Assembling Syntax: Modeling Constituent Questions in a Grammar Engineering Framework

Olga Zamaraeva

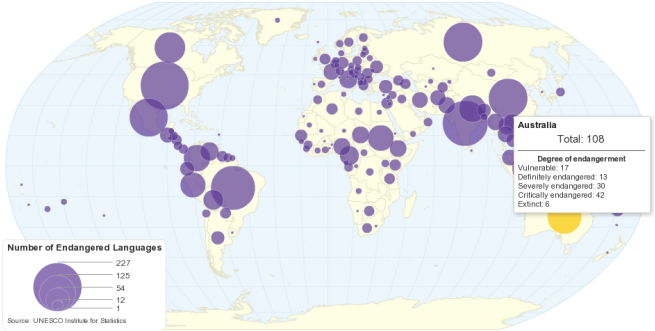
Department of Linguistics, University of Washington

SigTyp Lecture Series

May 14 2021

Range of language variation

- There are over 7000 languages in the world[1]
- 90% of people speak about 10% of them
  - Many languages are being spoken less and less and may disappear

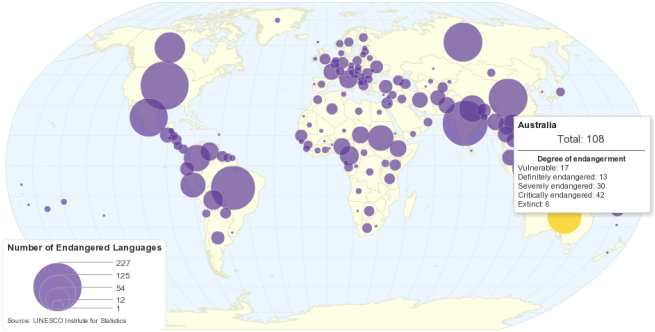


http://chartsbin.com/view/1339

1 *Language Files 12 ed.*

Understanding the range of language variation is important:

- For culture and society
- For science
  - One of the **fundamental goals of linguistics**
  - What about natural language processing?



http://chartsbin.com/view/1339

Natural Language Processing

Natural Language Processing

- Learn models from data

Natural Language Processing

- Learn models from data
- Use models to:

Natural Language Processing

- Learn models from data
- Use models to:
  - Perform language tasks

Natural Language Processing

- Learn models from data
- Use models to:
  - Perform language tasks
  - Learn something about faculties involved in those tasks

Natural Language Processing

- Learn models from data
- Use models to:
  - Perform language tasks
  - Learn something about faculties involved in those tasks
  - "Learn something about the world"[2] through analyzing language data

[2] **Mark Yatskar**; **p.c.**

Natural Language Processing

- Learn models from data
- Use models to:
  - Perform language tasks
  - Learn something about faculties involved in those tasks
  - "Learn something about the world"[2] through analyzing language data

[2] **Mark Yatskar**; **p.c.**

Natural Language Processing

- Learn models from data
- Use models to:
  - Perform language tasks
  - Learn something about faculties involved in those tasks
  - "Learn something about the world"[2] through analyzing language data
    - Knowledge bases?

[2] **Mark Yatskar; p.c.**

Does NLP care about range of language variation?

- Linguists:

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:
  - For learning about language faculties:

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:
  - For learning about language faculties:
    - Probably care, similar to like linguists do

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:
  - For learning about language faculties:
    - Probably care, similar to like linguists do
  - For performing tasks:

Does NLP care about range of language variation?

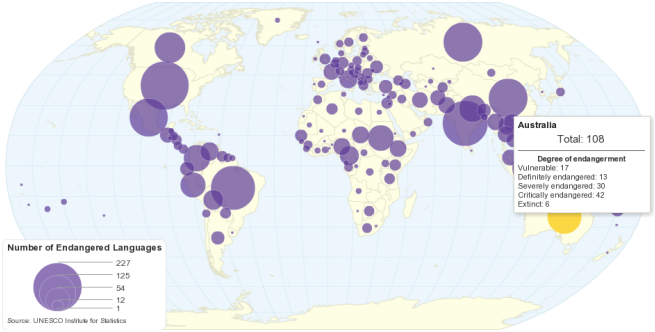
- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:
  - For learning about language faculties:
    - Probably care, similar to like linguists do
  - For performing tasks:
    - Maybe we care, but not if we can successfully transfer models from a few high-resource languages?

Does NLP care about range of language variation?

- Linguists:
  - Naturally care: Finding/describing/analyzing range of variation is a fundamental goal
- NLP:
  - For learning about language faculties:
    - Probably care, similar to like linguists do
  - For performing tasks:
    - Maybe we care, but not if we can successfully transfer models from a few high-resource languages?
    - Sure, but how do we **evaluate** the performance?

Understanding the range of language variation is important:

- For culture and society
- For science
  - One of the **fundamental goals of linguistics**
  - Core to evaluating NLP models
  - Can transfer a model from English to another language but still need systematic knowledge about that other language to evaluate



http://chartsbin.com/view/1339

We need to study more languages...

- To learn more about human language, via linguistics or via NLP
  - ...just like other sciences do in their domains [3]

Pongo tapanuensis. Pic credit: Tim Laman

[3] Nater et al. 2017

We need to study more languages...

- Studying a small set of languages is also valuable! (In both linguistics and NLP!)
- ...but it biases the questions we ask [4]

Pic credit: https://englishlive.ef.com/blog/language-lab/question-words/

[4] Idiatov 2007; Bybee 1985

How to study languages?!

- Linguistic typology

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power
    - Have a model of e.g. syntax s.t. the grammar generates only correct sentences

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power
    - Have a model of e.g. syntax s.t. the grammar generates only correct sentences
- Traditionally separate

How to study languages?!

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power
    - Have a model of e.g. syntax s.t. the grammar generates only correct sentences
- Traditionally separate
  - How do we combine them?

