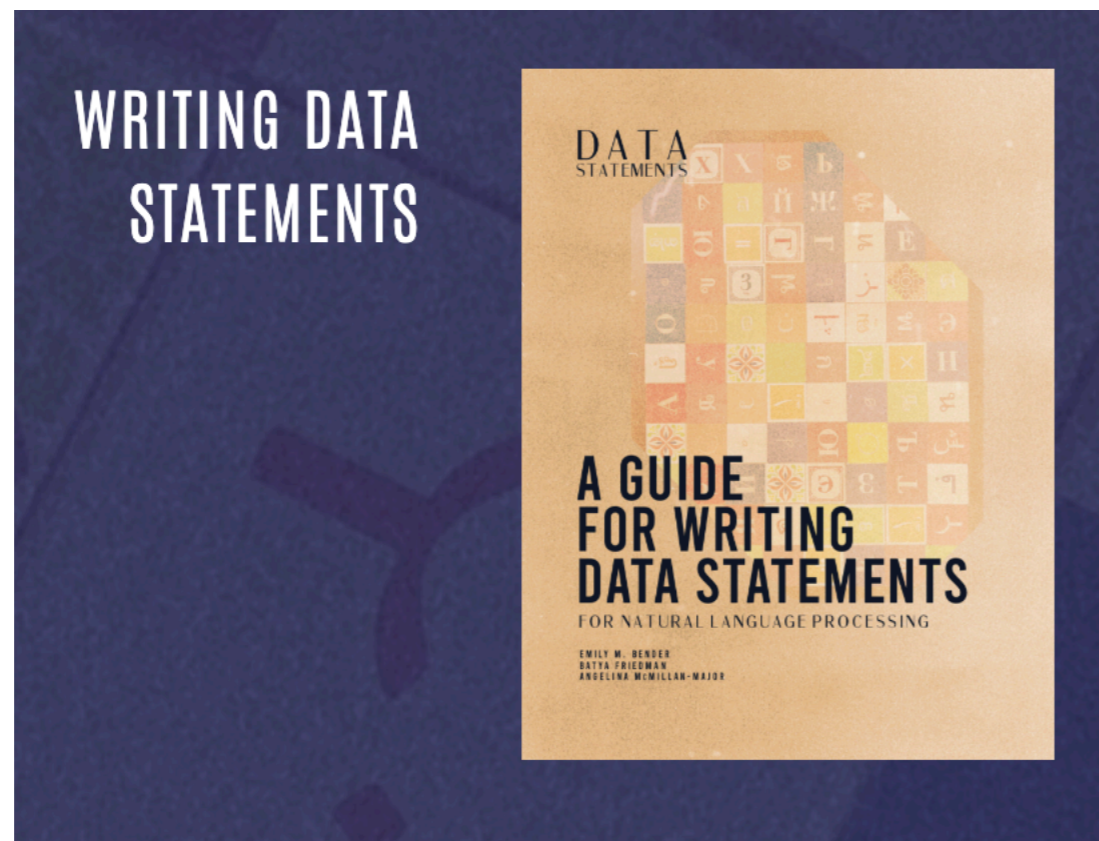
- Language models pick up bias from training data & amplify it
 - Biases come out in search engine results (Noble 2018), machine translation output (Caliskan & Lewis 2021), & more
- Systems trained for prestige varieties don't work as well for non-standard/stigmatized varieties (e.g. Wassink et al 2022, Sap et al 2019)
- ... but better coverage of language varieties enables more surveillance (Bender & Grissom II, 2024)
- Data annotation processes often involve exploitative labor practices (Fort et al 2011; see also news coverage in 2023 by Billy Perrigo and Karen Hao)

What can we do?

