Logical forms in broad-coverage grammar development

Ann Copestake

Computer Laboratory, University of Cambridge

August 17, 2017

Acknowledgments: Aurelie Herbelot, Alex Lascarides, Dan Flickinger, Emily Bender, Francis Bond, Stephan Oepen, Guy Emerson, Alexander Kuhnle, DELPH-IN

Outline.

Outline.

1. Preliminaries
2. Computational compositional semantics in DELPH-IN
3. Lexical meaning and inference in *MRS
4. Meaning representation and meaning

A personal perspective

- language: ‘alignment’ between sapient entities
- what is communicated depends on world context, status of the entities, amount of inference they do etc etc
- but we can use language in many ‘artificial’ ways, such as mathematics (perhaps surprising . . .)
- there are conventions which are (mostly) shared by speakers of a language: grammars and lexicons
- there are regularities across languages
- language has to be learnable and usable

And so on . . . Different fields of linguistics focus on different aspects.

The formal semantics tradition

- Takes the relationship between language and the world (really a microworld) as primary.
- Only fully covers very regular/limited contexts (potentially maths, with a lot more work: see Ganesalingam 2013).
- Considers some aspects of grammar in depth: formal syntax-semantics interface for grammar fragments.
- Usually trivializes the open-class lexicon.
- Some interest in regularities across languages: e.g., generalized quantifiers.
- Reasonably clear (though limited) methodology.

Computational formal semantics

- Long history: first(?) explicitly Montogovian work by Bronnenberg et al (1980), but note Woods et al (1972).
- SRI: Core Language Engine (Alshawi et al, 1992), Hobbs (1985) etc.
- Main application was Natural Language Interfaces to Databases: good match for formal semantics.
- More recently: several high-to-medium-throughput broad-coverage grammars with semantic output: e.g., C&C/Boxer, XLE, DELPH-IN.
- Classical formal semantics relevant but not simply transferable.

Compositional semantics for broad-coverage grammars

- Meaning representation for **every** sentence (and phrase).
- But not all sentences are logically interpretable. So:
  - Meaning representation supports logical interpretation in suitable contexts (incl. model to constrain interpretation).
  - Logical interpretation guides representation decisions.

Also:

- Capture all and only semantically-relevant information from syntax and morphology.
- Underspecify when information is absent (e.g., quantifier scope).
- No hidden syntactic assumptions in the representation.

Some other desiderata

- Cross-linguistically adequate
- Usable in realization and parsing
- Statistical ranking of analyses
- Support applications (robust inference)
- Usable for shallow parsing
- Incremental processing (e.g., Haugereid 2009)
- Lexical semantics . . .

Outline.

Broad-coverage computational syntax and compositional semantics

DELPH-IN collaboration (`www.delph-in.net`):

- English Resource Grammar (Flickinger 2000); Minimal Recursion Semantics (MRS: Copestake et al, 2005); English Resource Semantics (ERS: Bender et al, 2015).
- tools for processing (Oepen, Packard, Callmeier, Carroll, Copestake . . .)
- Statistical techniques for parse and realization ranking.
- Other resource grammars: Jacy (Japanese), GG (German), SRG (Spanish),
  also varying size grammars for Norwegian, Portuguese, Korean, Chinese . . .
- Grammar Matrix: Bender et al (2002).

ERG: some practicalities

- ERG: hand-written, domain-independent grammar
- Maxent parse selection models based on manual choice of analyses (Redwoods Treebanks: Oepen et al 2002, etc)
- ERG has about $80 \pm 10\%$ coverage on edited text
- Robustness: parsing: Packard and Flickinger, to appear; realization: Horvat, to appear.
- Downloadable corpora:
  - Manually selected/checked (Redwoods Treebank): DeepBank (PTB/WSJ data), WeScience etc
  - Automatically processed: Wikiwoods (Flickinger et al, 2010)
- All DELPH-IN resources are Open Source.
- Various output formats for syntax and semantics.
- Used on many projects since 1990s, including large-scale end-user applications.