Combining breadth and depth

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power
    - Have a model of e.g. syntax s.t. the grammar generates only correct sentences

Combining breadth and depth

- Linguistic typology
  - Study range of variation wrt broad characteristics
  - Pay special attention to diversity
- Syntactic (semantic, morphological, phonological, discourse...) theory
  - In-depth inquiry into modeling phenomena with generative power
    - Have a model of e.g. syntax s.t. the grammar generates only correct sentences
- Computational modeling
  - of the theory
  - ...for reproducibility and rigor
  - ...when assembling fragments of in-depth analyses
  - ...into a typologically diverse system

Computational syntax

- Implement grammars on the computer 👩‍💻

Computational syntax

- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄

Computational syntax

- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄
  - ...as many sentences as you have 📄📄📄📄

Computational syntax

- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄
  - ...as many sentences as you have 📄📄📄📄
  - ...from as many languages as you have data from

Computational syntax

- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄
  - ...as many sentences as you have 📄📄📄📄
  - ...from as many languages as you have data from
  - ...including typologically diverse languages 🌏

Computational syntax

- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄
  - ...as many sentences as you have 📄📄📄📄
  - ...from as many languages as you have data from
  - ...including typologically diverse languages 🌏
  - ...as many times as you need

Computational syntax

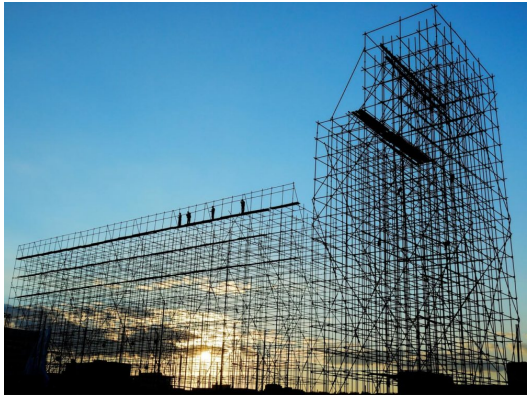
- Implement grammars on the computer 👩‍💻
- Run grammars automatically on sentences 📄
  - ...as many sentences as you have 📄📄📄📄
  - ...from as many languages as you have data from
  - ...including typologically diverse languages 🌏
  - ...as many times as you need
- Grow grammars and accumulate knowledge artifacts

Computational syntax

- Implement grammars on the computer
- Run grammars automatically on sentences
  - ...as many sentences as you have
  - ...from as many languages as you have data from
  - ...including typologically diverse languages
  - ...as many times as you need
- Grow grammars and accumulate knowledge artifacts
- Growing the area of applicability of a set of hypotheses which grammars represent

Philosophy: Method of fragments

- Fully explicit **grammar fragments**[5] that can be extended
  - constitute research artifacts that can be literally built upon
  - together and over time, contribute to our understanding of syntax



https://www.theinformationlab.co.uk/2017/08/09/data-scaffolding-easy-steps-fill-missing-data/

5 Montague 1974

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:
  - Typology

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:
  - Typology
  - Syntactic, semantic, phonological, acquisition, and other theories

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:
  - Typology
  - Syntactic, semantic, phonological, acquisition, and other theories
  - Computational linguistics

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:
  - Typology
  - Syntactic, semantic, phonological, acquisition, and other theories
  - Computational linguistics
  - Natural language processing

Wrapping up the big picture

- Big Question: What is the range of variation in human languages? 
- Many ways to approach and many relevant fields:
  - Typology
  - Syntactic, semantic, phonological, acquisition, and other theories
  - Computational linguistics
  - Natural language processing
- All of these fields are probably invested in understanding the range of variation in human language

Wrapping up the big picture

- Big Question: What is the range of variation in human languages?
- Many ways to approach and many relevant fields:
  - Typology
  - Syntactic, semantic, phonological, acquisition, and other theories
  - Computational linguistics
  - Natural language processing
- All of these fields are probably invested in understanding the range of variation in human language
  - Next: Combining syntactic theory and typology

Assembling Syntax: Part II

Computational syntax with Head-Driven Phrase Structure Grammar formalism

Computational syntax

- Implement **grammars** on the computer
- Run grammars automatically on sentences
  - ...as many sentences as you have
  - ...from as many languages as you have data from
  - ...including typologically diverse languages
  - ...as many times as you need
- Grow grammars and accumulate **knowledge artifacts**
- Growing the area of applicability of a **set of hypotheses** which grammars demonstrably represent

What is a grammar?

- Definition 1:
  - A set of rules which can be loaded into a parser to parse/generate sentences

What is a grammar?

- Definition 1:
  - A set of rules which can be loaded into a parser to parse/generate sentences
- Definition 2:
  - A set of rules which can be loaded into a parser to parse/generate sentences...
  - ...such that it gives **insight** into human linguistic faculties

What is a grammar?

- Definition 1:
  - A set of rules which can be loaded into a parser to parse/generate sentences
- Definition 2:
  - A set of rules which can be loaded into a parser to parse/generate sentences...
  - ...such that it gives **insight** into human linguistic faculties
- I am working somewhere in between...

Why talk about grammars to an NLP audience?

- Grammars aren't used very much in contemporary NLP

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary 🕸 NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD
  - ...someone 👩‍🔬 had to systematize, formalize, and record the knowledge **first**

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD
  - ...someone 👩‍🔬 had to systematize, formalize, and record the knowledge **first**
  - Hard to update/improve and does not scale

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD
  - ...someone 👩‍🔬 had to systematize, formalize, and record the knowledge **first**
  - Hard to update/improve and does not scale
  - ...unless you **use an implemented grammar**

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD
  - ...someone 👩‍🔬 had to systematize, formalize, and record the knowledge **first**
  - Hard to update/improve and does not scale
  - ...unless you **use an implemented grammar**
  - **Goal:** Do this for more languages and in a more systematic and easily update-able way

Why talk about grammars 📚 to an NLP audience?