- Value sensitive design (Friedman & Hendry 2019): collection of methodologies for
 - Identifying stakeholders (direct and indirect)
 - Eliciting stakeholder values
 - Designing to support stakeholder values
- Values of transparency and recourse + knowledge of sociolinguistic variation
=> Data Statements for Natural Language Processing
- See also: Gebru et al 2021, Mitchell et al 2019, Arnold et al 2019, Stoyanovich & Howe 2019

Data Statements for Natural Language Processing

- Bender & Friedman 2018, Bender et al 2021, McMillan-Major et al 2023, McMillan-Major & Bender forthcoming



SCHEMA

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

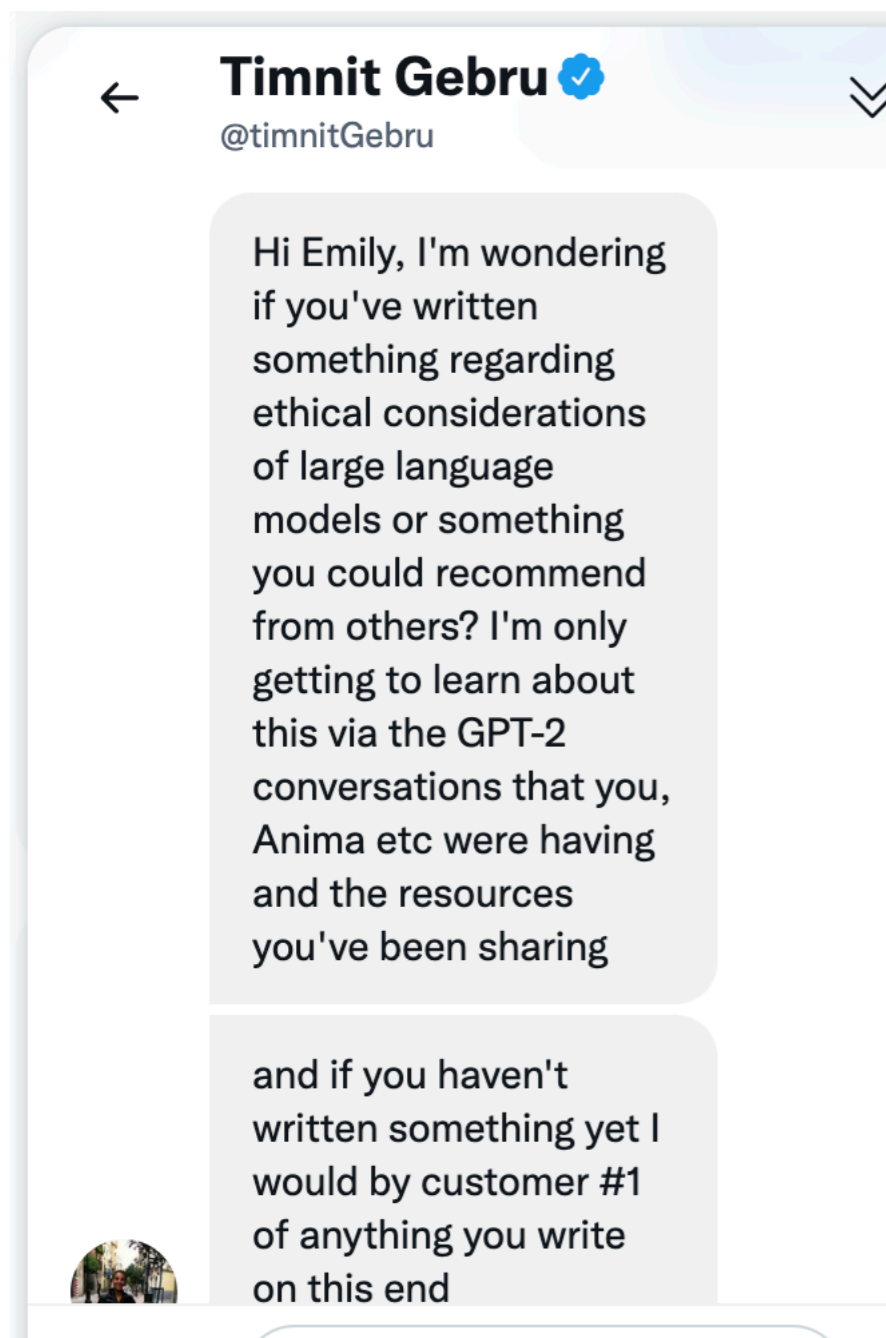
14 OTHER

15 GLOSSARY

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

2020: Origin story of the Stochastic Parrots paper

- Started off as a Twitter DM conversation, with Dr. Timnit Gebru:



This is something I'm trying to advocate for at Google

sent them some of your tweets as well haha

so many conversations about how we're not leading in large language models and should

GPT-3 is so impressive etc

and each time I'm like ANND see what Emily has to say

I can think of three other angles around ethical implications of GPT-3 and the like:

- 1) Carbon cost of creating the damn things (see Strubell et al at ACL 2019)
- 2) AI hype/people claiming it's understanding when it isn't (Bender & Koller at ACL 2020)
- 3) Deepfakes/random generated text that no one is accountable for but which is interpreted as meaningful.

Sep 8, 2020, 4:54 PM ✓

Bender, Gebru et al 2021

On the Dangers of Stochastic Parrots: Can Language Models be too big? 🦜



- *Prabhakaran*: Prabhakaran et al 2012, Prabhakaran & Rambow 2017, Hutchison et al 2020
- *Hutchinson*: Hutchinson et al 2019, 2020, 2021
- *Díaz*: Lazar et al 2017, Díaz et al 2018



History & context of this paper

- October 2020: Paper goes through Google-internal “pup-approve” before submission to FAccT
- Late November: Google asks Dr. Gebru to either retract the paper or remove the Google co-authors’ names from it
- Dr. Gebru pushes back, asking for information on what exactly was being objected to and objecting to how she & her team were being treated
- December 2, 2020: Google fires Dr. Gebru
- Dr. Margaret Mitchell starts documenting what happened to Dr. Gebru and calling on people within Google to apologize & fix systems
- February 19, 2021: Google fires Dr. Mitchell

History & context of this paper

- Google's actions led to intense media interest, both about their treatment of Dr. Gebru (and eventually Dr. Mitchell) and about our research
 - Selected media coverage
- Someone leaked the “pub-approve” version of the paper to Reddit
- FAccT 2021 primary reviewers were complete before the media story broke (preserving anonymous review)
- FAccT 2021 acceptances announced on December 22, 2020
- Camera ready due January 22, 2021

Risks of large language models (LLMs)



- Environmental impact (including environmental racism angle)
- Financial inaccessibility (impact on participation in research)
- Unmanageably large training data (no, the Web isn't representative)
- Reproduction and amplification of bias
- Misdirected research effort
- Harms of synthetic media

Stochastic Parrots

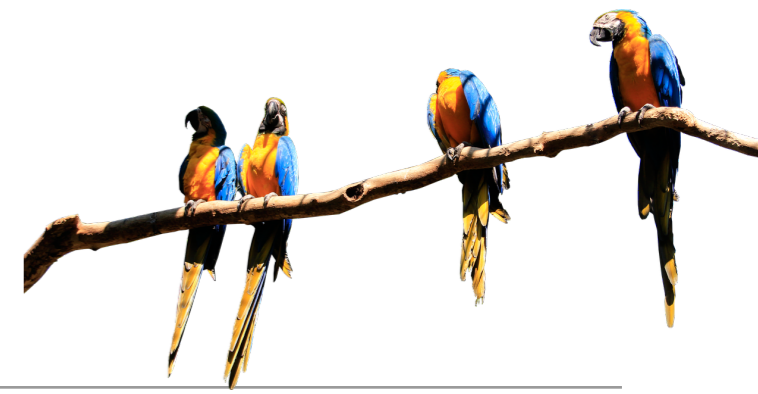


An LM is a system for haphazardly stitching together linguistic forms from its vast training data, without any reference to meaning: a stochastic parrot.

