

Multi-lingual Meaning Representation: trying to tame a force of nature

Lori Levin

Language Technologies Institute

Carnegie Mellon University



The mapping between form and meaning

- What are the ways in which it is complex?
- How is it different in different languages?

Outline

Part I: Theoretical foundations and definitions

Part II: A journey through form and meaning

An NLP Project	Lessons about the Typology of Form and Meaning	Good Practice
Interlingua-based MT	The typology of argument realization Morphosyntactic strategies The importance of constructions	Design by linguists Live meetings with developers of each language
A classifier for definiteness	The complexity of semantic maps Intersecting meaning systems	Broad knowledge of linguists Read reference grammars
A classifier for causality	Morphosyntactic recruitment Intersecting meaning systems	Two-phase: expert, less-expert

Part III: Opinions and advice for multi-lingual meaning representation projects

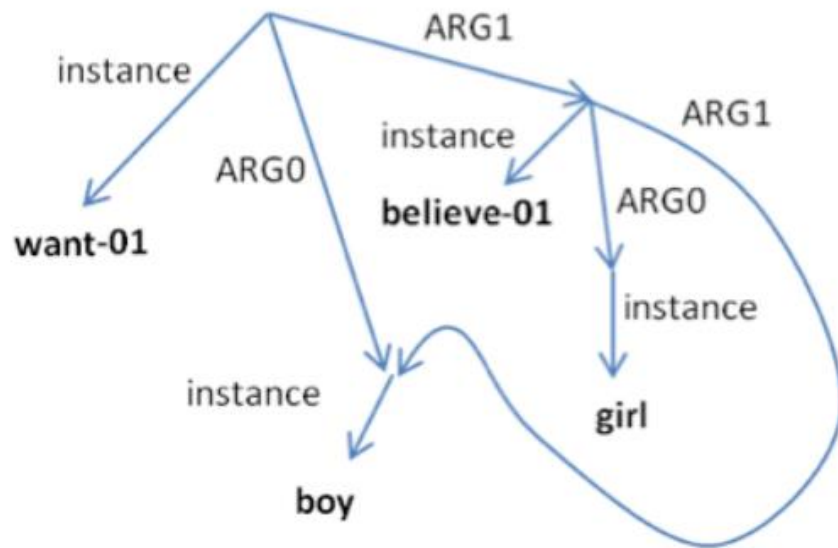
Foundation and Definitions

Foundations

- Cognitive Grammar
 - Grammatical categories like noun and subject are prototypes with radial categories.
- Construction Grammar
 - Constructions are conventionalized (grammaticalized) pairings of form and meaning.
 - Croft, *Radical Construction Grammar*,
 - universal and language-specific constructions
- Language Typology focusing on morphosyntax
 - The study of variation in human language
 - Croft, *Morphosyntax*
 - Typology and Radical Construction Grammar

Multi-lingual meaning representations for morphosyntax

- A discrete (not continuous), symbolic representation that is independent of different ways of *saying the same thing*.



1. The boy wants the girl to believe him.

(w / want-01

:ARG0 (b / boy)

:ARG1 (b2 / believe-01

:ARG0 (g / girl)

:ARG1 b))

What is a multi-lingual meaning representation?

- A discrete (not continuous), symbolic representation that is independent of different ways of *saying the same thing* in any human language.

- Otokonoko wa onnanoko ni kare o shinjite morai-tai.
- boy TOP girl DAT him ACC believing receive-want
- Literal: the boy wants to receive believing him *from* the girl.

1. The boy wants the girl to believe him.

(w / want-01

:ARG0 (b / boy)

:ARG1 (b2 / believe-01

:ARG0 (g / girl)

:ARG1 b))

Partial meaning representations

- PropBank (argument structure)
- FrameNet (frame semantics)
- WordNet (word senses)
- Many corpus annotation schemes for meanings like modality, temporal expressions, negation, genericity, definiteness, factivity

The uses of meaning representations

- Machine translation
- Communication with robots
- Automatic Question Answering

Granularity of meaning representations

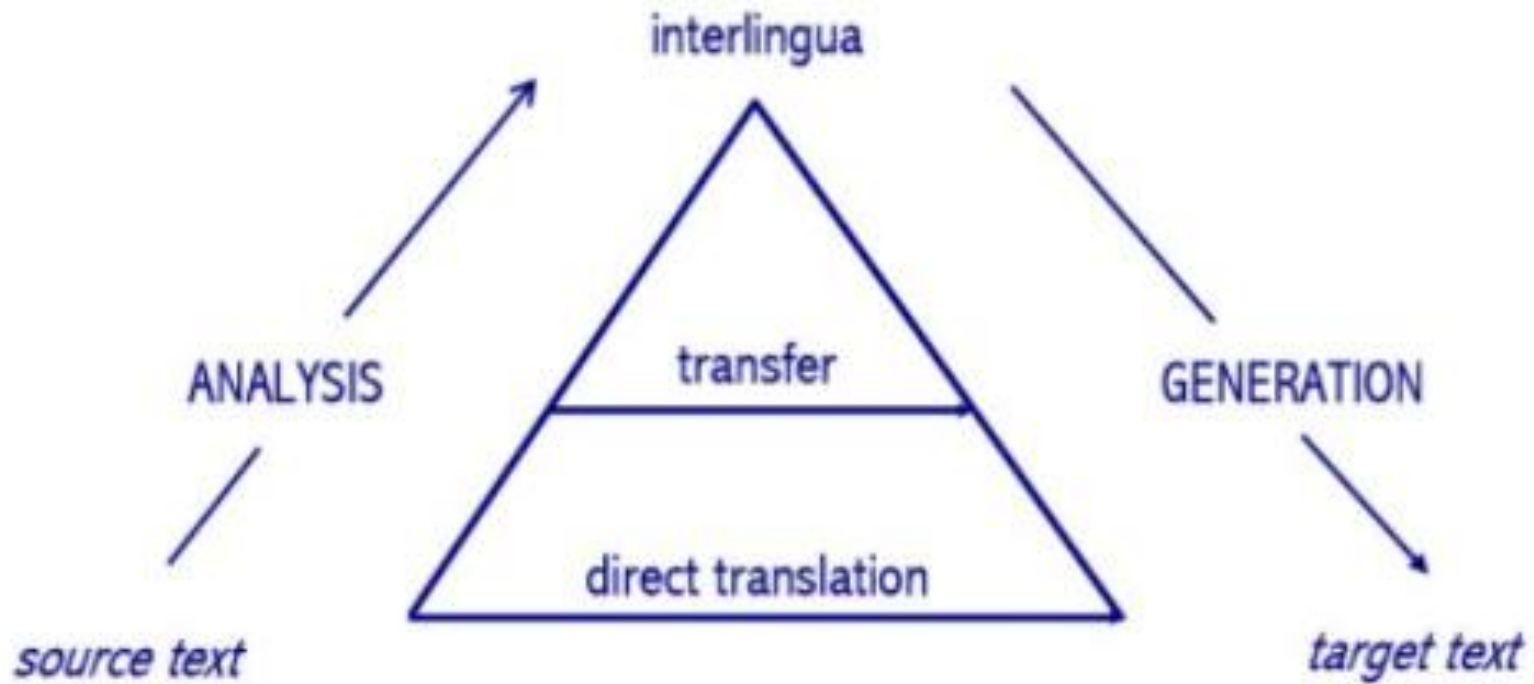
- The granularity is determined by what is typically lexicalized and grammaticalized.
 - No inferences or pragmatic meaning
- In a multi-lingual meaning representation the granularity is what *can* be lexicalized or grammaticalized in a human language.