- Grammars 📚 aren't used very much in contemporary ⚛ NLP
  - ...or are they? ☝
- Evaluation relies on annotation 👩‍🔬
- Many types of annotation rely on systematized knowledge 📚
  - Grammars are a type of systematized knowledge
  - e.g. PennTree bank or UD
  - ...someone 👩‍🔬 had to systematize, formalize, and record the knowledge **first**
  - Hard to update/improve and does not scale
  - ...unless you **use an implemented grammar**
  - **Goal:** Do this for more languages and in a more systematic and easily update-able way
  - Rely on indispensable human knowledge but support it with computational framework

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem
- Computational syntax:

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem
- Computational syntax:
  - Consider sentences (usually a few), hypothesize an analysis

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem
- Computational syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Implement the analysis in a grammar engineering framework 👩‍💻

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem
- Computational syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Implement the analysis in a grammar engineering framework 👩‍💻
  - Test on many sentences 📄📄📄📄in literal interaction 📜with analyses for other phenomena

Assembling grammars systematically

- Pen-and-paper syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Test in your head
  - Revise if you notice a problem
- Computational syntax:
  - Consider sentences (usually a few), hypothesize an analysis
  - Implement the analysis in a grammar engineering framework 👩‍💻
  - Test on many sentences 📄📄📄📄in literal interaction 📝with analyses for other phenomena
  - Revise hypotheses

Assembling grammars systematically

- **Goal**: Make implemented linguistic grammars bigger and more accessible to broader research community

Assembling grammars systematically

- **Goal**: Make implemented linguistic grammars bigger and more accessible to broader research community
- **Method**: Meta-grammar engineering with Head-Driven Phrase Structure Grammar

Assembling grammars systematically

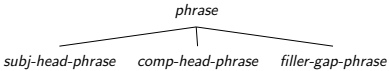
- **Goal**: Make implemented linguistic grammars bigger and more accessible to broader research community
- **Method**: Meta-grammar engineering with Head-Driven Phrase Structure Grammar
- **Project**: Analysis of constituent questions for a grammar engineering system

Assembling grammars systematically

- **Goal**: Make implemented linguistic grammars bigger and more accessible to broader research community
- **Method**: Meta-grammar engineering with Head-Driven Phrase Structure Grammar
- **Project**: Analysis of constituent questions for a grammar engineering system
- **Result**:
  - New library in the system; more complex hypotheses can be tested
  - Archived test suites and analyses for several languages
  - Some takeaways regarding the interaction of different analyses

Head-Driven Phrase Structure Grammar

- **Fully explicit** formalism[6]
- Lexicalist and surface-oriented
- A grammar is a hierarchy of **types** encoded as **feature structures** where features are constrained to have some values
- A structure licensing a sentence must be **well-formed**

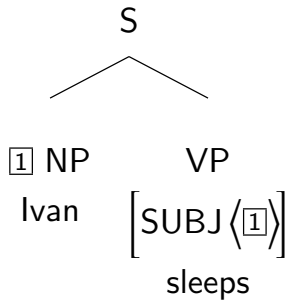


$$\begin{bmatrix} \textit{subj-head-phrase} & \\ \text{SUBJ} & \langle\rangle \\ \text{HEAD-DTR} & \boxed{2}\begin{bmatrix}\text{SUBJ} & \langle\, \boxed{1}\,\rangle\end{bmatrix} \\ \text{NON-HEAD-DTR} & \boxed{1} \\ \text{ARGS} & \langle\, \boxed{1}, \boxed{2}\,\rangle \end{bmatrix}$$

6 Pollard and I. A. Sag 1994

HPSG Phrase Structure Rule

- Describes a feature structure that is a phrase and can be visualized as a tree
- "Mother" and "daughter" nodes
- Identities (tags)



$$\begin{bmatrix} \textit{subj-head-phrase} & \\ \text{SUBJ} & \langle\rangle \\ \text{HEAD-DTR} & \boxed{2}\begin{bmatrix}\text{SUBJ} & \langle\,\boxed{1}\,\rangle\end{bmatrix} \\ \text{NON-HEAD-DTR} & \boxed{1} \\ \text{ARGS} & \langle\,\boxed{1},\,\boxed{2}\,\rangle \end{bmatrix}$$

HPSG: Theoretical and Computational

- Both types of research exist
- HPSG formalism can be used to posit multiple theories
- DELPH-IN HPSG 🐬
  - International research consortium[7]
  - Restricted version of HPSG formalism

[7] https://github.com/delph-in, http://moin.delph-in.net/wiki/

DELPH-IN main projects

- The English Resource Grammar (ERG) [8]
  - **Broad** coverage; used in NLP[9]
    - Semantic representations (ERS, MRS, DMRS) used widely for **evaluating semantic parsers**[10]

8 Flickinger 2000, 2011

9 Hajdik et al. 2019; Zamaraeva, Howell, and Rhine 2018; Buys and Blunsom 2017; Packard 2014

10 Oepen and Flickinger 2019

DELPH-IN main projects

- The English Resource Grammar (ERG) [8]
  - **Broad** coverage; used in NLP[9]
    - Semantic representations (ERS, MRS, DMRS) used widely for **evaluating semantic parsers**[10]

- **Medium**-size grammars of Japanese, Chinese, German, Spanish...[11]

8 Flickinger 2000, 2011

9 Hajdik et al. 2019; Zamaraeva, Howell, and Rhine 2018; Buys and Blunsom 2017; Packard 2014

10 Oepen and Flickinger 2019

11 Siegel et al. 2016; Fan 2018; Crysmann 2003; Marimon 2010

DELPH-IN main projects

- The English Resource Grammar (ERG) [8]
  - **Broad** coverage; used in NLP[9]
    - Semantic representations (ERS, MRS, DMRS) used widely for **evaluating semantic parsers**[10]

- **Medium**-size grammars of Japanese, Chinese, German, Spanish...[11]
- The Grammar Matrix:[12] Automated **starter** grammars; **typologically**-driven (Part III)