Stochastic Parrots coda (2024 update)

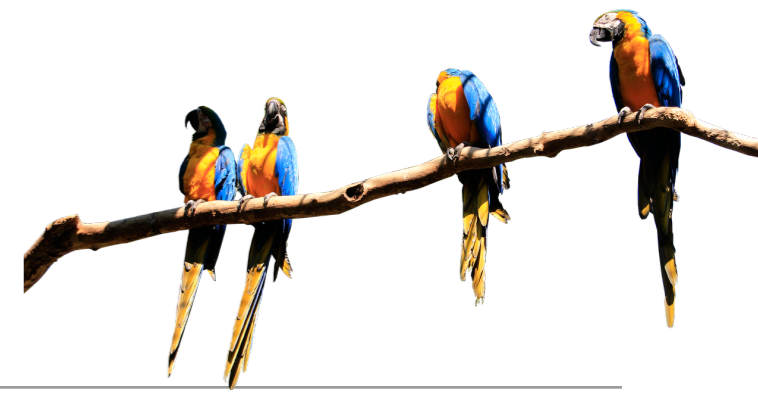
- "How do you feel now that your predictions have come true?"
- Those weren't predictions, they were warnings!
- What we didn't predict/notice at the time:
 - Exploitative labor practices
 - Just how enthusiastic people would be about synthetic text
 - Pollution of the information ecosystem (Shah & Bender 2023)
 - The transition to treating LLMs as “everything machines”, i.e. an “unscoped technology” (Gebru & Torres 2023)

Stochastic Parrots



An LM is a system for haphazardly stitching together linguistic forms from its vast training data, without any reference to meaning: a stochastic parrot.

Stochastic parrots and what we know about how language works



- Language as a system of signs (pairings of form and meaning; de Saussure 1959)
- Language use as a joint activity, involving intersubjectivity (Clark 1996, Baldwin 1995)
- => Non-understanding of LLMs is obvious
- But there's a huge need to share this understanding
- ... and increasingly so as venture capital is pushing the use of LLMs in every aspect of modern life

Public scholarship: AI hype takedowns

- Started with Twitter threads, then blog posts
- Mostly focused on media coverage of “AI” claims
- Collected here: <https://bit.ly/antiAIhype>

=> Watching for and calling out implausible claims

=> Close reading of how people are talking about “AI” and where the discourse goes astray