Interlingua- based Machine Translation

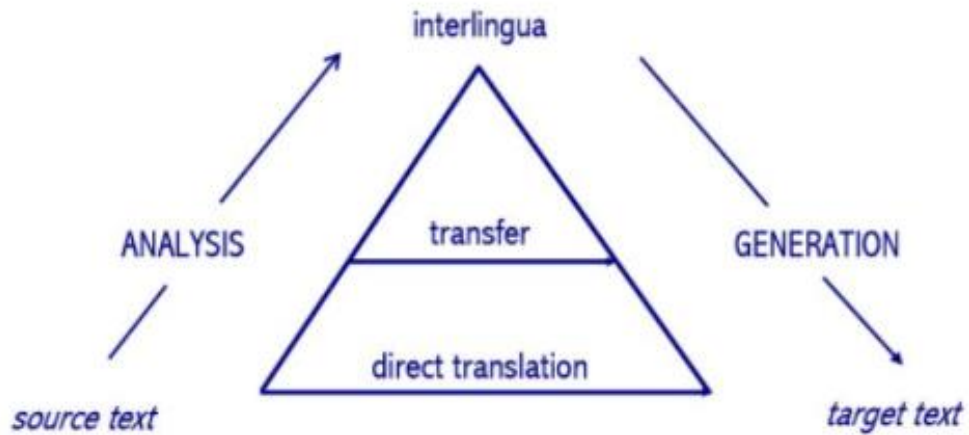
- The typology of argument realization
 - Morphosyntactic strategies
 - The importance of constructions
-
- Design by linguists
 - Have meetings with developers for different languages

The Vauquois Triangle

(Bernard Vauquois, 1968)



The Vauquois Triangle



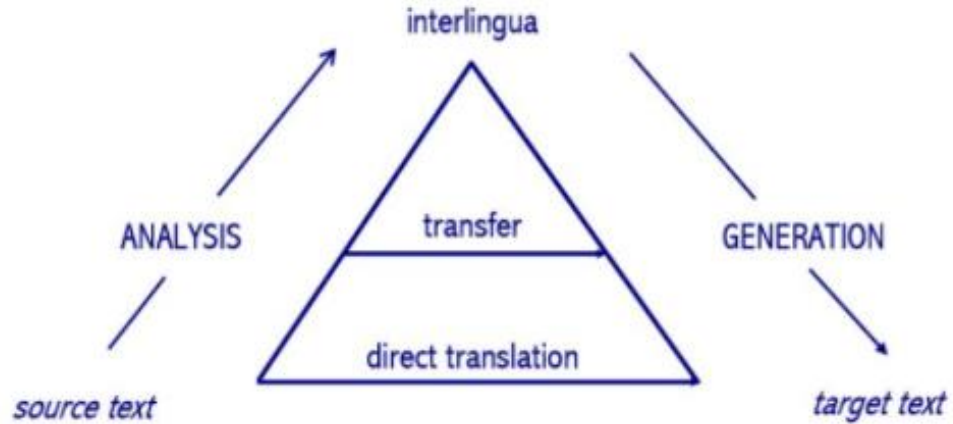
Direct translation (phrasebook):

Source language sentence: I read it



Target language sentence: Je l'ai lu

The Vauquois Triangle



Transfer-based Machine Translation

Source language sentence: I read it

Analysis

Source language structure:

Pronoun-1-agr1 V-pst pronoun-2-agr2

Transfer

Target language structure:

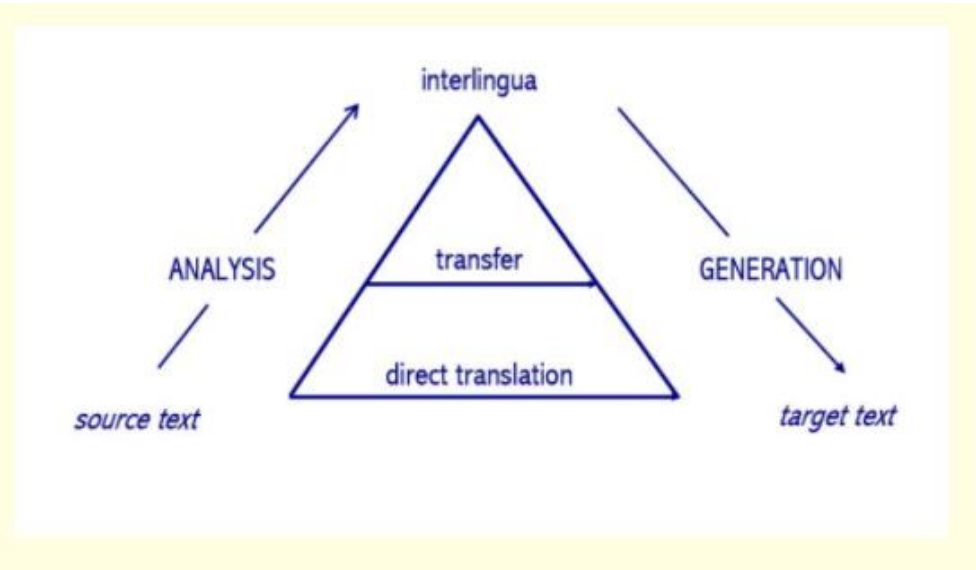
Pronoun-1-agr1 pronoun-2-agr2 AVOIR-agr1 V-pastpart

Generation

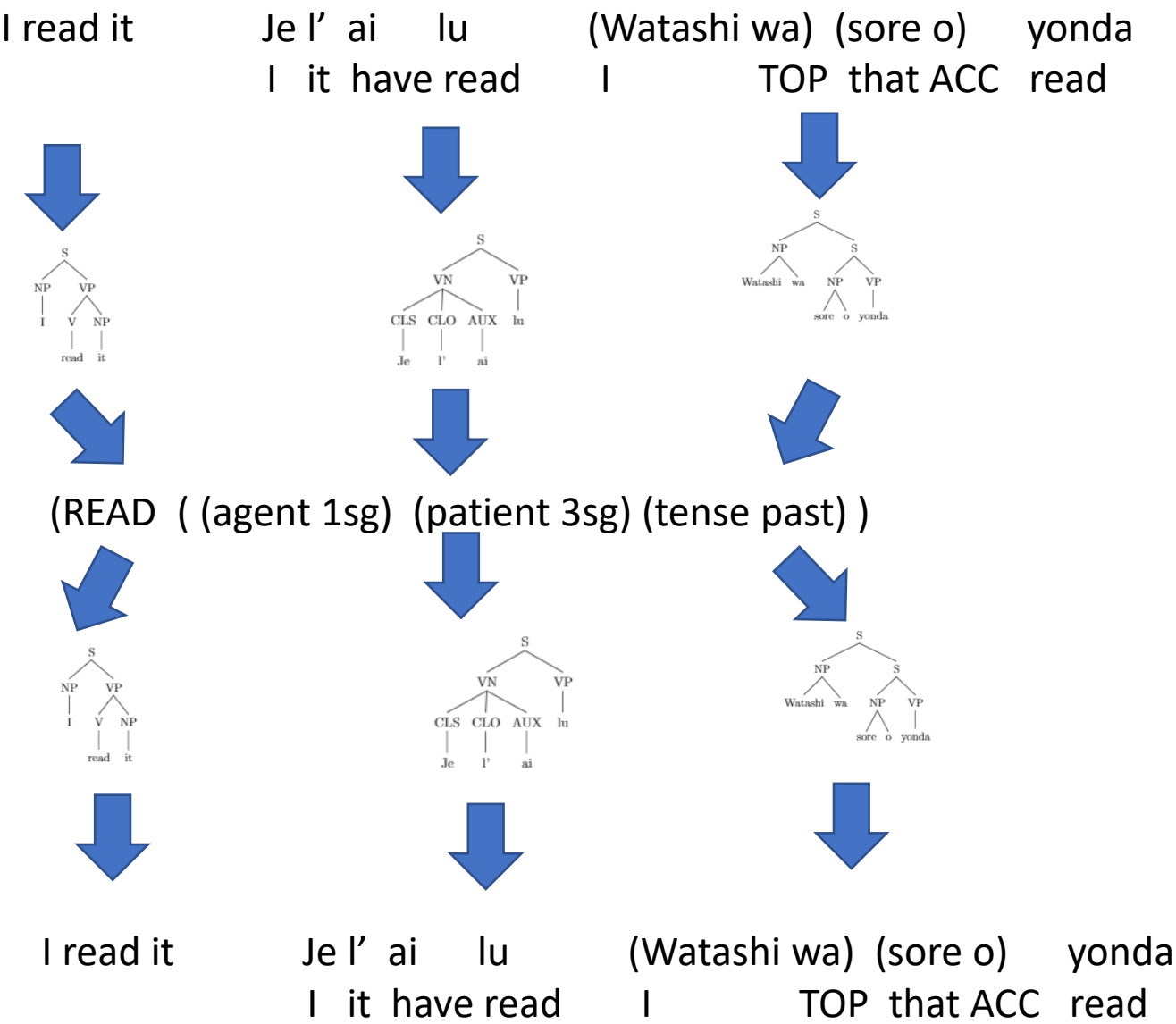
Target language sentence:

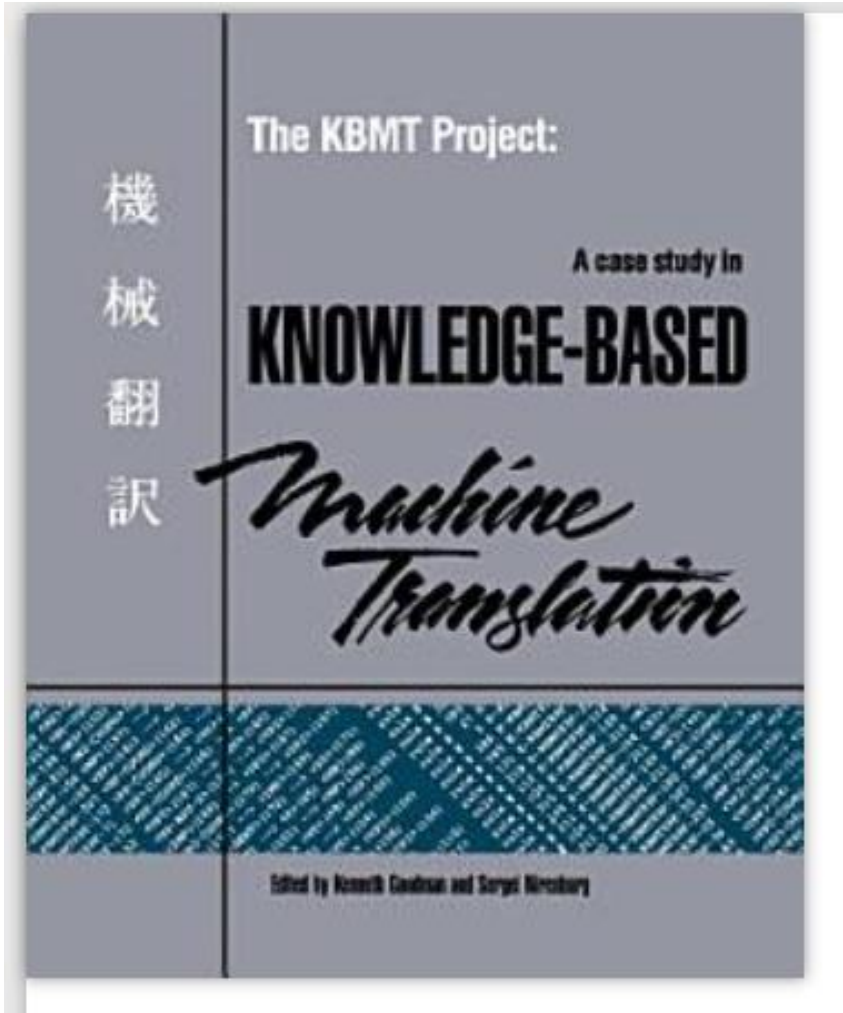
Je l'ai lu

The Vauquois Triangle



Interlingua-based machine translation





KBMT 89: Carnegie Mellon University and IBM Japan

- Sergei Nirenburg, Lori Levin, Teruko Mitamura, Eric Nyberg, Donna Gates, Koichi Takeda, and many others
- English and Japanese

Morphosyntactic Strategies

- Informally: morphosyntactic strategies are different grammar for doing the same thing.
- Example of morphosyntactic strategies:
 - To express subject and object: word order, case marking, agreement.
 - To express location: case marker, preposition, postposition, or genitive plus noun.
 - To express possession: mark on head, mark on possessor
 - To express negation: affix, verb, adverb
 - And thousands more

Typology of argument realization, verb classes, and transitivity alternations in linguistics

Causative-Inchoative Class

I (subj) opened the door (obj).

The door (subj) opened.

Conative Class

I kicked the ball (obj).

I kicked at the ball (obl).

Argument Realization Divergences

I (subj) like cake (obj).

Me (oblique) gusta el pastel (subject).

Argument Realization is the mapping from syntactic relations like subject and object to semantic roles like agent and patient.

Transitivity alternations are different mappings for a verb with the same semantic roles. There are classes of verbs that undergo the same set of transitivity alternations. (B. Levin, 1993)

Verbs with similar meanings may have different argument realization in different languages. (Dorr 1993)

Following Lexical
Functional Grammar,
KBMT 89 separated
parsing from semantic
mapping.

- Morphosyntactic strategies were handled during parsing
- Argument realization was handled during semantic mapping

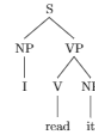
Interlingua-based machine translation

Parser: handles morphosyntactic strategies

Mapping Rules: handle typological differences in argument realization

Syntactic Generator: handles morphosyntactic strategies

I read it

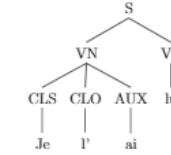


(READ (agent 1sg) (patient 3sg) (tense past))



I read it

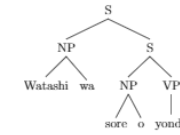
Je l' ai lu



Je l' ai lu
I it have read

(Watashi wa) (sore o) yonda

I TOP that ACC read



(Watashi wa) (sore o) yonda
I TOP that ACC read

Morphosyntactic Strategies for identifying subject and object in English and Japanese

```
(<S> <--> (<V>
  ((x0 = x1)))
```

```
(<S> <--> (<NP> <S>)
  (((x2 subj-case) = *defined*)
   ((x2 subj-case) = (x1 case))
   (x0 = x2)
   ((x0 subj) = x1)))
```

```
(<S> <--> (<NP> <S>)
  (((x2 obj-case) = *defined*)
   ((x2 obj-case) = (x1 case))
   (x0 = x2)
   ((x0 obj) = x1)))
```

```
(<S> <--> (<NP> <S>)
  (((x2 obj2-case) = *defined*)
   ((x2 obj2-case) = (x1 case))
   (x0 = x2)
   ((x0 obj2) = x1)))
```

Japanese: free word order and case marking

Neko **ga** inu **wo** mi-ta

Cat SUBJ dog OBJ see-PAST

```
(S <--> ( <NP> <VP> )
  ((x0 subj) = x1)
  (x0 = x2)))
```

```
(VP <--> ( <V> <NP> )
  ((x0 obj) = x2)
  (x0 = x1)))
```

English: word order

The cat saw the dog ≠ The dog saw the cat

A syntax-semantics mapping rule for interlingua-based MT

In KBMT 89, each language had an inheritance hierarchy for semantic mapping rules verb classes.

```
(emap *insert
  <=l=> insert ((CAT v) (SUBCAT trans))
  (role =sem (*physical-action))
  (:agent =syn (SUBJECT))
  (:theme =syn (DOBJECT))
  (:goal =syn (PPADJUNCT
    ((PREP into) (CAT n)))))
```

Takeda, Uramoto, Nasukawa, and Tsutsumi: Shalt2 – a symmetric machine translation system with conceptual transfer, COLING 1992

The importance of constructions

Constructional Divergences

(10) a. You'd better go.

b. Itta hoo ga ii.
go-PAST alternative SUBJ good
“The alternative that you went is good.”

c. Tebe stoit pojti.
you-DATIVE cost-IMPERSONAL go-INFINITIVE
“To you costs to go.”

Typology in the days of interlingua-based machine Translation

Constructional Divergences

- (10) a. You'd better go.
- auxiliary verb and adverb
- b. Itta hoo ga ii.
go-PAST alternative SUBJ good
“The alternative that you went is good.”
- c. Tebe stoit pojti.
you-DATIVE cost-IMPERSONAL go-INFINITIVE
“To you costs to go.”