8 Flickinger 2000, 2011

9 Hajdik et al. 2019; Zamaraeva, Howell, and Rhine 2018; Buys and Blunsom 2017; Packard 2014

10 Oepen and Flickinger 2019

11 Siegel et al. 2016; Fan 2018; Crysmann 2003; Marimon 2010

12 Bender, Flickinger, and Oepen 2002; Bender, Drellishak, et al. 2010

DELPH-IN main projects

- The English Resource Grammar (ERG) [8]
  - **Broad** coverage; used in NLP[9]
    - Semantic representations (ERS, MRS, DMRS) used widely for **evaluating semantic parsers**[10]

- **Medium**-size grammars of Japanese, Chinese, German, Spanish...[11]
- The Grammar Matrix:[12] Automated **starter** grammars; **typologically**-driven (Part III)
  - Bootstrap grammar development for more languages[13]

8 Flickinger 2000, 2011

9 Hajdik et al. 2019; Zamaraeva, Howell, and Rhine 2018; Buys and Blunsom 2017; Packard 2014

10 Oepen and Flickinger 2019

11 Siegel et al. 2016; Fan 2018; Crysmann 2003; Marimon 2010

12 Bender, Flickinger, and Oepen 2002; Bender, Drellishak, et al. 2010

13 Bender 2010; Crowgey 2019; Inman 2019

An analysis of constituent (wh) questions in DELPH-IN HPSG

- Classic 🏛 set of questions for syntactic theory:
  - How are question words ("wh-") distributed?
  - How to represent interrogative semantics?
    - Quantification, scope, wh-words as question parameters of different clauses...
  - How to model question word fronting (4)?
  - How to model optional fronting (5)?

An analysis of constituent (wh) questions in DELPH-IN HPSG

- Classic 🏛 set of questions for syntactic theory:
  - How are question words ("wh-") distributed?
  - How to represent interrogative semantics?
    - Quantification, scope, wh-words as question parameters of different clauses...
  - How to model question word fronting (4)?
  - How to model optional fronting (5)?
- Previously:
  - Nobody had put forth an analysis of **multiple fronting** as in Slavic languages
  - Nobody had offered a **unified** HPSG account of wh-questions in typologically different languages
  - ...and tested it rigorously 🤦 with a parser
  - 14

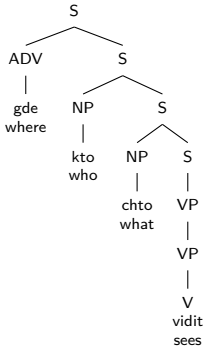
(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]

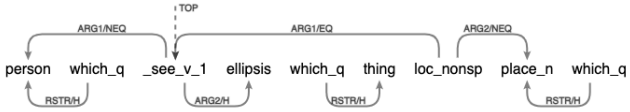
(5) Ty gde rabotaesh?
you where work.3SG
'Where do you work?'[rus]

14 Zamaraeva 2021

Questionnaire – Analysis – Semantics

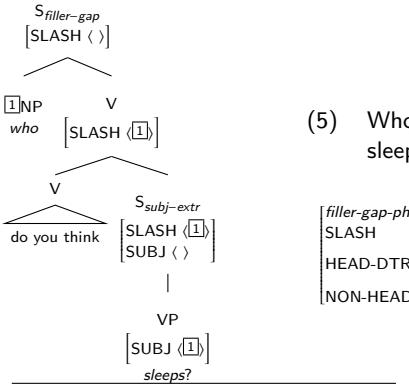
(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]





Fronting analysis in HPSG

- "Nonlocal" features are propagated up the tree[15]
- Feature SLASH creates a long-distance dependency at the level of the verb
- The *filler-gap* rule discharges the dependencies



(5) Who do you think sleeps? [eng]

$$\begin{bmatrix} \textit{filler-gap-phrase} & \\ \text{SLASH} & \langle\ \rangle \\ \text{HEAD-DTR} & \begin{bmatrix}\text{SLASH} & \langle\ \boxed{1}\ \rangle\end{bmatrix} \\ \text{NON-HEAD-DTR} & \boxed{1} \end{bmatrix}$$

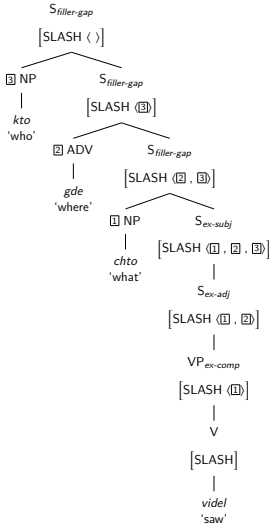
15 Pollard and I. A. Sag 1994; Ginzburg and I. Sag 2000

Extending the fragment: Multiple question fronting

- Allow multiple extraction and recursive application of *filler-gap*[16]
- Takeaways:
  - "Optional" fronting is hard![17]
  - Simpler analysis of multiple fronting → less simple morphological marking[18]

$$\begin{bmatrix} \textit{filler-gap-phrase} & \\ \text{SLASH} & \boxed{1} \\ \text{ARGS} & \left\langle \boxed{2}, \left[\text{SLASH} \left\langle \boxed{2} \right\rangle \oplus \boxed{1} \right] \right\rangle \end{bmatrix}$$

(6) kto gde chto vidit?
who.NOM where what.ACC see.3SG
'Who sees what where?'[rus]



16 Zamaraeva and Emerson 2020; Crysmann 2015
17 Zamaraeva 2021
18 Zamaraeva to appear

An analysis of Slavic fronting in HPSG

- Open questions:

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?
  - **Essential** vs. **incidental** complexity in formalisms

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?
  - **Essential** vs. **incidental** complexity in formalisms
  - **Tensions in elegance** in a **typologically diverse** system

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?
  - **Essential** vs. **incidental** complexity in formalisms
  - **Tensions in elegance** in a **typologically diverse** system
  - Can we have one framework for everything? 🌍

An analysis of Slavic fronting in HPSG

- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?
  - **Essential** vs. **incidental** complexity in formalisms
  - **Tensions in elegance** in a **typologically diverse** system
  - Can we have one framework for everything? 🌏
  - **Must** a system be elegant?