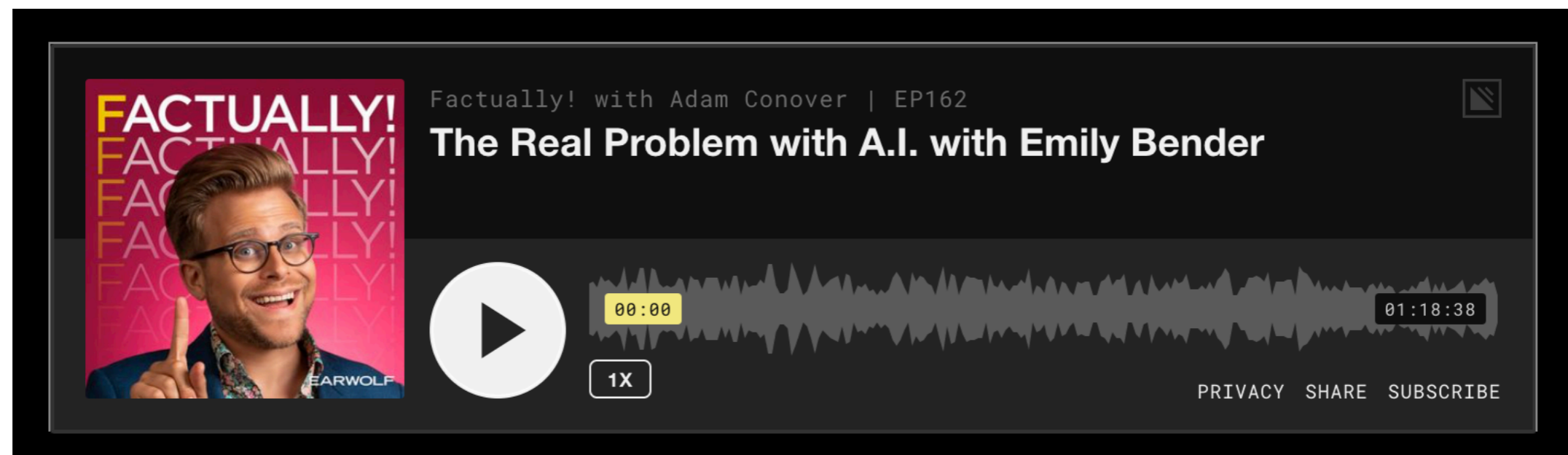
Public scholarship: AI hype takedowns



On NYT Magazine on AI: Resist the Urge to be Impressed

22 min read · Apr 17, 2022

 2.2K  24



FACTUALLY! Factly with Adam Conover | EP162
The Real Problem with A.I. with Emily Bender

00:00 01:18:38

1X

PRIVACY SHARE SUBSCRIBE

An accidental podcast (with Alex Hanna)



<https://www.dair-institute.org/maiht3k/>

November 2022: (Galactica then) ChatGPT



Now LLM talk wasn't just in tech circles

- Intense media attention, now from the non-tech press
- Approached by Liz Weil for her *New York Magazine* piece
- Policymakers (city, federal, UN, IMF)

Can feel like saying the same thing over and over again, but:

- Each new media contact is a chance to reach a new audience
- Continually working on how to make the message resonate
- Continually learning from other scholars and activists in the “AI” ethics space
- Continually asking: What specifically am I bringing to this conversation
 - based on what I know about how language works

Linguistics is good for more than combatting AI hype

- AI hype is all over the news right now, but it's far from the only arena where the world can benefit from what we know about how language works
- Public scholarship can take many forms, and doesn't have to involve engaging on social media or talking with reporters
- Sometimes the relevant public is a small, tight-knit community
- Sometimes it's teachers and/or parents
- There's a long tradition of public scholarship in linguistics

Some of my linguistics heroes

- Dr. Anne Charity Hudley, redefining spaces from the K-12 classroom, to college campuses, to the LSA to make them truly inclusive
- Dr. Wesley Leonard, developing approaches to language reclamation and language science (teaching) that center an Indigenous perspectives
- Dr. John Rickford & Dr. Sharese King, finding interventions for when the legal system can't "hear" vernacular speakers
- Dr. Kirby Conrod, providing practical pronoun guidance, showing folks they can in fact get this right
- Dr. Kelly Elizabeth Wright, using sociolinguistics to dismantle institutionalized racism

An explicitly political activity

- My date of entry into the ethics and NLP space is not coincidental (2016)
 - I was asking myself: How can I use my expertise to do something meaningful?
- Explicitly political moves can be scary
- But scholarship is inherently political—to say otherwise is also a (covert) political move

An explicitly political activity

- The more privilege one has, the more easily one can lay claim to the apocryphal “view from nowhere”
- ... and the more tempting it is to stay there.
- But it's worth it to take the leap

What are you doing or will you do with what you know about how language works?

