Semantic role labeling

Hakyung Sung
University of Hawaii at Manoa

November 4, 2026

Topics to cover

- Review
- Semantic role labeling
- Presentation
- Lab

Review:
Linguistic annotations



- **Process my corpus sample**
- 1. raw texts
- 2. tagged texts
- 3. annotated texts
- 4. texts turned into embedding

Some annotation examples

- Syntactic annotations (e.g., *treebanks*)
- Phonetic and prosodic annotations
- **Semantic annotations (more on this today!)**
- Sentiment/Discourse annotations (more on this next class)

When no tool fits the task

- How can we find the linguistic feature we want to study?
- **Rule-based approach: Write the rules**
  - Use what you know about the language to define rules.
  - Apply those rules to find the feature in new texts.
- **Data-driven approach: Learn from examples**
  - Label examples of the feature by hand.
  - Train a model on those examples to find the feature in new texts.

Good annotation practice

- Keep the original text (e.g., L2 errors?)
- Document the annotation process
  - Make your project replicable/accessible to other researchers

Evaluating annotation

- Validity: Does the scheme capture meaningful linguistic distinctions?
- **Accuracy**: Do predicted labels match the gold labels?
- Consistency: Do independent annotators apply the scheme in the same way?

Accuracy

- Compare assigned labels with a reference or “gold” standard.
- Accuracy: The proportion of evaluated units labeled correctly.

$$A = \frac{C}{N}$$

- C: Number of correctly labeled units
- N: Total number of evaluated units

Precision/Recall: Identifying nouns

- Suppose `NOUN` is the target category.

	Gold: NOUN	**Gold: Not NOUN**
Predicted: NOUN	True positive (TP)	False positive (FP)
Predicted: Not NOUN	False negative (FN)	True negative (TN)

- **Precision**: Of the tokens labeled as nouns, how many really are nouns?

$$P = \frac{TP}{TP + FP}$$

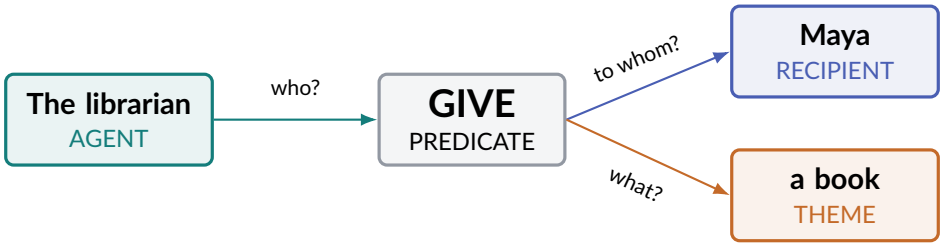
- **Recall**: Of all actual nouns, how many did the model identify?

$$R = \frac{TP}{TP + FN}$$

Semantic roles

Semantic roles: *Who did what to whom?*

The librarian **gave** Maya a book.



Syntax describes grammatical relations (subject, object, etc.).
Semantic roles describe participants' relationships to an event.

Other semantic roles

- **Force** (non-volitional cause)
  The wind opened the door.

- **Agent** (intentional actor)
  Maya opened the door.

- **Instrument** (tool used)
  Maya opened the door with a key.

- **Beneficiary** (who benefits)
  Ann made reservations for her boss.

- **Source** and **Goal** (movement)
  I flew from Boston to Portland.

Different syntax, Same semantic roles

- **Sentence A: Double-object construction**

  The librarian gave Maya a book.

- **Sentence B: Prepositional dative** (dative alternations)

  The librarian gave a book to Maya.

Agent **Theme** **Recipient**

Diathesis alternations

- The same semantic roles can appear in different syntactic structures.

- Consider the verb *melt*:
  1. Sarah melted the ice.
  2. Sarah melted the ice with a heat lamp.
  3. The heat lamp melted the ice.
  4. The ice melted.
  5. The ice was melted by Sarah.