An analysis of Slavic fronting in HPSG

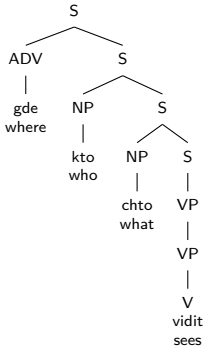
- Open questions:
  - Modeling **optional** fronting leads to **spurious** ambiguity
    - How to get rid of it while still accounting for data?
    - **Direction:** information-structure
    - **Gist:** There is **no** optional fronting!
    - Nobody has done this in DELPH-IN HPSG yet
  - But isn't this a problem with **DELPH-IN HPSG** rather than with HPSG or syntactic theory in general?
  - **Essential** vs. **incidental** complexity in formalisms
  - **Tensions in elegance** in a **typologically diverse** system
  - Can we have one framework for everything? 🌎
  - **Must** a system be elegant?
  - **Part III:** Assembling typologically diverse analyses and evaluating the result

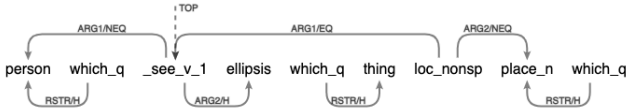
Assembling Syntax, Part III

Assembling and evaluating typologically diverse analyses

Questionnaire – Analysis – Semantics

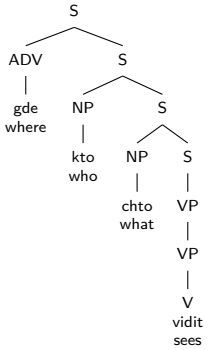
(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]





Questionnaire – Analysis – Semantics

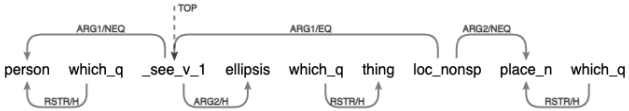
(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]



Choices regarding the position of question phrases

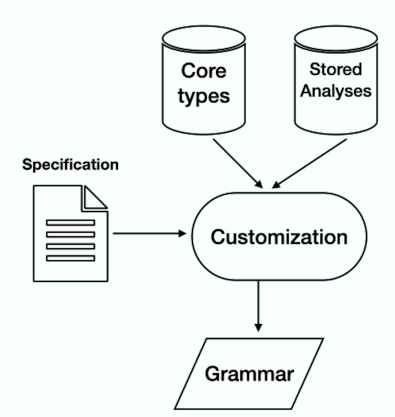
Question phrases can appear at the left edge of the sentence regardless of the position the questioned constituent would appear in (*Who did you see? I know who you saw* etc.):

- ○ Only one question phrase can be fronted
- ◉ All question phrases can be fronted
- ○ Question phrases cannot be fronted (stay *in situ*)



The Grammar Matrix

- Meta-grammar engineering framework[19]
- Input: Typological specification, lexicon, morphological rules
- Output: Implemented HPSG grammar fragment
  - Parse and generate sentences
  - Output syntactic and **semantic** representations
- Many syntactic phenomena are supported[20]

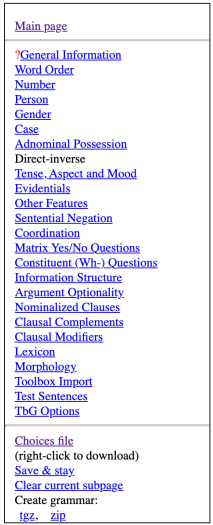


19 https://matrix.ling.washington.edu/customize/matrix.cgi

20 Zamaraeva 2021; Zamaraeva, Howell, and Bender 2019; Howell and Zamaraeva 2018; Saleem 2010; Song 2014; Nielsen 2018; Drellishak and Bender 2005; Crowgey 2013; Bender and Flickinger 2005

Matrix libraries

- Specify several phenomena at the same time
- Click to **get a grammar fragment** covering all of them[21]
- Test hypotheses **in interaction**
- Parse and **generate** data within fragment's area of coverage
  - Large lexicons can be imported



[21] There may be bugs 🐞

Matrix library development

- Mapping typological specifications to customized grammar fragments is supported by Matrix **libraries** 

Matrix library development

- Mapping typological specifications to customized grammar fragments is supported by Matrix **libraries** 
  - Questionnaires are designed based on **surveys of typological literature** 

Matrix library development

- Mapping typological specifications to customized grammar fragments is supported by Matrix **libraries** 
  - Questionnaires are designed based on **surveys of typological literature** 
  - Libraries are evaluated with **held-out languages**

Matrix library development

- Mapping typological specifications to customized grammar fragments is supported by Matrix **libraries** 
  - Questionnaires are designed based on **surveys of typological literature** 
  - Libraries are evaluated with **held-out languages**
  - Growing **regression testing** base
    - Language specs + test suites paired with "gold" semantic representations
    - Check **automatically** how any small change affects the **all** of the pairings

Matrix library development

- Mapping typological specifications to customized grammar fragments is supported by Matrix **libraries**
  - Questionnaires are designed based on **surveys of typological literature**
  - Libraries are evaluated with **held-out languages**
  - Growing **regression testing** base
    - Language specs + test suites paired with "gold" semantic representations
    - Check **automatically** how any small change affects the **all** of the pairings
- Latest addition: Constituent questions[22]

(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]

22 Zamaraeva 2021

Meta-grammar engineering research question

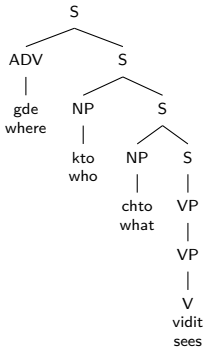
RQ: *What constitutes a model of a range of typologically attested ways of forming constituent questions within the given framework?*

Evaluation: How well does the analysis generalize to a set of randomly picked "held-out" languages?