MRS

Some big angry dog barks loudly

Fully specified logical form:
some(x4, big(e8,x4) $\wedge$ angry(e9,x4) $\wedge$ dog(x4), bark(e2,x4) $\wedge$ loud(e10,e2))

ERS, generalized quantifiers, lots of ‘event’ variables

MRS:
l1:_some_q (x4, h5, h6),
l2:_big_a(e8,x4),
l2:_angry_a(e9,x4),
l2:_dog_n(x4),
l4:_bark_v(e2,x4),
l4:_loud_a(e10,e2),
h5 $=_q$ l2

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,h2), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,h4), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,h2), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,h4), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, h2=l5, h4=l4

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,**h2**), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), **l5**:every(y,h3,h4), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, **h2=l5**, h4=l4

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,h2), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,**h4**), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, h2=l5, **h4=l4**

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,h2), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,h4), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, h2=l5, h4=l4

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))
h1=l2, h3=l6

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,**h2**), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,h4), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, h2=l5, h4=l4

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))
h1=l2, h3=l6, **h2=l4**, h4=l1

Scope underspecification

Some big dog chased every cat

l1:some(x,h1,h2), h1 qeq l2, l2:big(x), l2:dog(x),
l4:chase(e,x,y), l5:every(y,h3,**h4**), h3 qeq l6, l6:cat(y)

Elementary predications (EPs) and scope constraints (qeqs)

some(x, big(x) $\wedge$ dog(x), every(y, cat(y), chase(e,x)))
h1=l2, h3=l6, h2=l5, h4=l4

every(y, cat(y), some(x, big(x) $\wedge$ dog(x), chase(e,x)))
h1=l2, h3=l6, h2=l4, **h4=l1**

RMRS

Some big angry dog barks loudly

MRS:
l1:_some_q (x4, h5, h6), l2:_big_a(e8,x4),
l2:_angry_a(e9,x4), l2:_dog_n(x4),
l4:_bark_v(e2,x4), l4:_loud_a(e10,e2),
h5 $=_q$ l2

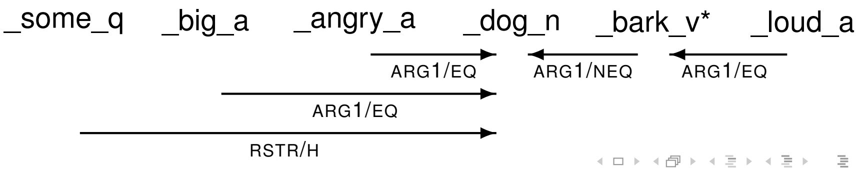
RMRS:
l1:a1:_some_q, BV(a1,x4), RSTR(a1,h5), BODY(a1,h6),
l2:a2:_big_a(e8), ARG1(a2,x4),
l2:a3:_angry_a(e9), ARG1(a3,x4),
l2:a4:_dog_n(x4),
l4:a5:_bark_v(e2), ARG1(a5,x4),
l4:a6:_loud_a(e10), ARG1(a6,e2), h5 $=_q$ l2

DMRS

Some big angry dog barks loudly

RMRS:
l1:a1:_some_q, BV(a1,x4), RSTR(a1,h5), BODY(a1,h6),
l2:a2:_big_a(e8), ARG1(a2,x4),
l2:a3:_angry_a(e9), ARG1(a3,x4),
l2:a4:_dog_n(x4),
l4:a5:_bark_v(e2), ARG1(a5,x4),
l4:a6:_loud_a(e10), ARG1(a6,e2), h5 $=_q$ l2

DMRS:



A real example sentence

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

A real example sentence

Very few of the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

modified quantifier

A real example sentence

Very few **of** the Chinese construction companies consulted were even remotely interested in entering into such an arrangement with a local partner.

partitive

A real example sentence

Very few of the Chinese **construction companies** consulted were even remotely interested in entering into such an arrangement with a local partner.

compound nominal

A real example sentence

Very few of the Chinese construction companies **consulted** were even remotely interested in entering into such an arrangement with a local partner.

reduced relative

A real example sentence

Very few of the Chinese construction companies consulted were **even remotely** interested in entering into such an arrangement with a local partner.

modified modifier

A real example sentence

Very few of the Chinese construction companies consulted were even remotely interested in entering into **such an** arrangement with a local partner.

predeterminer

What can I do with an ERS?

Applications investigated include:

- Machine translation: e.g., Bond et al (2011)
- Information extraction and QA: e.g., MacKinlay et al (2009)
- Ontology extraction: e.g., Herbelot and Copestake (2006)
- Question generation: e.g., Yao et al (2012)
- Entailment recognition: e.g., Lien and Kouylekov (2014)
- Preprocessing for distributional semantics: e.g., Herbelot (2013)
- Detection scope of negation: e.g., Packard, Bender, Read, Oepen and Dridan (2014)
- Robot control interface: e.g., Packard (2014)
- Logic to English (for teaching logic): e.g. Flickinger (2017)

MRS vs (deep) syntax

MRS more abstract, less language-dependent than (detailed) syntax: e.g., Bender (2008) on Wambaya.