Typology in the days of interlingua-based machine Translation

Constructional Divergences

(10) a. You'd better go.

b. Itta hoo ga ii.
go-PAST alternative SUBJ good
“The alternative that you went is good.”

Highly grammaticalized relative clause

c. Tebe stoit pojti.
you-DATIVE cost-IMPERSONAL go-INFINITIVE
“To you costs to go.”

Typology in the days of interlingua-based machine Translation

Constructional Divergences

(10) a. You'd better go.

b. Itta hoo ga ii.
go-PAST alternative SUBJ good
“The alternative that you went is good.”

c. Tebe stoit pojti.
you-DATIVE cost-IMPERSONAL go-INFINITIVE
“To you costs to go.”

Dative Case

(10) a. You'd better go.

b. Itta hoo ga ii.
go-PAST alternative SUBJ good
"The alternative that you went is good."

c. Tebe stoit pojti.
you-DATIVE cost-IMPERSONAL go-INFINITIVE
"To you costs to go."

clause-1

head: go-1
agent: *hearer*
destination: *unknown*

aspect:

phase: none

duration: *unspecified*

iteration: single

attitude-1

type: deontic

value: 0.8-1.0

scope: clause-1

attributed-to: *speaker*

time: time-of-speech

relation-1

type: temporal-before

from: time-of-speech

to: time-of (clause-1)

Figure 6. TMR for the sentences in (10).

Good Practice

- The design phase of the KBMT 89 project included with two linguists (Levin and Mitamura) who had written dissertations on argument realization in multiple languages.
- The KBMT 89 system was clean and modular based on Lexical Functional Grammar:
 - Grammatical encoding: morphosyntactic strategies
 - Lexical mapping: argument realization
- The C-STAR and NESPOLE! consortia:
 - Task-oriented spoken language: meeting scheduling and travel reservations
 - The interlingua represented domain-specific speech acts
 - Developers for seven languages met for 3-5 days twice a year
 - Ensured that every decision we made worked for all languages

A Classifier for Definiteness

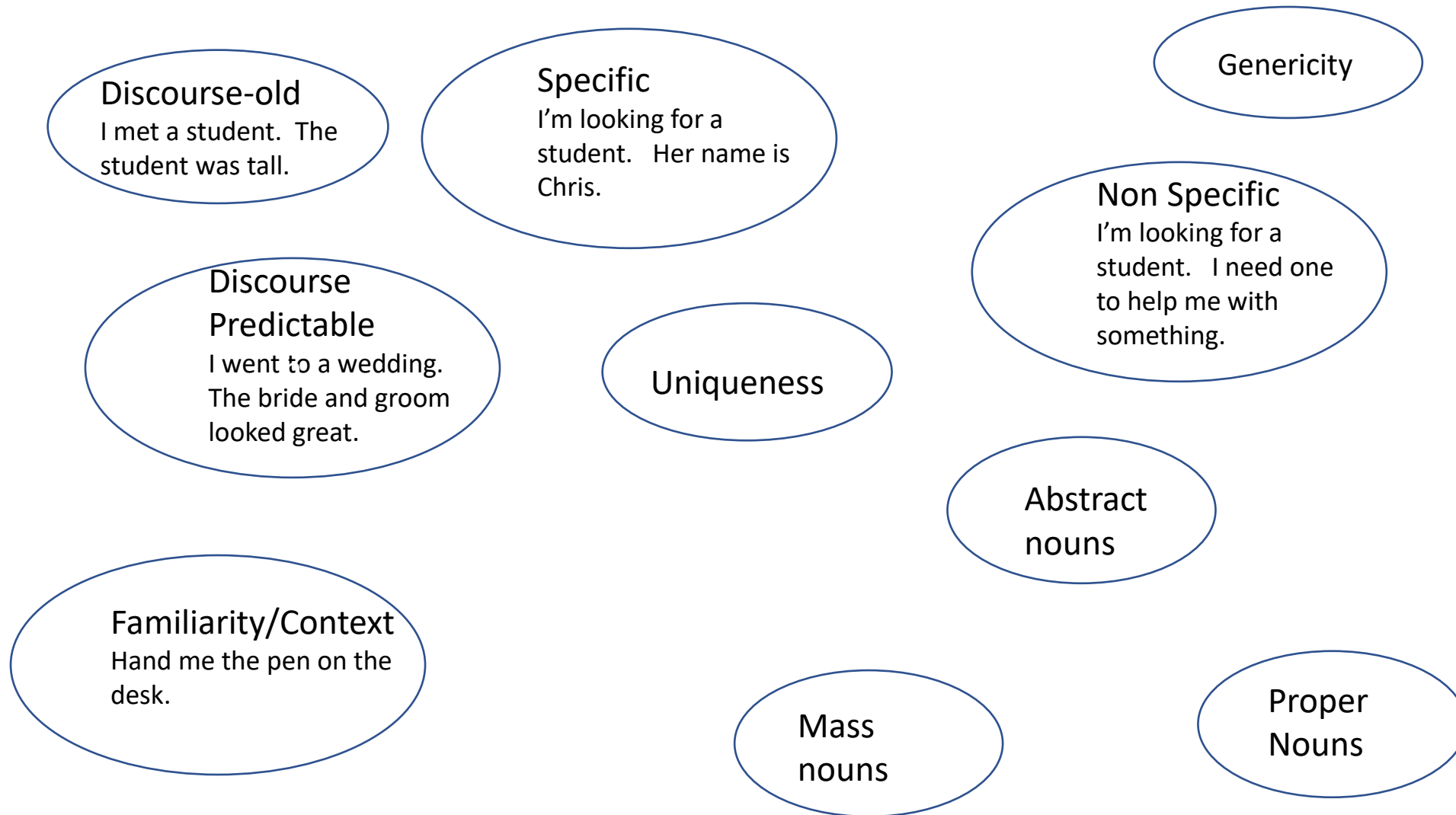
- The complexity of semantic maps
- Intersecting meaning systems