Constituent questions in the Matrix

(4) Gde kto chto vidit?
where who.NOM what.ACC see.3SG
'Who sees what where?'[rus]

- New library[23]
- Typological scope:[24]
  - Position of question phrase
    - **Fronting**, *in situ*
    - Fronting optionality
  - Morphological marking
  - Question particles
    - position
    - obligatoriness
  - Question words



23 Zamaraeva 2021

24 König and Siemund 2007; Hagège 2008; Miyagawa 1987, *inter alia*

Assembling 🖼 fronting

Specification	Types	Core?	New?	Custom features
single oblig. front.	*wh-ques-phrase*	no	yes	SLASH
	subj-, obj-, adj-ex.	no	no	none
multi oblig. front.	*wh-ques-phrase*	no	yes	HDR|QUE MODIFIED *hasmod*
	subj-, obj-, adj-ex.	no	no	none
single opt. front.	*wh-ques-phrase*	no	yes	SLASH
	in-sutu-phrase	no	yes	none
	subj-, obj-, adj-ex.	no	no	none
multi opt. front.	*wh-ques-phrase*	no	yes	MODIFIED *hasmod*
	in-sutu-phrase	no	yes	HDR|L-QUE –
	subj-, obj-, adj-ex.	no	no	HDR|L-QUE –
in situ (no front.)	*in-sutu-phrase*	no	yes	none

The position of question phrases customization summary[25]

25 Zamaraeva 2021

Assembling particles

Specification	Types	Core?	New?	Custom features
clause-final or init.	*qpart-comp-lex*	no	no	none
	head-comp-phrase	yes	no	INIT
2nd pos.	*ques-clitic-lex*	no	yes	none
	non-local	yes	no	YNQ
	basic-binary-phrase	yes	no	L-PERIPH
	int-cl-phrase	no	yes	none
	in-situ-phrase	no	yes	none
obligatory	*in-situ-phrase*	no	yes	MC –
only in embed.	*qpart-comp-lex*	no	no	MC 1
				COMPS|MC 1
only in polar	*qpart-comp-lex*	no	no	WH|BOOL –
	ques-clitic-lex	no	yes	

Question particles customization summary[26]

26 Zamaraeva 2021

Assembling morphological marking

Specification	Types	Core?	New?	Custom features
Same mkg	*interrog-lex-rule*	no	yes	none
for polar and *wh-*	*indicative-lex-rule*	no	yes	none
Separate paradigms	*indicative-lex-rule*	no	yes	none
	polar-lex-rule	no	yes	none
	wh-subj-lex-rule	no	yes	none
	wh-obj-lex-rule	no	yes	none

Morphological question marking customization summary[27]

27 Zamaraeva 2021

Evaluation

- “Freeze” the analyses and the development

Evaluation

- “Freeze” the analyses and the development
- Grab reference grammars from “unseen” language families

Evaluation

- "Freeze" the analyses and the development
- Grab reference grammars from "unseen" language families
- Grab all examples of constituent questions from those books

Evaluation

- "Freeze" the analyses and the development
- Grab reference grammars from "unseen" language families
- Grab all examples of constituent questions from those books
- Fill out the questionnaire so as to cover the examples

Evaluation

- "Freeze" the analyses and the development
- Grab reference grammars from "unseen" language families
- Grab all examples of constituent questions from those books
- Fill out the questionnaire so as to cover the examples
- Click "create grammar", run the grammar on the examples

Evaluation

- "Freeze" the analyses and the development
- Grab reference grammars from "unseen" language families
- Grab all examples of constituent questions from those books
- Fill out the questionnaire so as to cover the examples
- Click "create grammar", run the grammar on the examples
- **Coverage**: % of examples the grammar actually covered
- **Overgeneration** % of ungrammatical examples parsed
  - covered = gave **correct** semantic representation

Evaluation

- "Freeze" the analyses and the development
- Grab reference grammars from "unseen" language families
- Grab all examples of constituent questions from those books
- Fill out the questionnaire so as to cover the examples
- Click "create grammar", run the grammar on the examples
- **Coverage**: % of examples the grammar actually covered
- **Overgeneration** % of ungrammatical examples parsed
  - covered = gave **correct** semantic representation
- **How well can the system handle examples from an "unseen" language, as it is described in the reference grammar?**

Evaluating the constituent questions library[28]

Language	ISO-639-3	Family	Gram. items	Coverage%	Question typology
Apinajé	[apn]	Macro-Jê	17	82.3	single front.
Makah	[myh]	Wakashan	14	78.5	morphological,int. verbs
Pacoh	[pac]	Austroasiatic	26	84.6	single opt. front.
Paresi-Haliti	[pab]	Arawakan	64	56.0	single front., int. verbs
Jalkunan	[bxl]	Mande	33	78.8	*in situ*, particle, int. verbs

✓ Single fronting, particles, morphological marking, interrogative verbs

✗ Question words as predicates

− Did not come up: Multiple fronting and LDDs

35 Zamaraeva 2021

Evaluating the constituent questions library[29]

Language	ISO-639-3	Family	Gram. items	Coverage%	Handled phenomena
Apinajé	[apn]	Macro-Jê	17	82.3	evidentials, arg. drop, clausal modifiers
Makah	[myh]	Wakashan	14	78.5	clausal complements, arg. drop
Pacoh	[pac]	Austroasiatic	26	84.6	arg. drop
Paresi-Haliti	[pab]	Arawakan	64	56.0	**focus**, adnom. poss, coordination
Jalkunan	[bxl]	Mande	33	78.8	adnom. poss.

- ✓ Single fronting, particles, morphological marking, interrogative verbs
- ✗ Question words as predicates
- − Did not come up: Multiple fronting and LDDs
- 🐞 Bugs: in interaction with **information structure** and clausal complements

36 Zamaraeva 2021

Interim Summary

- The Grammar Matrix facilitates grammar creation for a wide variety of languages 🌍

Interim Summary

- The Grammar Matrix facilitates grammar creation for a wide variety of languages 🌍
- Bigger grammar fragments now possible to obtain from the Matrix[30]
  - Further projects focused on polysynthetic languages are needed
  - What's next?