- Sometimes called **verb** alternations

Verb alternations and verb classes

- **Verbs** with similar meanings often share patterns of argument realization.
- Levin (1993) classified 3,000+ **English verbs** based on their syntactic alternations.
- This work influenced lexical resources such as *VerbNet* (Schuler, 2005).

Challenges in defining semantic roles

- But can we use the same set of semantic roles across all verbs?
  - Broad roles (e.g., agent, theme) may not capture all differences across predicates.
    - She pushed the cart.
    - She feared the dog.
  - There is no universally agreed-upon inventory of semantic roles.

- Other approaches:
  - Proto-roles (Dowty, 1991):
    - Semantic roles are clusters of properties, not fixed categories.
    - Proto-Agent: Volition, sentience, causation, etc.
    - Proto-Patient: Change of state, affectedness, etc.
    - e.g., She melt the ice.
  - **Predicate-specific roles:** PropBank

PropBank

Semantic role annotation

- PropBank (Palmer et al., 2005) annotates predicates and their arguments with semantic roles. Each predicate sense has its own roleset (a set of argument definitions).

- Core arguments are labeled using numbers:
  - ARG0: Typically Proto-Agent
  - ARG1: Typically Proto-Patient
  - ARG2–ARG4: Other predicate-specific roles

- **Note that the meaning of each numbered argument depends on the predicate.**

Palmer, M., Gildea, D., & Kingsbury, P. (2005). The Proposition Bank: An annotated corpus of semantic roles. Computational Linguistics, 31(1), 71–106.

Core arguments

- Consider the predicate *increase*.

- **Active voice**

  [ARG0 Big Fruit Co.] increased [ARG1 the price of bananas].

- **Passive voice**

  [ARG1 The price of bananas] was increased again [ARG0 by Big Fruit Co.].

Modifier arguments (ARGM)

- PropBank also annotates modifier arguments using ARGM labels.
- Unlike numbered arguments, their meanings are relatively consistent across predicates.
- Common examples:
  - **ARGM-TMP**: Time (when?)
  - **ARGM-LOC**: Location (where?)
  - **ARGM-DIR**: Direction (where to/from?)
  - **ARGM-MNR**: Manner (how?)
  - **ARGM-CAU**: Cause (why?)
  - **ARGM-PRP**: Purpose (for what purpose?)

Let’s practice

- Consider the sentence:

  (1) Yesterday, Maya opened the door carefully at home.

  (2) After carefully reviewing the proposal, the committee gave the student a research grant at the annual meeting.

Frame-based semantic roles

- FrameNet organizes words around semantic frames (Baker et al., 1998; Fillmore @ UC Berkeley).

- A frame represents background knowledge about a particular event or situation.

- Each frame includes:
  - Lexical units: Words that evoke the frame
  - Frame elements: Semantic roles specific to the frame

- e.g., *Change_position_on_a_scale* (Frame)
  - The price increased from $5 to $8.
  - Item: the price
  - Initial_value: $5
  - Final_value: $8

- `https://framenet.icsi.berkeley.edu/intro`

Semantic Role Labeling (SRL)

Semantic Role Labeling (SRL)

- SRL: Automatically identify predicates and assign semantic roles to their arguments.
- Traditionally trained on annotated resources (e.g., PropBank).

ID	FORM	LEMMA	UPOS	XPOS	FEAT	HEAD	DEPREL	UP:PREDS	UP:ARGHEADS	UP:ARGSPANS
1	Ich	ich	PRON	PPER	_	2	nsubj	_	_	_
2	hatte	haben	VERB	VAFIN	_	0	root	have.03	A0:1|A1:3	A0:1-1|A1:3-3
3	Gelegenheit	Gelegenheit	NOUN	NN	_	2	obj	_	_	_
4	eines	ein	DET	PIS	_	6	det	_	_	_
5	seiner	sein	DET	PPOSAT	_	6	det:poss	_	_	_
6	Seminare	Seminar	NOUN	NN	_	8	obj	_	_	_
7	zu	zu	PART	PTKZU	_	8	mark	_	_	_
8	besuchen	besuchen	VERB	VVINF	_	3	xcomp	attend.01	A0:1|A1:4	A0:1-1|A1:4-6
9	.	.	PUNCT	.	_	2	punct	_	_	_