1. Construction semantics: e.g., relative clauses:
every cat who slept snored
l5:every(y,h3,h4), h3 qeq l6, l6:cat(y), l6:sleep(e,y), l7:snore(e1,y)
2. Construction semantics: additional predications:
tree house
l1:house(x), l3:**udef_q**(y,h2,h3), h2 qeq l2, l2:tree(y), l2:**cmpd**(e,x,y)
house in a tree
l1:house(x), l3:a(y,h2,h3), h2 qeq l2, l2:tree(y), l2:in(e,x,y)
3. Words with no direct semantic contribution:
relative clause *who*, infinitival *to*, expletive *it* etc
4. Multiword expressions: verb-particle, idioms etc.

MRS vs predicate calculus

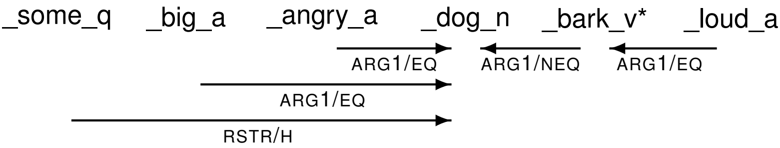
Copestake et al (2005) formally describe MRS as a meta-language for predicate calculus object language.
As used in ERS:

- NOT a fragment: produce some sort of MRS for everything including: generics, liar sentences, *circular square*, greetings ...
- contradictions, speakers with different word uses ...
- interpretation of 'logical' vocabulary isn't determined: *or* (exclusive or not?), *all* (domain of quantification, really universal?) and so on.

Much of this is not new to MRS, but rarely explicit ...

DMRS as a more meta meta-language?

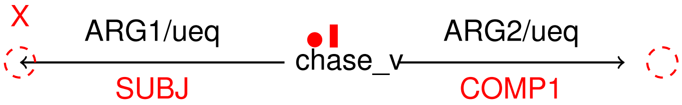
Some big angry dog barks loudly



Current ERS: some(x4, big(e8,x4) $\wedge$ angry(e9,x4) $\wedge$ dog(x4), bark(e2,x4) $\wedge$ loud(e10,e2))
Normal Davidsonian: some(x4, big(x4) $\wedge$ angry(x4) $\wedge$ dog(x4), bark(e2,x4) $\wedge$ loud(e2))
No event variables: some(x4, big(x4) $\wedge$ angry(x4) $\wedge$ dog(x4), loud(bark)(x4))

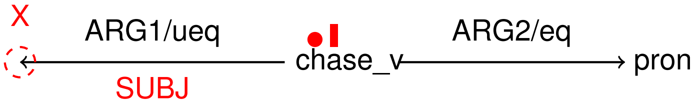
DMRS composition

chase:



red dashed circle — slot
red circle — ltop
red rectangle — index

chase it:



The ‘logical’ fragment of ERS

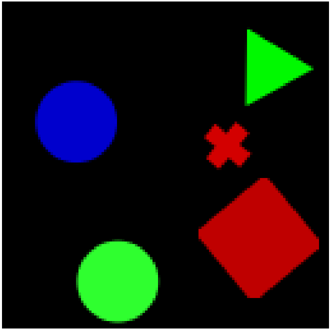
- Cannot produce model-theoretic interpretation for all ERS.
- But: reasonable semantics for a (substantial) fragment.
- Methodology:
  - Use intuitions about truth conditions to develop ERS for a small test set (fragment).
  - Assume similar structures outside fragment.
  - Note: there are some structures which don’t simply follow from syntax: e.g., generalized quantifiers, ‘small clauses’.
- Even without model-theoretic semantics, we want compositionality (motivation from learnability, substitution).
- Think of *MRS as annotation, not replacement.

Exploiting models

- Given a limited domain, expressed as a model, map (some) MRS relations to concepts in model.
- Classic Natural Language Interface (e.g., Woodley Packard demo).
- Other applications:
  - Flickinger (2017): teaching logic.
  - Shapeworld: generating material for training and testing neural network models (Kuhnle and Copestake 2017).

Shapeworld

Training and testing NNs with grounded language:



All circles are to the left of a red cross.

$$\forall s_1 \in W: \text{circle}(s_1.\text{shape}) \Rightarrow$$
$$\left(\exists s_2 \in W: \text{cross}(s_2.\text{shape}) \wedge \text{red}(s_2.\text{colour}) \wedge s_1.\text{x} < s_2.\text{x}\right)$$

Shapeworld (cont.)