30 Zamaraeva 2021

Assembling syntax: Part IV

Future directions

Applying grammars in Natural Language Processing

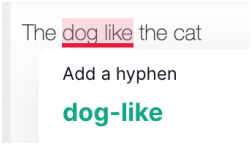
- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌍 as a long-term investment

Applying grammars in Natural Language Processing

- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages as a long-term investment
  - Meanwhile...

Applying grammars in Natural Language Processing

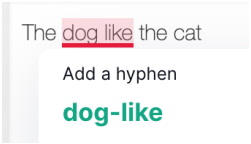
- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌏as a long-term investment
  - Meanwhile...
- NLP for grammar coaching



  - Statistical systems are imprecise ☝

Applying grammars in Natural Language Processing

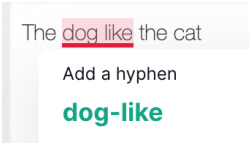
- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌏 as a long-term investment
  - Meanwhile...
- NLP for grammar coaching



  - Statistical systems are imprecise ☝
  - Adding more training data still won't help with **explanations**

Applying grammars in Natural Language Processing

- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌍 as a long-term investment
  - Meanwhile...
- NLP for grammar coaching

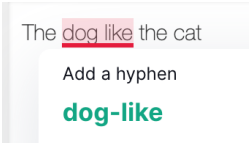


  - Statistical systems are imprecise ☝
  - Adding more training data still won't help with **explanations**
  - Grammars: Incorporate L2 productions[31]

[31] Schneider and McCoy 1998

Applying grammars in Natural Language Processing

- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌍 as a long-term investment
  - Meanwhile...
- NLP for grammar coaching

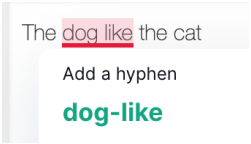


  - Statistical systems are imprecise ☝
  - Adding more training data still won't help with **explanations**
  - Grammars: Incorporate L2 productions[31]

[31] Schneider and McCoy 1998

Applying grammars in Natural Language Processing

- Ultimately, NLP relies on formal annotation
  - ERG is more robust than PTB/PSD and can be automatically rerun
  - We need this for more languages 🌍 as a long-term investment
  - Meanwhile...
- NLP for grammar coaching



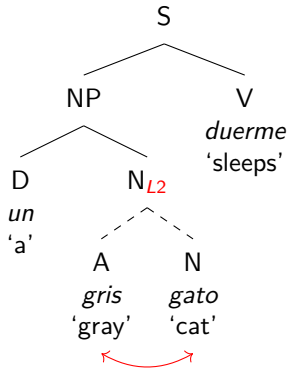
  - Statistical systems are imprecise ☝
  - Adding more training data still won't help with **explanations**
  - Grammars: Incorporate L2 productions[31]
  - Map L2 productions to **useful** feedback[32]

[31] Schneider and McCoy 1998

[32] Bender, Flickinger, Oepen, et al. 2004; Morgado da Costa, Bond, et al. 2016; Morgado da Costa, Winder, et al. 2020

HPSG grammars for grammar coaching

- First step: Spanish and Galician
  - with Gómez Rodríguez and Alonso Ramos, U. of A Coruña
- Next step: More languages



Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP

Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP
- Fully explicit formalisms like HPSG allow us to implement grammars on the computer and rigorously test them

Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP
- Fully explicit formalisms like HPSG allow us to implement grammars on the computer and rigorously test them
- **Example:** An analysis of constituent questions integrated into the Grammar Matrix system
  - Clear area of applicability as archived in the specifications, test suites, and the version of the Matrix system

Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP
- Fully explicit formalisms like HPSG allow us to implement grammars on the computer and rigorously test them
- **Example:** An analysis of constituent questions integrated into the Grammar Matrix system
  - Clear area of applicability as archived in the specifications, test suites, and the version of the Matrix system
- **Future:** Robust NLP evaluation and precision in NLP applications for more languages 

Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP
- Fully explicit formalisms like HPSG allow us to implement grammars on the computer and rigorously test them
- **Example:** An analysis of constituent questions integrated into the Grammar Matrix system
  - Clear area of applicability as archived in the specifications, test suites, and the version of the Matrix system
- **Future:** Robust NLP evaluation and precision in NLP applications for more languages
- **Future:** Encoded and tested sets of hypotheses for more languages

Conclusion

- Formal approaches to syntax are an important part of linguistics and NLP
- Fully explicit formalisms like HPSG allow us to implement grammars on the computer and rigorously test them
- **Example:** An analysis of constituent questions integrated into the Grammar Matrix system
  - Clear area of applicability as archived in the specifications, test suites, and the version of the Matrix system
- **Future:** Robust NLP evaluation and precision in NLP applications for more languages
- **Future:** Encoded and tested sets of hypotheses for more languages
- **Assembling fragments of our understanding of language**

Acknowledgments

- The work presented today was partially funded by the United States National Science Foundation under Grant No. BCS-1561833 (PI Bender).
- This work would not be possible without the DELPH-IN community

References I

Nater, Alexander et al. (2017). "Morphometric, behavioral, and genomic evidence for a new orangutan species". In: *Current Biology* 27.22, pp. 3487–3498.

Idiatov, Dmitry (2007). "A typology of non-selective interrogative pronominals". PhD thesis. University of Antwerp.

Bybee, Joan L (1985). *Morphology: A study of the relation between meaning and form*. Vol. 9. John Benjamins Publishing.

Montague, Richard (1974). *Formal Philosophy: Selected Papers, RH Thomason, Ed*. Yale University Press, New Haven, Conn.

Pollard, Carl and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago, IL and Stanford, CA: The University of Chicago Press and CSLI Publications.

Flickinger, Dan (2000). "On building a more efficient grammar by exploiting types". In: *Natural Language Engineering* 6.01, pp. 15–28.

— (2011). "Accuracy v. Robustness in Grammar Engineering". In: *Language from a Cognitive Perspective: Grammar, Usage and Processing*. Ed. by Emily M Bender and Jennifer E. Arnold. Stanford, CA: CSLI Publications, pp. 31–50.