- Background in linguistics literature
- Read reference grammars

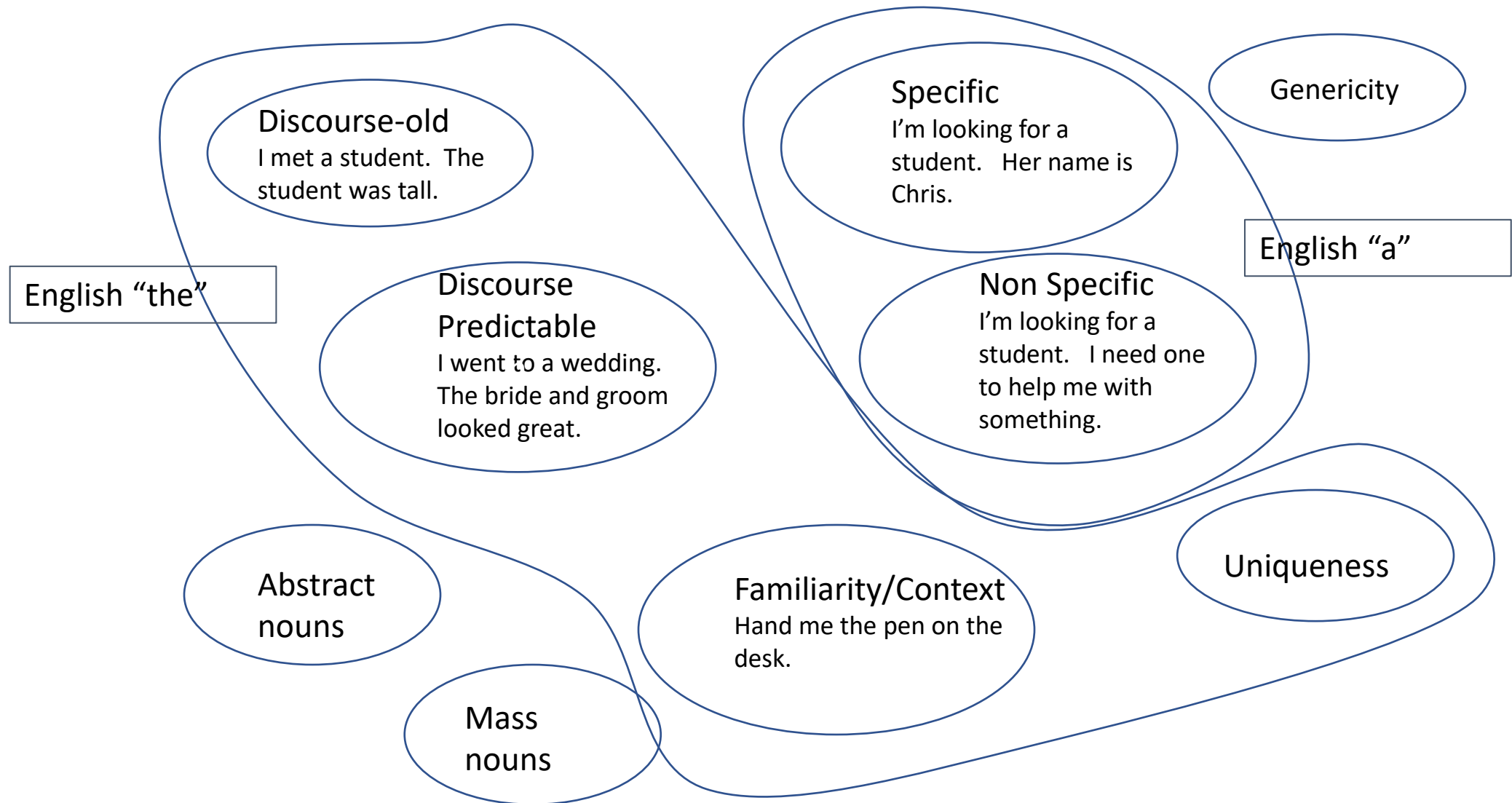
A Classifier for Definiteness

- ‘A Unified Annotation Scheme for the Semantic/Pragmatic Components of Definiteness’, Archana Bhatia, Mandy Simons, Lori Levin, Yulia Tsvetkov, Chris Dyer, Jordan Bender, LREC 2014
- ‘Automatic Classification of **Communicative Functions of Definiteness**’, Archana Bhatia, Chu-Cheng Lin, Nathan Schneider, Yulia Tsvetkov, Fatima Talib Al-Raisi, Laleh Roostapour, Jordan Bender, Abhimanu Kumar, Lori Levin, Mandy Simons, Chris Dyer, COLING 2014

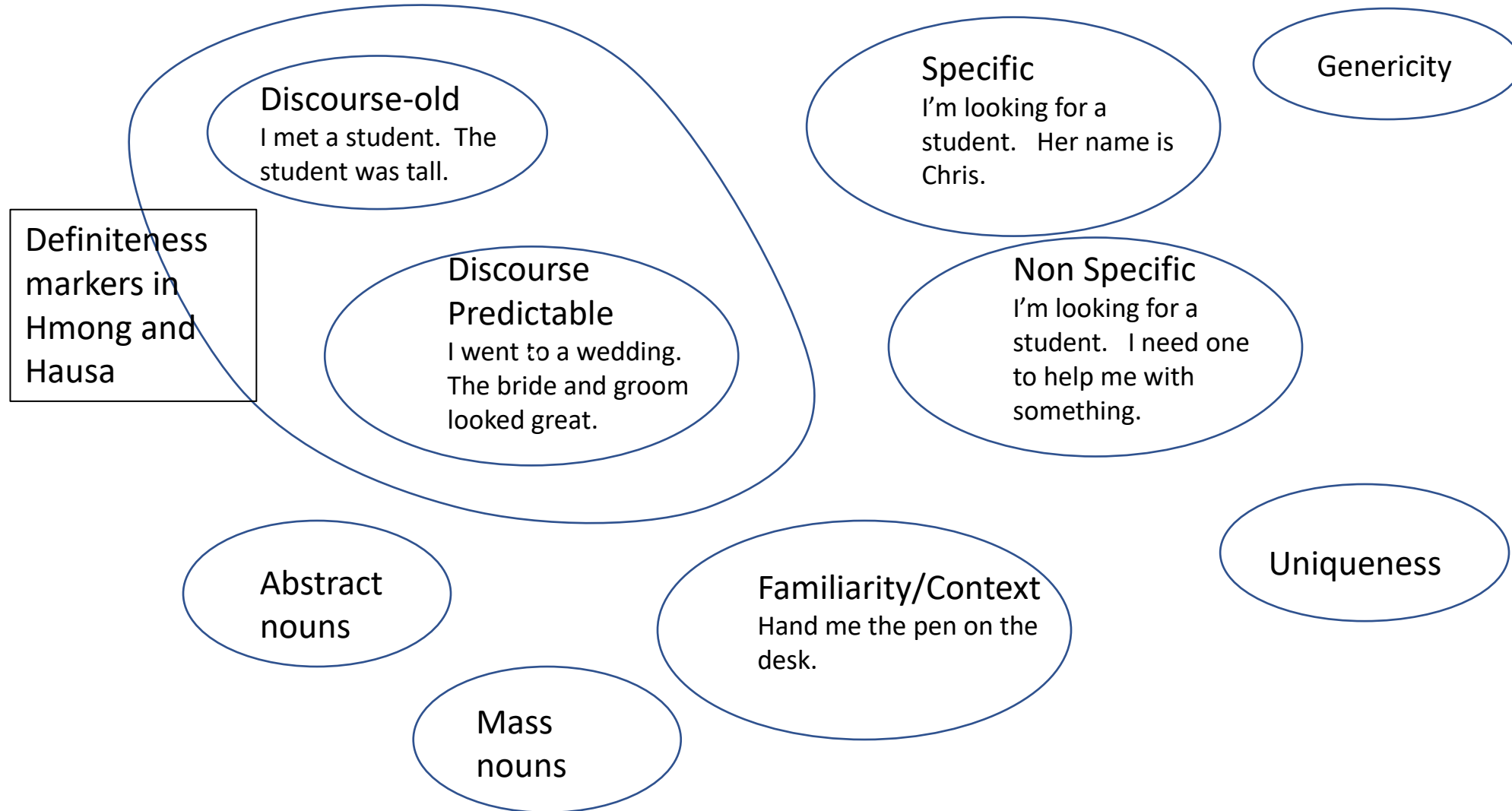
A semantic map for some aspects of definiteness