The German verbs

- 'hatte' is labeled as evoking the '**have.03**' frame with two roles: "Ich" (*I*) is labeled **A0** (owner) and "Gelegenheit" (*opportunity*) is labeled **A1** (possession).
- 'besuchen' is labeled as evoking the '**attend.01**' frame with two roles: "Ich" (*I*) is labeled **A0** (thing attending) and "eines seiner Seminare" (*one of his seminars*) is labeled **A1** (thing attended).

https://universalpropositions.github.io/; https://vimeo.com/161718580

Traditional SRL: Feature engineering

- Earlier SRL systems relied on manually designed linguistic features.
- **Step 1**: Parse the sentence and identify candidate arguments.
- **Step 2**: Extract features such as:
  - Phrase type (e.g., NP, PP)
  - Head word
  - Position relative to the predicate
  - Syntactic path to the predicate
  - Predicate properties
- **Step 3**: Train a classifier to assign semantic role labels (including NONE).

Independent predictions fail...

- Consider: *Sarah baked her friend a cake.*
- **Independent predictions:**
  - her friend $\rightarrow$ ARG1
  - a cake $\rightarrow$ ARG1
- **Problem**: Duplicate ARG1 labels!
- **Solution**: Joint/constrained decoding
  - Consider predictions for all arguments together.
  - Apply constraints on argument structure.
  - Select the best overall combination of labels.

Neural SRL: Sequence Labeling

- Neural SRL can be formulated as a token-level sequence labeling task.
- One common approach uses BIO tagging:
  - **B**: Beginning of an argument span
  - **I**: Inside an argument span
  - **O**: Outside an argument span
- Example: *The cats love hats.*

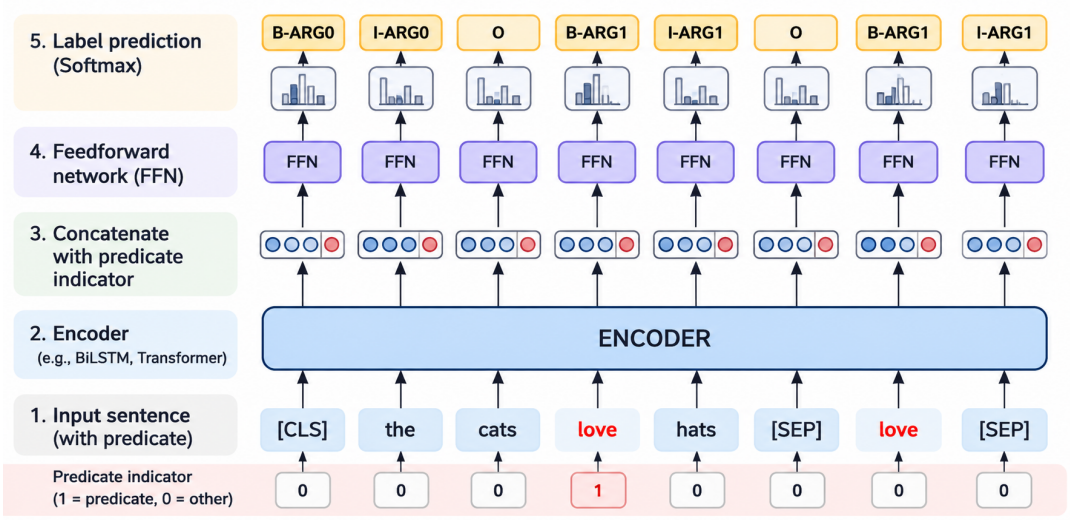
Token	**Label**
The	B-ARG0
cats	I-ARG0
love	B-V
hats	B-ARG1

Neural SRL: How Does It Work?