Hajdik, Valerie et al. (2019). "Neural text generation from rich semantic representations". In: *arXiv preprint arXiv:1904.11564*.

Zamaraeva, Olga, Kristen Howell, and Adam Rhine (2018). "Improving Feature Extraction for Pathology Reports with Precise Negation Scope Detection". In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 3564–3575.

Buys, Jan and Phil Blunsom (2017). "Robust Incremental Neural Semantic Graph Parsing". In: *ACL (1)*.

Packard, Woodley (2014). "UW-MRS: Leveraging a Deep Grammar for Robotic Spatial Commands". In: *SemEval 2014*, p. 812.

Oepen, Stephan and Dan Flickinger (2019). "The ERG at MRP 2019: Radically Compositional Semantic Dependencies". In: *CoNLL 2019*, p. 40.

Siegel, Melanie, Emily M Bender, and Francis Bond (Nov. 2016). *Jacy: An Implemented Grammar of Japanese*. CSLI Studies in Computational Linguistics. Stanford CA: CSLI Publications. ISBN: 9781684000180.

Fan, Zhenzhen (2018). "Building an HPSG Chinese grammar (Zhong)". PhD thesis. Nanyang Technological University.

Crysmann, Berthold (2003). "On the Efficient Implementation of German Verb Placement in HPSG". In: *Proceedings of RANLP 2003*. Borovets, Bulgaria, pp. 112–116.

Marimon, Montserrat (2010). "The Spanish Resource Grammar.". In: *LREC*.

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". In: *Proceedings of the Workshop on Grammar Engineering and Evaluation at the 19th International Conference on Computational Linguistics*. Ed. by John Carroll, Nelleke Oostdijk, and Richard Sutcliffe. Taipei, Taiwan, pp. 8–14.

Bender, Emily M, Scott Drellishak, et al. (2010). "Grammar Customization". In: *Research on Language & Computation* 8.1, pp. 23–72. ISSN: 1570-7075.

References II

Bender, Emily M (2010). "Reweaving a grammar for Wambaya". In: *Linguistic Issues in Language Technology* 3.1.

Crowgey, Joshua (2019). "Braiding Language (by Computer): Lushootseed Grammar Engineering". PhD thesis. University of Washington.

Inman, David Anthony (2019). "Multi-predicate Constructions in Nuuchahnulth". PhD thesis. University of Washington.

Zamaraeva, Olga (2021). "A cross-linguistic analysis of constituent questions for the Grammar Matrix". PhD thesis. University of Washington.

Ginzburg, Jonathan and Ivan Sag (2000). *Interrogative investigations*. Stanford: CSLI publications.

Zamaraeva, Olga and Guy Emerson (2020). "Multiple Question Fronting without Relational Constraints: An analysis of Russian as a basis for cross-linguistic modeling". In: *Proceedings of the 27th International Conference on Head-Driven Phrase Structure Grammar*. Virtual conference, to appear.

Crysmann, Berthold (2015). "Resumption and Extraction in an Implemented HPSG grammar of Hausa". In: *Proceedings of the Grammar Engineering Across Frameworks (GEAF) 2015 Workshop*, pp. 65–72.

Zamaraeva, Olga (to appear). "Morphological marking of questions: A case for nonlocal amalgamation". In: *Proceedings of the 28th International Conference on Head-Driven Phrase Structure Grammar*, to appear.

Zamaraeva, Olga, Kristen Howell, and Emily M Bender (2019). "Modeling clausal complementation for a grammar engineering resource". In: Proceedings of the 2nd meeting of the Society for Computation in Linguistics, pp. 39–49.

Howell, Kristen and Olga Zamaraeva (2018). "Clausal modifiers in the Grammar Matrix". In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 2939–2952.

Saleem, Safiyyah (2010). "Argument optionality: A new library for the Grammar Matrix customization system". MA thesis. University of Washington.

Song, Sanghoun (2014). "A grammar library for information structure". PhD thesis. University of Washington.

Nielsen, Elizabeth K (2018). "Modeling adnominal possession in the lingo grammar matrix". MA thesis. University of Washington.

Drellishak, Scott and Emily M Bender (2005). "Coordination Modules for a Crosslinguistic Grammar Resource". In: *Departamento de Informática Faculdadede Ciências da Universidadede Lisboa Campo Grande, 1749–016 Lisboa Portugal*, p. 29.

Crowgey, Joshua (2013). "The syntactic exponence of sentential negation: A model for the LinGO Grammar Matrix". PhD thesis. University of Washington.

Bender, Emily M and Dan Flickinger (2005). "Rapid Prototyping of Scalable Grammars: Towards Modularity in Extensions to a Language-Independent Core". In: *Proceedings of the 2nd International Joint Conference on Natural Language Processing IJCNLP-05 (Posters/Demos)*. Jeju Island, Korea.

References III

König, Ekkehard and Peter Siemund (2007). "Speech act distinctions in grammar". In: *Language typology and syntactic description* 1, pp. 276–324.

Hagège, Claude (2008). "Towards a typology of interrogative verbs". In: *Linguistic typology* 12, pp. 1–44.

Miyagawa, Shigeru (1987). "LF affix raising in Japanese". In: *Linguistic Inquiry* 18.2, pp. 362–367.

Schneider, D. and K. McCoy (1998). "Recognizing Syntactic Errors in the Writing of Second Language Learners". In: *ACL*, pp. 1198–1204.

Bender, Emily M, Dan Flickinger, Stephan Oepen, et al. (2004). "Arboretum: Using a precision grammar for grammar checking in CALL". In: *InSTIL/ICALL*.

Morgado da Costa, Luis, Francis Bond, and Xiaoling He (2016). "Syntactic well-formedness diagnosis and error-based coaching in computer assisted language learning using machine translation". In: *NLPTEA*, pp. 107–116.

Morgado da Costa, Luis, R. Winder, et al. (2020). "Automated Writing Support Using Deep Linguistic Parsers". In: *LREC*, pp. 369–377.