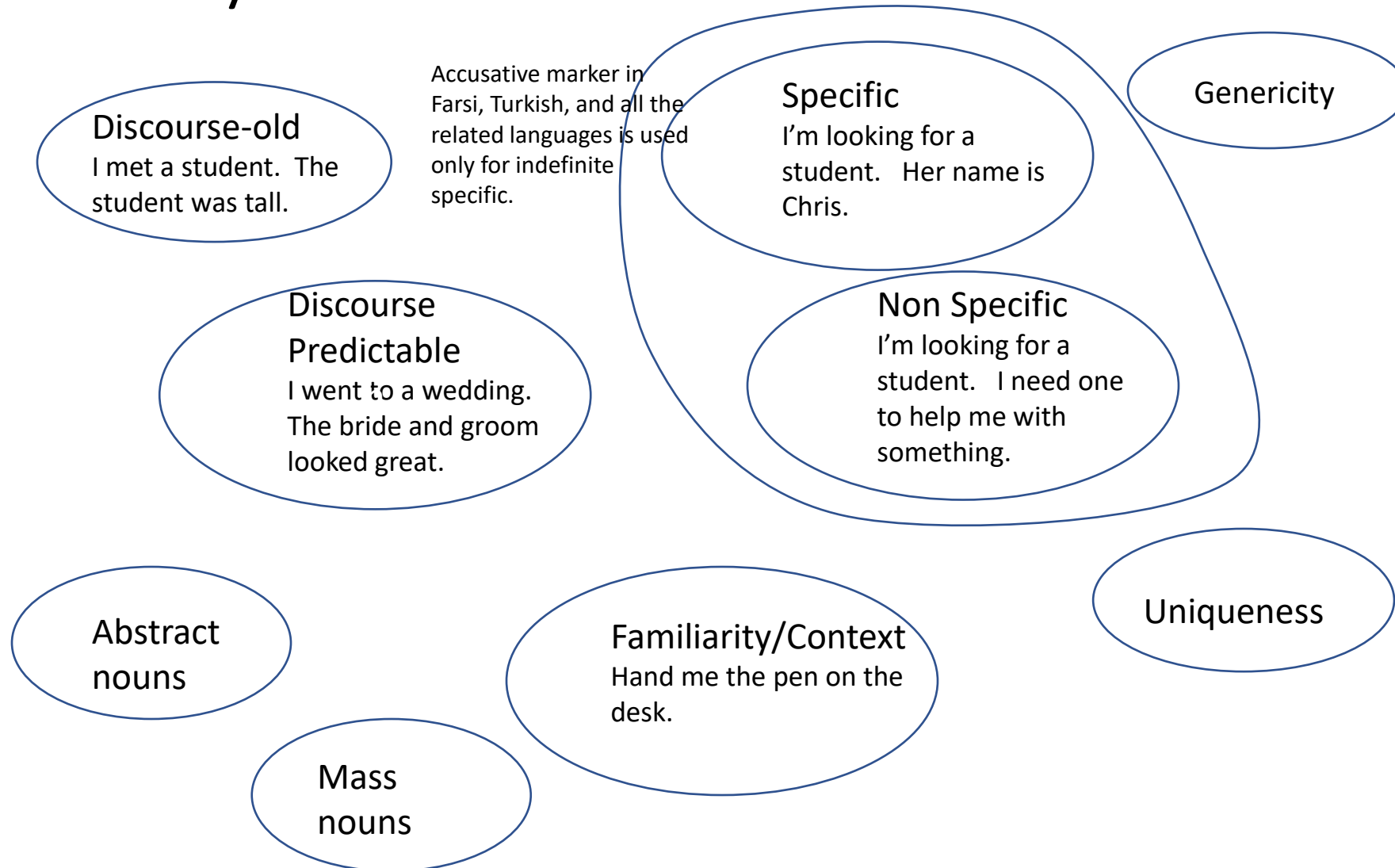
English “the” and “a”



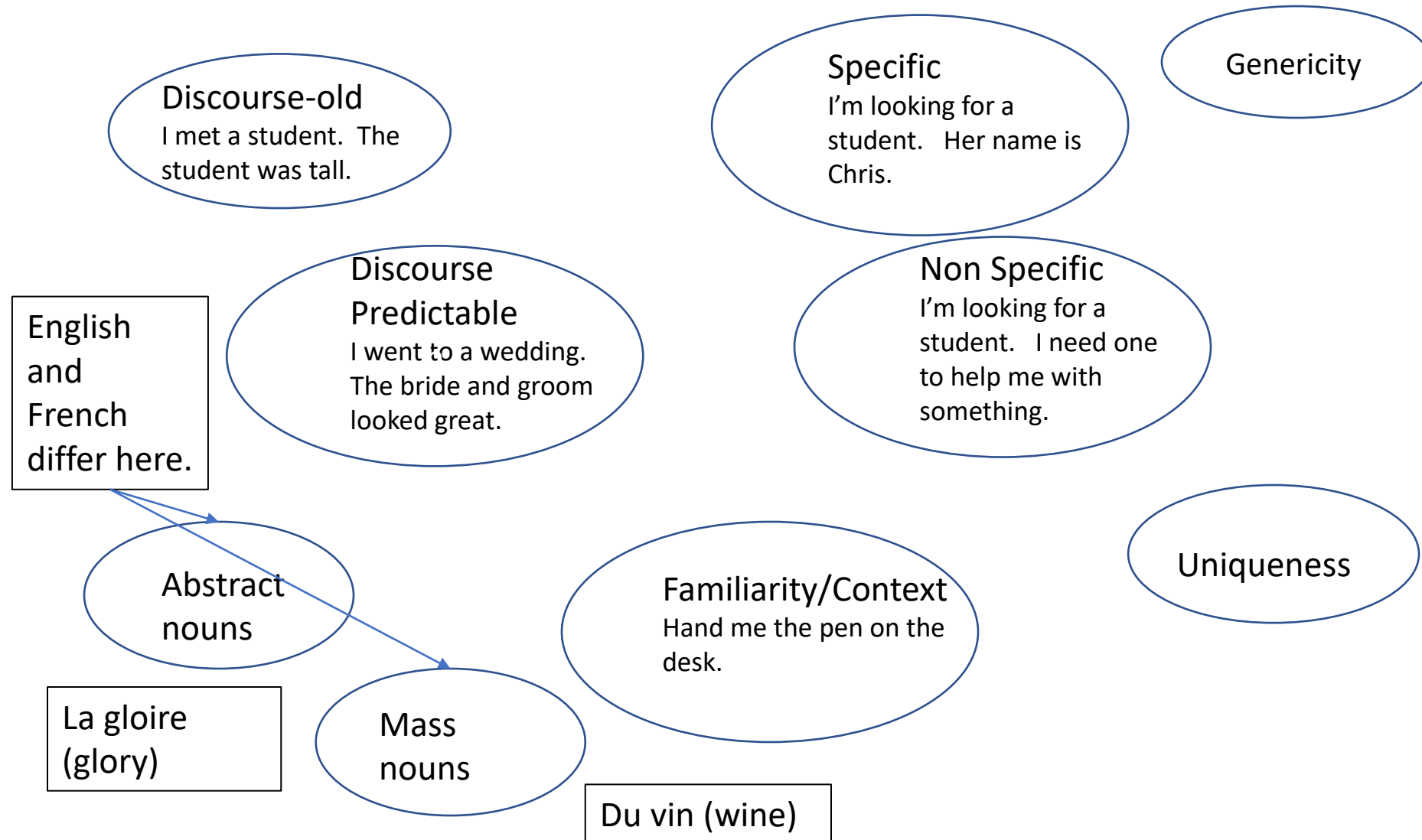
Hmong and Hausa “aforementioned” definiteness markers



Farsi and Turkish differential object marking for specificity



Even English and French do not have the same contour



Communicative Functions of Definiteness: semantic map/annotation scheme

• NONANAPHORA [-A, -B]	999	• ANAPHORA [+A]	1574
– UNIQUE [+U]	287	– BASIC_ANAPHORA [-B, +F]	795
* UNIQUE_HEARER_OLD [+F, -G, +S]	251	* SAME_HEAD	556
• UNIQUE_PHYSICAL_COPRESENCE [+R]	13	* DIFFERENT_HEAD	329
• UNIQUE_LARGER_SITUATION [+R]	237	– EXTENDED_ANAPHORA [+B]	779
• UNIQUE_PREDICATIVE_IDENTITY [+P]	1	* BRIDGING_NOMINAL [-G, +R, +S]	43
* UNIQUE_HEARER_NEW [-F]	36	* BRIDGING_EVENT [+R, +S]	10
– NONUNIQUE [-U]	581	* BRIDGING_RESTRICTIVE_MODIFIER [-G, +S]	614
* NONUNIQUE_HEARER_OLD [+F]	169	* BRIDGING_SUBTYPE_INSTANCE [-G]	0
• NONUNIQUE_PHYSICAL_COPRESENCE [-G, +R, +S]	39	* BRIDGING_OTHER_CONTEXT [+F]	112
• NONUNIQUE_LARGER_SITUATION [-G, +R, +S]	117	• MISCELLANEOUS [-R]	732
• NONUNIQUE_PREDICATIVE_IDENTITY [+P]	13	– PLEONASTIC [-B, -P]	53
* NONUNIQUE_HEARER_NEW_SPEC [-F, -G, +R, +S]	231	– QUANTIFIED	248
* NONUNIQUE_NONSPEC [-G, -S]	181	– PREDICATIVE_EQUATIVE_ROLE [-B, +P]	58
– GENERIC [+G, -R]	131	– PART_OF_NONCOMPOSITIONAL_MWE	100
* GENERIC_KIND_LEVEL	0	– MEASURE_NONREFERENTIAL	125
* GENERIC_INDIVIDUAL_LEVEL	131	– OTHER_NONREFERENTIAL	148

	+	-	0		+	-	0		+	-	0		+	-	0
<u>Anaphoric</u>	1574	999	732	<u>Generic</u>	131	1476	1698	<u>Predicative</u>	72	53	3180	<u>Specific</u>	1305	181	1819
<u>Bridging</u>	779	1905	621	<u>Familiar</u>	1327	267	1711	<u>Referential</u>	690	863	1752	<u>Unique</u>	287	581	2437

Example of definiteness annotation

Once upon a time there was a dear little girl who was loved by everyone who looked at her, but most of all by her grandmother, and there was nothing that she would not have given to the child.