- **Step 1**: Input a sentence and identify the target predicate.
- **Step 2**: Use a neural encoder to generate contextualized token representations.
- **Step 3**: Predict possible semantic role labels for each token.
- **Step 4**: Decode the best label sequence.

$$\hat{y} = \arg\max_{y \in T} P(y \mid w)$$

- Select the highest-scoring valid sequence of semantic role labels.


5. Label prediction (Softmax)
4. Feedforward network (FFN)
3. Concatenate with predicate indicator
2. Encoder (e.g., BiLSTM, Transformer)
1. Input sentence (with predicate)
Predicate indicator (1 = predicate, 0 = other)
B-ARG0
I-ARG0
O
B-ARG1
I-ARG1
O
B-ARG1
I-ARG1
FFN
ENCODER
[CLS]
the
cats
love
hats
[SEP]
love
[SEP]
0
0
0
1
0
0
0
0

Beyond SRL: Abstract Meaning Representation (AMR)

- **SRL**: Identifies semantic arguments of individual predicates.
- **AMR**: Represents sentence meaning as a *structured* graph.
- Example: *The boy wants to go.*

```
(w / want-01
   :ARG0 (b / boy)
   :ARG1 (g / go-02
            :ARG0 b))
```

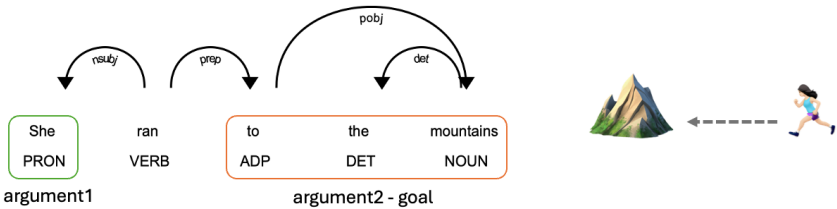
- The same boy is associated with both *want* and *go*.
- AMR also represents relations among events, negation, modifiers, and other semantic information.

Presentation: Ettinger et al. (2023)

"You Are An Expert Linguistic Annotator":
Limits of **LLMs as Analyzers** of **Abstract Meaning Representation**

Lab:
Training models

Background



- Argument structure constructions (ASCs) are (lexico-)grammatical patterns at the clausal level.
- They consist of an argument structure and a main verb, with each argument contributing to the clause's meaning.

Background

ASCs must be reliably identified in target texts.

ASC categories

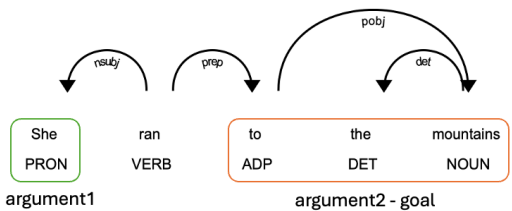
Annotation Tag	**Syntactic Frame**	**Semantic Frame**
Intransitive simple (INTRAN_S)	*nsubj*-root	*agent*-V
Intransitive motion (INTRAN_MOT)	*nsubj*-root-*obl*	*theme*-V-*goal*
Intransitive resultative (INTRAN_RES)	*nsubj*-root-*advmod*	*patient*-V-*goal*
Transitive simple (TRAN_S)	*nsubj*-root-*obj*	*agent*-V-*theme*
Ditransitive (DITRAN)	*nsubj*-root-*iobj*-*obj*	*agent*-V-*recipient*-*theme*
Caused-motion (CAUS_MOT)	*nsubj*-root-*obj*-*obl*	*agent*-V-*theme*-*destination*
Transitive resultative (TRAN_RES)	*nsubj*-root-*obj*-*xcomp*	*agent*-V-*result*-*result*
Attributive (ATTR)	*nsubj*-*cop*-root	*theme*-V-*attribute*
Passive (PASSIVE)	*nsubj:pass*-*aux:pass*-root	*theme*-*aux*-V_{passive}

Potential approaches

1. Manual annotation of all ASC types:
   - *She ran to the mountain*
   - Read the sentence
   - Annotate based on the judgment
   - 👍High control of the data
   - 👎Impossible to do with large data