- Automatically generate huge number of models in various classes: generate diagrams and DMRS using models.
- Generate English captions from DMRS using ERG (both true and false captions).
- Use pictures and captions to train NNs: evaluate performance on examples including unseen combinations (e.g., red triangle).
- Finding: performance of successful standard VQA approaches surprisingly bad (need new models).
- In progress: more languages.
- Compared with alternatives: no need for human annotation, less limited than simple template generation.

Outline.

Compositional semantics and lexical semantics

One version . . .

- Lexical semantics via Distributional Semantics (DS).
- Predicates are pointers into a context-dependent semantic space.
- Every dog barked — predicate points to the space corresponding to contextually-relevant dogs.
- Conventional meaning usually bounds the space (otherwise language wouldn't be learnable and novel utterances wouldn't be interpretable).
- Word 'senses' are somewhat conventionalized subspaces, systematically arising in particular contexts.

Using DELPH-IN resources in DS and deep learning

- Lots of treebanked data: automatically parsed data, checked by humans, annotated with grammar constraints. Much wider range of genres than standard NLP treebanks.
- Large quantities of automatically parsed data (e.g,. Wikiwoods — Wikipedia dump): currently using this in experiments on distributional semantics, SMT, sentence chunking.
- Implemented theoretically-grounded models which connect DELPH-IN style compositional semantics with distributions.

Ideal words

- Central discrepancy: DS gives observations about use, while formal semantics talks about denotation/reference.
- Copestake and Herbelot (2014), Herbelot and Copestake (in progress)
- **ideal corpus** — everything you could truthfully say about a situation (generative grammar, filtered by model).
- **ideal distribution**: a distribution derived from the ideal corpus.
- Direct formal link between distributional and model-theoretic semantics.

Emerson and Copestake (2016, 2017)

- Functional Distributional Semantics: functions mapping from semantic space (representing entities) to truth values.
- Distinguish between probabilistic truth values and observed text.
- DMRS gives joint distribution between entities.
- Implementation using deep learning techniques.
- Inference via conditional probabilities, also distributional similarity.
- e.g., *lion*, *stone lion*; *roses*, *plastic roses*, *stone roses*

Functional Distributional Semantics

$$\text{dog} \xleftarrow{\textsc{arg}1} \text{chase} \xrightarrow{\textsc{arg}2} \text{cat}$$

Functional Distributional Semantics

$$x \xleftarrow{\textsc{arg}1} y \xrightarrow{\textsc{arg}2} z$$

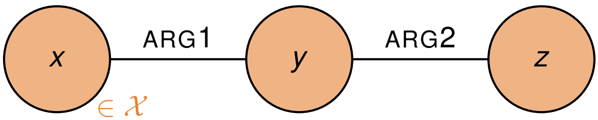
dog(x) chase(y) cat(z)