Once she gave her a little riding hood of red velvet, which suited her so well that
SAME_HEAD DIFFERENT_HEAD OTHER_NONREFERENTIAL SAME_HEAD
NONUNIQUE_HEARER_NEW_SPEC

she would never wear anything else; so she was always called ‘Little Red Riding Hood.’
SAME_HEAD QUANTIFIED SAME_HEAD UNIQUE_HEARER_NEW

Figure 2: An annotated sentence from “Little Red Riding Hood.” The previous sentence is shown for context.

The annotated corpus:

17 documents: TED talks (75%), presidential speech (16%), fictional narrative (5%).

13,860 words

868 sentences

3,422 NPs

Details of the definiteness classifier

- Models
 - Log Linear (interpretable)
 - Random Forest (more accurate)
- Percepts (features)
 - The head of the NP (token, lemma, POS, token length, token position)
 - The dependents of the NP (token, lemma, POS)
 - The governor (immediate parent) of the NP
 - The dependency label attaching the NP to the governor
 - The closest dominating verb (left or right of NP)
 - The auxiliaries of the closest dominating verb
 - Whether the closest dominating verb is negated
 - Path length from the head of NP to the root of the tree
 - And more
- Goal: predict a communicative function given a set of percepts (features)

+Specific		Percepts	-Specific	
First dependent's POS	PRP\$		First dependent's lemma	a
Head's left neighbor's POS	PRP\$		Last dependent's lemma	a
Last dependent's lemma	you		Num. of dependents, dependent's lemma	1, a
Num. of dependents, dependent's lemma	1, you		Head's left neighbor's POS	JJR
Num. of dependents, dependent's POS	1, PRP\$		Last dependent's POS	JJR
Governor's right neighbor's POS	PRP\$		Num. of dependents, dependent's lemma	2, a
Last dependent's POS	NNP		First dependent's lemma	new
Last dependent's POS	PRP\$		Last dependent's lemma	new
First dependent's lemma	the		Num. of dependents, dependent's POS	2, JJR
Governor's lemma	from		Governor's left neighbor's POS	VB

Figure 3: Percepts receiving highest positive weights in association with values of the Specific attribute.

Example	Relevant percepts from fig. 3	CFD annotation
This is just for <i>the United States of America</i> .	Last dependent's POS: NNP First dependent's lemma: the	Unique_Larger_Situation
We were driving from <i>our home in Nashville</i> to a little farm we have 50 miles east of Nashville — driving ourselves.	First dependent's POS: PRP\$ Head's left neighbor's POS: PRP\$ Governor's right neighbor's POS: PRP\$ Governor's lemma: from	Bridging_Restrictive_Modifier

Figure 4: Sentences from our corpus illustrating percepts fired for gold NPs and their CFD annotations.

Overlapping semantic spaces

- The limits of our work on definiteness
 - In languages that do not have non-deictic determiners, some communicative functions of definiteness are accomplished with different morphosyntactic mechanisms such as word order, special constructions, and differential object marking.
 - But it is not clear that these morphosyntactic mechanisms are actually definiteness. They may be grammaticalizations of other semantic spaces that happen to intersect with definiteness:
 - Old and new information, affectedness, completedness