Potential approaches

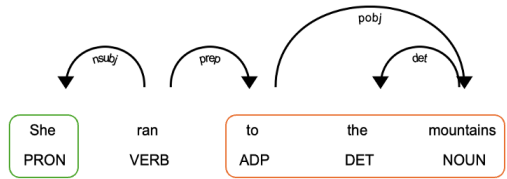
2 Automatic extraction of ASCs using **dependency** structures:



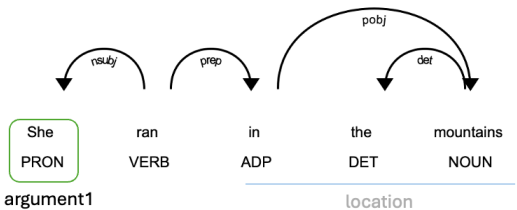


Background

2 Automatic extraction of ASCs using **dependency** structures:









- 👍Scale up the ASC extractions
- 👎Lack the semantic information needed to disambiguate certain ASC types

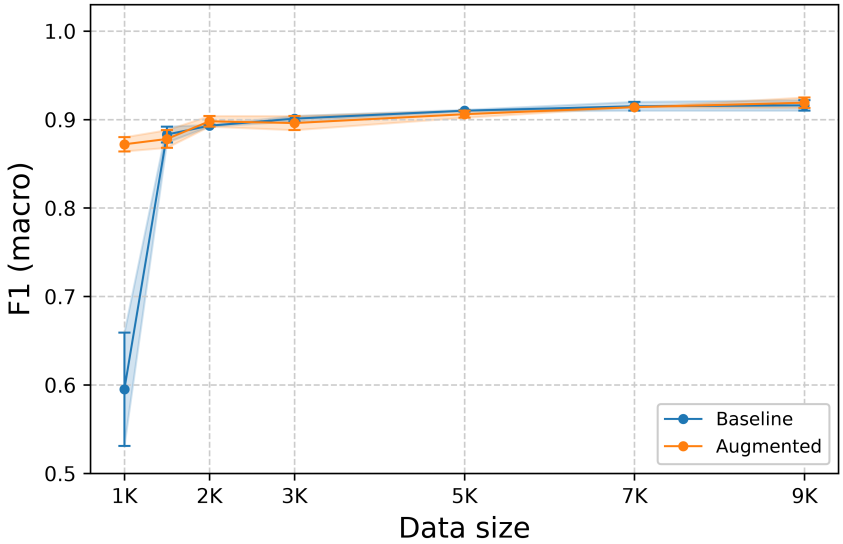
Background

3 Automatic extraction of ASCs using **SRL**:

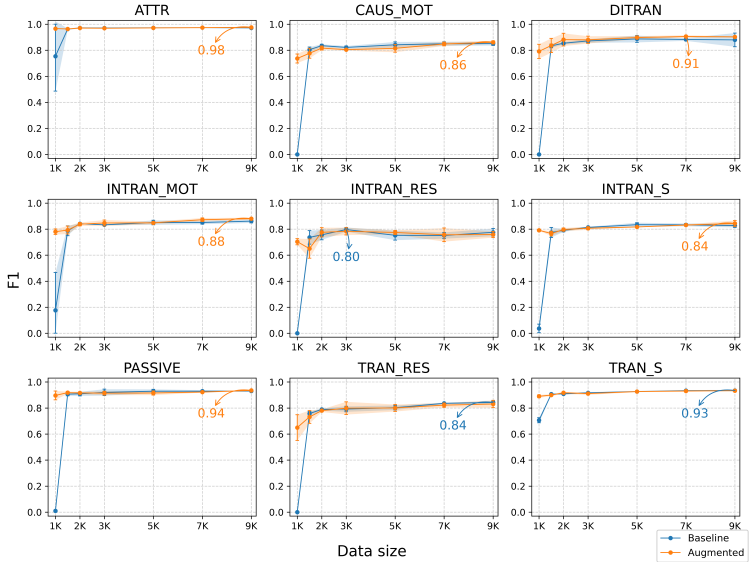
- The meaning of each numbered argument depends on the predicate
- Less accurate than dependency parser