Functional Distributional Semantics

$$x \xleftarrow{\text{ARG}1} y \xrightarrow{\text{ARG}2} z$$

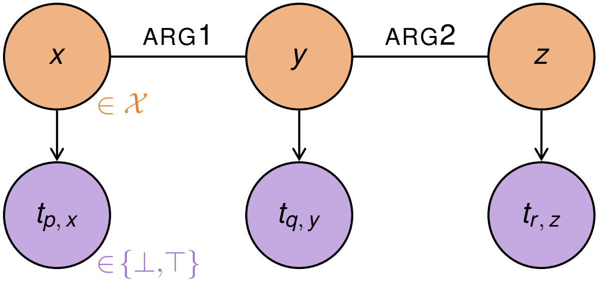
$p(x)$ $q(y)$ $r(z)$

Functional Distributional Semantics

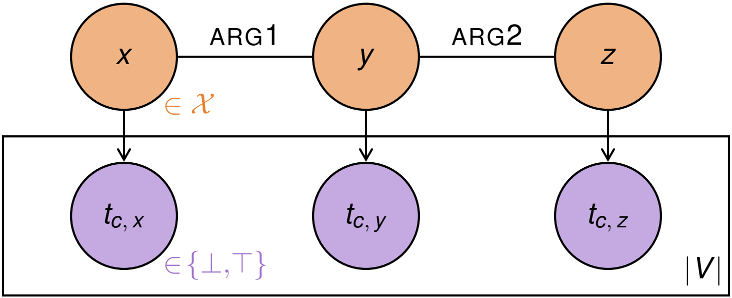


$p(x)$ $q(y)$ $r(z)$

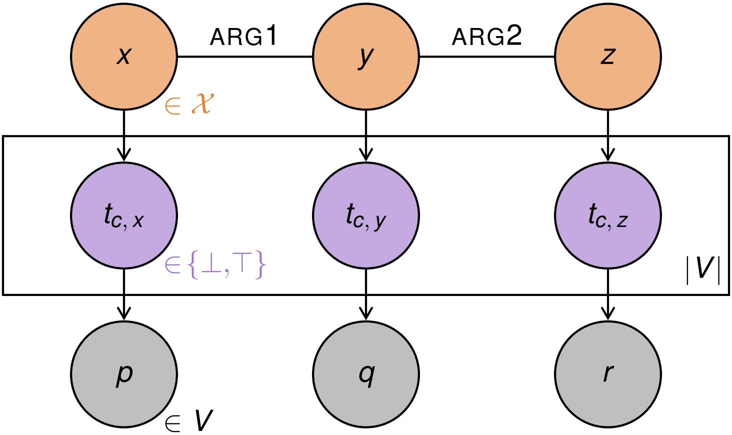
Functional Distributional Semantics



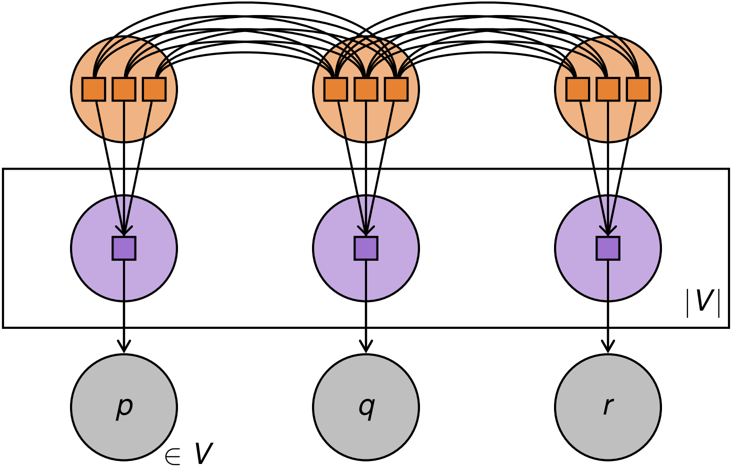
Functional Distributional Semantics



Functional Distributional Semantics



Functional Distributional Semantics



Outline.

Alternative philosophical accounts?

- Fregean tradition has problems if we assume we want a meaning representation for every utterance.
- Also has problems as a psycholinguistically plausible account (e.g., generics learned earlier than quantifiers).
- CL can use explicit models for interfaces to databases etc, but no obvious counterpart in broad-coverage systems.
- Rare to see full Montague Grammmar (intensional contexts etc), and only done for smallish fragments.
- Meaning as use (late Wittgenstein): explicit in some early Computational Linguistics (Masterman/CLRU).
- But late Wittgenstein much more about what we can't do than what we can . . .

One alternative: Brandom's version of Inferentialism

- Brandom (1994, 2000): non-Platonist, non-representationalist philosophical approach.
- cf 'meaning as use' but prioritizes 'giving and asking for reasons'.
- 'good inference' as prior to truth (cf early Frege).
- Logical inferences are a subset of material inferences.

  Pittsburgh is to the west of Philadelphia
  Philadelphia is to the east of Pittsburgh

- Top-down: propositions decomposable but not built from atomic meanings (cf Frege's Context Principle).
- Emphasis on pragmatics.

Inferentialism for computational linguists?

- Methodology of using human judgements (RTE etc) fits better with Brandom's 'commitment' to propositions than model-theoretic account: no theoretical problem with differing judgements.
- Not much in Brandom about differences in lexical semantics between speakers, but not obviously inconsistent.
- Lexical semantics: material inferences without further justification (e.g., 'east' and 'west').
- Explicitly logical vocabulary has important role: no need for us to abandon the stuff that works.
- MRS is a representation but use for decomposition/substitution consistent with inferentialism.

Shopping for philosophy?

- Not at all helpful for immediate grammar engineering!
- Philosophers and linguists taking us seriously (or not) ...
- Less contingent explanations for why we DON'T do things: e.g., intensional contexts.
- The point isn't whether or not Brandom (or others) are right, but what it leads us to investigate.
  e.g., use of language in more varied social contexts.
- Computational linguistics as empirical investigation of approaches to language semantics.

Concluding

- Lots of work on building grammars and compositional semantics in DELPH-IN . . .
- Reasonable confidence that this works at scale and across languages
- At some point, get away from hand-built grammars, but not urgent for English!
- Theoretical and practical work which relies on combination of compositional semantics and DS.