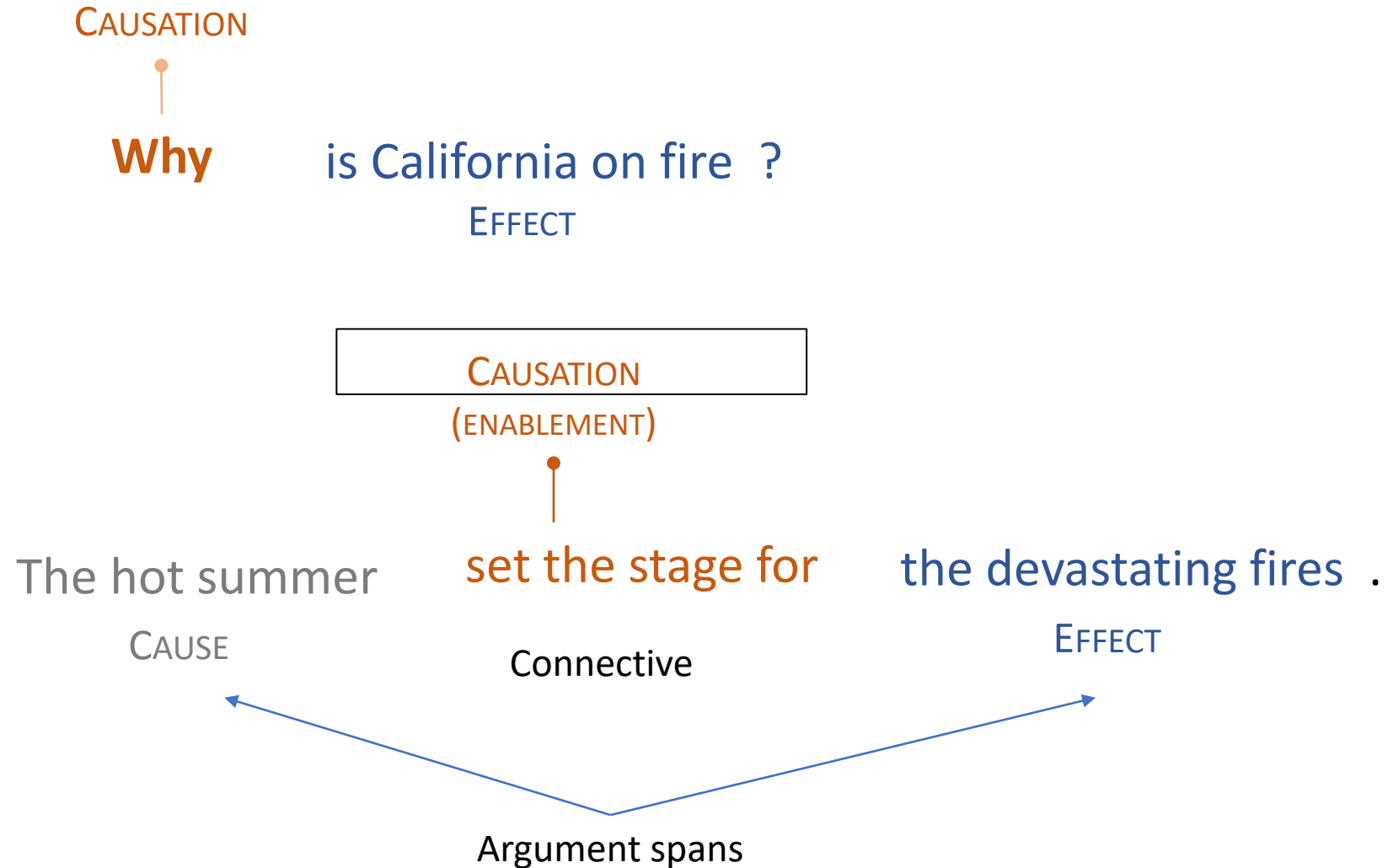
A Classifier for Causality

- Morphosyntactic Recruitment
- Two-phase annotation: expert and less expert

Collaborators and acknowledgements

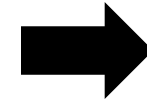
- Collaborators: Jesse Dunietz, Jaime Carbonell,
 - Dunietz, Jesse, Lori Levin, and Jaime Carbonell. "Automatically Tagging Constructions of Causation and Their Slot-Fillers." *Transactions of the Association for Computational Linguistics* 5 (2017). 117–133
 - Dunietz, Jesse, Lori Levin, and Jaime Carbonell. "The BECauSE Corpus 2.0: Annotating Causality and Overlapping Relations." *Proceedings of LAW XI – The 11th Linguistic Annotation Workshop* (2017). 95–104.
 - Dunietz, Jesse, Lori Levin, and Jaime Carbonell. "Annotating Causal Language Using Corpus Lexicography of Constructions." *Proceedings of LAW IX – The 9th Linguistic Annotation Workshop*(2015): 188–196.
 - Dunietz, Jesse, Lori Levin, and Miriam R. L. Petruck. "Construction Detection in a Conventional NLP Pipeline." *AAAI Spring Symposium Technical Report SS-17-02: Computational Construction Grammar and Natural Language Understanding* (2017).
 - Jesse Dunietz, Annotating and Automatically Tagging Constructions of Causal Language, Ph.D. Thesis, Carnegie Mellon University, 2018.
 - Jesse Dunietz, Jaime Carbonell and Lori Levin, "DeepCx: A transition-based approach for shallow semantic parsing with complex constructional triggers". EMNLP 2018
- Thank you: Nathan Schneider, Miriam Petruck, Alexis Palmer

Annotating Causality



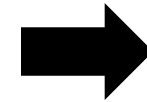
We annotate three types of causality.

The system failed **because of**
a loose screw.



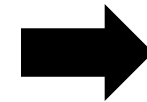
CONSEQUENCE

Mary left **because** John was
there.



MOTIVATION

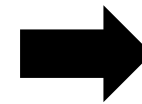
Mary left **in order to** avoid John.



PURPOSE

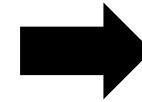
Causation can be positive or negative.

This has often **caused**
problems elsewhere.



FACILITATE

He **kept** the dog
from leaping at her.



INHIBIT

The Because Corpus

	Documents	Sentences	Causal
New York Times Washington section (Sandhaus, 2014)	59	1924	717
Penn TreeBank WSJ	47	1542	534
2014 NLP Unshared Task in PolilInformatics (Smith et al., 2014)	3	772	324
Manually Annotated Sub-Corpus (Ide et al., 2010)	12	629	228
Total	121	4790	1803

BECAUSE = **B**ank of **E**ffects and **C**auses **S**tated **E**xplicitly

Annotators were guided by
a “**construction**.”

Connective pattern	<cause> prevents <effect> from <effect>	<enough cause> for <effect> to <effect>
Annotatable words	prevent, from	enough, for, to
WordNet verb senses	prevent.verb.01 prevent.verb.02	
Type	Verbal	Complex
Degree	INHIBIT	FACILITATE
Type restrictions	Not PURPOSE	
Example	His actions prevented disaster.	There’s enough time for you to find a restroom.

Causality uses structures from other domains

After a drink, she felt much better.	(Temporal)
They're too big to fail.	(Extremity)
The more I read his work, the less I like it.	(Correlation)
The police let his sister visit him briefly.	(Permission)
As voters get to know Mr. Romney, his poll numbers will rise.	(Temporal + (Correlation)

We annotate 7 different types of overlapping relations.

TEMPORAL	After; once; during
CORRELATION	As; the more...the more...
HYPOTHETICAL	If...then...
OBLIGATION/PERMISSION	Require; permit
CREATION/TERMINATION	Generate; eliminate
EXTREMITY/SUFFICIENCY	So...that...; sufficient...to...
CONTEXT	Without; when (circumstances where...)

Overlapping Relation Examples

- Temporal
 - Within minutes after the committee released its letter, Torricelli took the senate floor to apologize to the people of New Jersey.
- Correlation
 - Auburn football players are reminded of last year's losses every time they go into the weight room.
- Hypothetical
 - Previously, he allowed increases in emissions as long as they did not exceed the rate of economic growth.

Overlapping Relation Examples

- Obligation/Permission

- He will roll back a provision known as a new source review that **compels** utilities to install modern pollution controls whenever the significantly upgrade older plants.
 - “whenever” is also a connective

- Creation/Termination

- Many expected synergies of financial service activities **gave rise to** conflicts and excessive risk taking.

- Context

- **With** Hamas controlling Gaza, it was not clear that Mr. Abbas had the power to carry out his decrees.

Temporal with and without causal interpretation

After last year's fiasco, everyone is being cautious.
MOTIVATION + TEMPORAL ARGC ARG E

After last year's fiasco, they've rebounded this year.
TEMPORAL ARGC ARG E

Conditional hypotheticals don't have to be causal, but most are.

Non-causal:	If he comes, he'll bring his wife.
Causal:	If I told you, I'd have to kill you.

84% carry causal meaning

Causality has seeped into
the temporal and hypothetical domains.

Of the causal expressions in the corpus:

- > 14% are piggybacked on temporal relations
- ~7% are expressed as hypotheticals

Opinions and Advice

How to run a multi-lingual design project

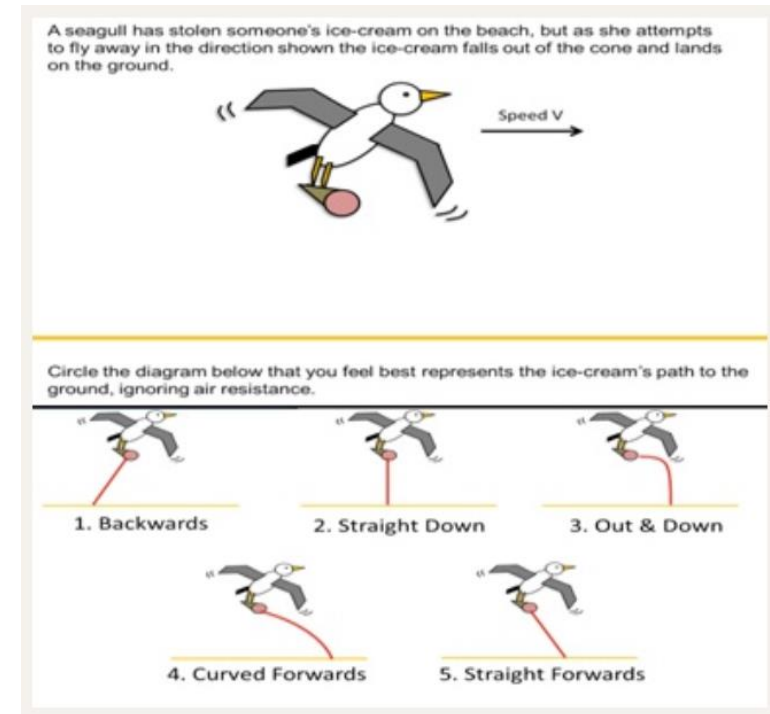
- Know your phenomenon and your languages: read reference grammars and talk to linguists
 - Semantic maps
 - Morphosyntactic strategies
 - Overlapping semantic domains
 - Morphosyntactic Recruitment
- Iterations
 - Do not get too big to fail all at once. Make a few preliminary versions that you intend to revise.
- Reconcile together:
 - Developers for different languages should meet regularly with each other.
- How do you know when it is good enough?

Your new hobbies

- Reading typology papers and reference grammars
 - If you can't do it: always consult a typologist
- Reading papers on classifiers in Semeval and Non-prop/ex-prop, and LAW

Opinion on crowdsourcing

- I don't recommend it for most types of meaning
- But you can factor out some of the more expert part from some of the less expert part.
 - PropBank
 - Because Constructicon



Naïve science: not recommended for real-world applications

Lobby for increased funding for resource creation

- Resource creation is not funded by the National Science Foundation in the US
- There is intellectual merit
- There are intrinsic and extrinsic evaluation metrics
 - We need to make this case to the funding agencies

End