3 Annotate a “small” sample + train LLMs

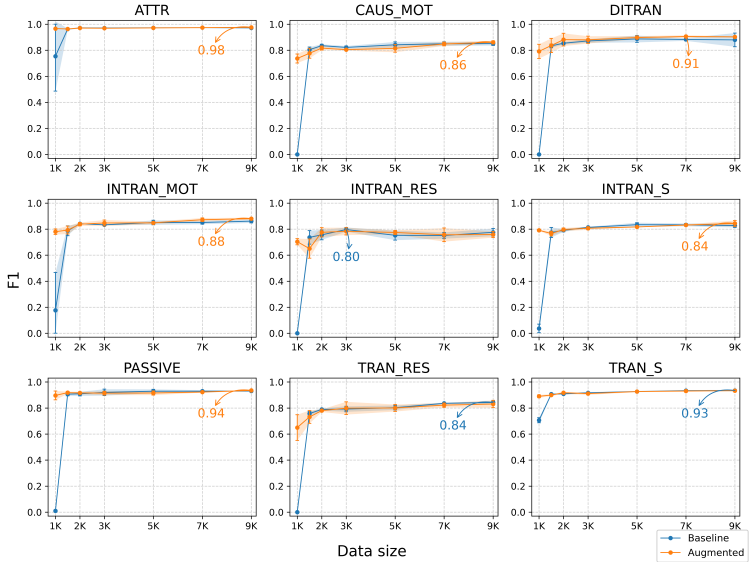
- How small?



Kyle, K., & Sung, H., (*forthcoming*). Using encoder-based and generative large language models to facilitate linguistic annotation. Applied Linguistics.

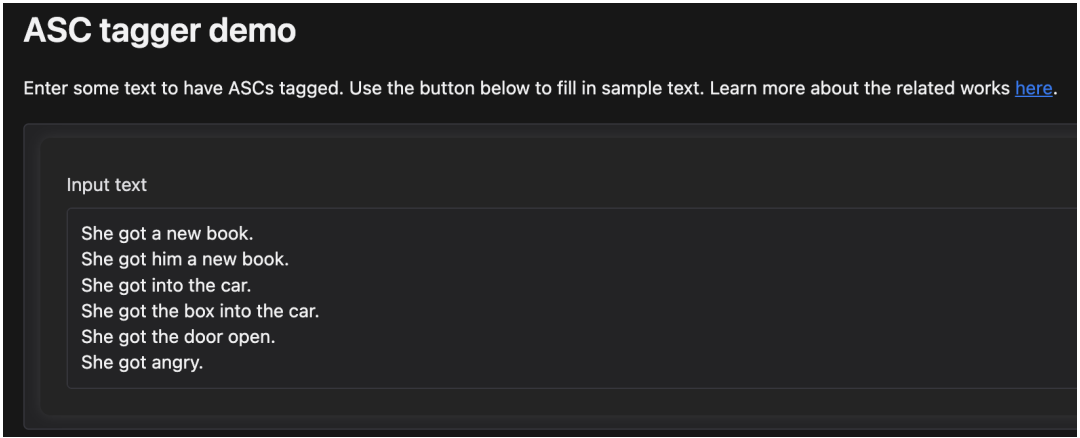


Kyle, K., & Sung, H., (*forthcoming*). Using encoder-based and generative large language models to facilitate linguistic annotation. Applied Linguistics.



Kyle, K., & Sung, H., (*forthcoming*). Using encoder-based and generative large language models to facilitate linguistic annotation. Applied Linguistics.

Task: Token-level annotation



https://huggingface.co/spaces/hksung/ASC_tagger_V2

Task: Token-level annotation



https://huggingface.co/spaces/hksung/ASC_tagger_V2