References

- Arnold, M., Bellamy, R. K., Hind, M., Houde, S., Mehta, S., Mojsilović, A., Nair, R., Ramamurthy, K. N., Olteanu, A., Piorkowski, D., et al. (2019). Factsheets: Increasing trust in AI services through supplier’s declarations of conformity. *IBM Journal of Research and Development*, 63(4/5):6:1–6:13.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. In Moore, C. and Dunham, P. J., editors, *Joint Attention: Its Origins and Role in Development*, pages 131–158. Psychology Press.
- Bender, E. M. (2001). *Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence*. PhD thesis, Stanford University.
- Bender, E. M. (2013). *Linguistic Fundamentals for Natural Language Processing: 100 Essentials from Morphology and Syntax*. Morgan & Claypool.
- Bender, E. M. (2019). The #BenderRule: On naming the languages we study and why it matters. *The Gradient*. <https://thegradient.pub/the-benderrule-on-naming-the-languages-we-study-and-why-it-matters/>.
- Bender, E. M., Drellishak, S., Fokkens, A., Goodman, M. W., Mills, D. P., Poulson, L., and Saleem, S. (2010). Grammar prototyping and testing with the LinGO Grammar Matrix customization system. In *Proceedings of the ACL 2010 System Demonstrations*, pages 1–6, Uppsala, Sweden. Association for Computational Linguistics.
- Bender, E. M., Flickinger, D., and Oepen, S. (2002). The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars. In Carroll, J., Oostdijk, N., and Sutcliffe, R., editors, *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*, pages 8–14, Taipei, Taiwan.
- Bender, E. M., Freidman, B., and McMillan-Major, A. (2021a). A guide for writing data statements for natural language processing.
- Bender, E. M. and Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics*, 6:587–604.
- Bender, E. M., Gebru, T., McMillan-Major, A., Shmitchell, S., and et al (2021b). On the dangers of stochastic parrots: Can language models be too big? 🦜. In *Proceedings of FAccT 2021*.
- Bender, E. M. and Grissom II, A. (2024). Power shift: Toward inclusive natural language processing. In *Inclusion in Linguistics*. Oxford University Press.
- Bender, E. M. and Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5185–5198, Online. Association for Computational Linguistics.
- Bender, E. M. and Lascarides, A. (2019). *Linguistic Fundamentals for Natural Language Processing II: 100 Essentials from Semantics and Pragmatics*. Morgan & Claypool.
- Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., and Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. In Lee, D. D., Sugiyama, M., Luxburg, U. V., Guyon, I., and Garnett, R., editors, *Advances in Neural Information Processing Systems 29*, pages 4349–4357. Curran Associates, Inc.
- Brooks, R. and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8(6):535–543.
- Caliskan, A., Bryson, J. J., and Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186.
- Caliskan, A. and Lewis, M. (2021). Social biases in word embeddings and their relation to human cognition. In Deghani, M. and Boyd, R., editors, *The Handbook of Language Analysis in Psychology*. Guilford Press.
- de Saussure, F. (1959). *Course in General Linguistics*. The Philosophical Society, New York. Translated by Wade Baskin.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional

- transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Díaz, M., Johnson, I., Lazar, A., Piper, A. M., and Gergle, D. (2018). Addressing age-related bias in sentiment analysis. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 114. Association for Computing Machinery, New York, NY, USA.
- Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*.
- Flickinger, D. (2000). On building a more efficient grammar by exploiting types. *Natural Language Engineering*, 6(1 Special Issue on Efficient Processing with HPSG):15–28.
- Flickinger, D. (2011). Accuracy v. robustness in grammar engineering. In Bender, E. M. and Arnold, J. E., editors, *Language from a Cognitive Perspective: Grammar, Usage and Processing*, pages 31–50. CSLI Publications, Stanford, CA.
- Fort, K., Adda, G., and Cohen, K. B. (2011). Last words: Amazon Mechanical Turk: Gold mine or coal mine? *Computational Linguistics*, 37(2):413–420.
- Friedman, B. and Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.
- Garg, N., Schiebinger, L., Jurafsky, D., and Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16):E3635–E3644.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., and Crawford, K. (2021). Datasheets for datasets. *Commun. ACM*, 64(12):8692.
- Gebru, T. and Torres, É. P. (2023). Eugenics and the promise of utopia through artificial general intelligence. Talk presented by Gebru at SaTML 2023. Recording available at <https://www.youtube.com/watch?v=P7XT4TWLzJw>.
- Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3):146–162.
- Herbelot, A., von Redecker, E., and Müller, J. (2012). Distributional techniques for philosophical enquiry. In *Proceedings of the 6th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 45–54, Avignon, France. Association for Computational Linguistics.
- Hutchinson, B., Pittl, K. J., and Mitchell, M. (2019). Interpreting social respect: A normative lens for ML models. *CoRR*, abs/1908.07336.
- Hutchinson, B., Prabhakaran, V., Denton, E., Webster, K., Zhong, Y., and Denuyl, S. (2020). Social biases in NLP models as barriers for persons with disabilities. *CoRR*, abs/2005.00813.
- Hutchinson, B., Smart, A., Hanna, A., Denton, E., Greer, C., Kjartansson, O., Barnes, P., and Mitchell, M. (2021). Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’21, page 560575, New York, NY, USA. Association for Computing Machinery.
- Kuhl, P. K. (2007). Is speech learning ‘gated’ by the social brain? *Developmental Science*, 10(1):110–120.
- Lazar, A., Díaz, M., Brewer, R., Kim, C., and Piper, A. M. (2017). Going gray, failure to hire, and the ick factor: Analyzing how older bloggers talk about ageism. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW ’17, page 655668, New York, NY, USA. Association for Computing Machinery.
- McMillan-Major, A. and Bender, E. M. (forthcoming). A guide for creating and documenting language datasets with data statements schema version 3.
- McMillan-Major, A., Bender, E. M., and Friedman, B. (2023). Data statements: From technical concept to community practice. *ACM J. Responsib. Comput.* Just Accepted.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., and Gebru, T. (2019). Model cards for model reporting. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* ’19, pages 220–229, New York, NY, USA. ACM.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press.
- Prabhakaran, V. and Rambow, O. (2017). Dialog structure through the lens of gender, gender environment, and power. *Dialogue & Discourse*, 8(2):21–55.

- Prabhakaran, V., Rambow, O., and Diab, M. (2012). Predicting overt display of power in written dialogs. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 518–522, Montréal, Canada. Association for Computational Linguistics.
- Raji, I. D., Scheuerman, M. K., and Amironesei, R. (2021). You can’t sit with us: Exclusionary pedagogy in ai ethics education. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’21, page 515525, New York, NY, USA. Association for Computing Machinery.
- Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In *Metaphor and Thought*, pages 164–201.
- Sap, M., Card, D., Gabriel, S., Choi, Y., and Smith, N. A. (2019). The risk of racial bias in hate speech detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1668–1678, Florence, Italy. Association for Computational Linguistics.
- Shah, C. and Bender, E. M. (2023). Envisioning information access systems: What makes for good tools and a healthy web? Under review. Available from: https://bit.ly/Env_IAS.
- Snow, C. E., Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., and Vorster, J. (1976). Mothers’ speech in three social classes. *Journal of Psycholinguistic Research*, 5(1):1–20.
- Stoyanovich, J. and Howe, B. (2019). Nutritional labels for data and models. *A Quarterly bulletin of the Computer Society of the IEEE Technical Committee on Data Engineering*, 42(3).
- Tomasello, M. and Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57(6):1454–1463.
- Veres, C. (2022). Large language models are not models of natural language: They are corpus models. *IEEE Access*, 10:61970–61979.
- Wassink, A. B., Gansen, C., and Bartholomew, I. (2022). Uneven success: Automatic speech recognition and ethnicity-related dialects. *Speech Communication*, 140:50–70.
- Zamaraeva, O., Curtis, C., Emerson, G., Fokkens, A., Goodman, M., Howell, K., Trimble, T., and Bender, E. M. (2022). 20 years of the Grammar Matrix: Cross-linguistic hypothesis testing of increasingly complex interactions. *Journal of Language Modelling*, 10(1):49137.