English ASC Treebank Annotation Manual

Hakyung Sung and Kristopher Kyle
LCR-ADS Lab, Linguistics, University of Oregon
Last updated: March 1, 2024

Contents

1 Introduction

The ASC treebank is a corpus in which sentences are annotated with argument structure constructions (ASCs) in relation to the predicate. This project builds upon earlier related works, such as PropBank (Palmer et al., 2005), FrameNet (Fillmore et al., 2003), VerbNet (Schuler, 2005), and Universal PropBank (Akbik et al., 2015). To explore more about our research, please refer to the following papers:

- Kyle, K. & Sung, H. (2023). An argument structure construction treebank, In *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, 51–62, Association for Computational Linguistics.
- Sung, H. & Kyle, K. (2024). Annotation scheme for English argument structure constructions treebank, In *Proceedings of the 18th Linguistic Annotation Workshop (LAW-XVIII)*.

1.1 Development phases of the ASC treebank

The ASC Treebank has been developed through a multi-phase approach, detailed below:

v1: The first version of the ASC treebank was developed using a semi-automatic (silver) approach, comprising the following steps. For more detailed information, see Kyle and Sung, 2023.

1. The Universal Dependencies version of the English Web Treebank (EWT; Bies et al., 2012; Silveira et al., 2014) was used from the Universal PropBank project Akbik et al. (2015).
2. The large-grained argument structures were extracted based on PropBank's semantic role labels (e.g., *ARG0-Verbsense-ARG1*) from the dataset.
3. These large-grained semantic tags were then converted into fine-grained semantic role frames (e.g., *agent-Verbsense-theme*) using mapping protocols from PropBank frame files, which incorporate insights from both FrameNet (Fillmore et al., 2003) and VerbNet (Schuler, 2005).
4. Each fine-grained semantic role frame was manually assigned an ASC (e.g., *agent-Verbsense-theme* $\rightarrow$ Transitive simple ASC).
5. After addressing ambiguities and conducting spot checks for errors, 94.1% of the ASCs in the EWT Treebank were categorized.

v2: The second version of the ASC treebank was annotated using a manual (gold) approach. For more detailed information, see Sung and Kyle (2024).

1. The remaining uncategorized ASCs in the EWT treebank, constituting 5.9% of the ASCs, were classified. For both categorized and uncategorized ASCs in the EWT treebank, manual annotations were carried out with two distinct purposes: (1) the annotation of the uncategorized ASCs, and (2) the verification of the tagging accuracy of the semi-automatic approach used in the first phase by comparing it with the manual annotations.
2. The L2 English (ESL) written corpus (the Treebank of Learner English [ESL-WR]; Berzak et al.,

2016) and the L2 spoken corpus (the Treebank of Spoken L2 English [ESL-SP]; Kyle et al., 2022) were manually annotated.

3. In total, **22,069** ASCs (**10,204** sentences) were annotated (EWT: 5,936 with 104,640 word tokens; ESL-WR: 1,948 with 37,055 tokens; ESL-SP: 2,320 with 21,312 tokens).

Tag	EWT	ESL-WR	ESL-SP
INTRAN_S	1,395	662	525
INTRAN_MOT	607	250	240
INTRAN_RES	213	44	23
TRAN_S	6,094	2,488	1,385
DITRAN	285	160	37
CAUS_MOT	766	87	53
TRAN_RES	763	76	16
ATTR	2,539	1,289	760
PASSIVE	1,058	224	50
Total	13,720	5,260	3,089

Table 1: Overview of ASCs distribution in the treebank

1.2 Manual outline

This manual aims to help ongoing projects in achieving reliable and valid ASC annotations within English datasets. Herein, we outline the annotation scheme (§2) and the steps involved in the annotators training process (§3). Subsequently, we enumerate the nine ASC categories we have identified in our projects, defining each with specific features and illustrating them with examples from L1 and L2 datasets (§4). For the examples from the semi-automatic annotation version (v1), we clarify them using specific semantic frames (based on the semantic role labels from previous projects). Additionally, we address particular examples that posed challenges for our annotators in a separate discussion section (§5).

2 Annotation Scheme

2.1 Theoretical background

- The **usage-based constructionist approach** posits that language development emerges from the statistical induction of form-meaning pairings within linguistic structures. In this process, learners cultivate and categorize abstract entities known as **constructions**, as they interact with varied language inputs and outputs.
- Among the variety of construction types, **argument structure constructions (ASCs)** have become particularly noteworthy. These clause-level units articulate core concepts of sentence meaning (i.e., (i.e., *who* did[VERB] *what* to *whom*, with arguments being italicized) (Fillmore, 1975) and serve an

important role in communication (Bencini and Goldberg, 2000; Goldberg, 1995, 2003, 2006, 2013; O'Connor and Kay, 2003; Rappaport Hovav and Levin, 1998).

- In **cognitive linguistics**, the significance of ASCs is highlighted by their role in human cognition and perception. At the heart of our experiences are fundamental scenes, such as motion events or causative actions, that are crucial to human understanding. ASCs are closely related to these *conceptual archetypes* (Langacker, 1987); Bowerman, 1996; Clark, 1978; Slobin, 1985).
- In **language learning**, understanding the relationship between ASCs and the verb is important. First, ASCs are often generalized based on verbs that are frequently and strongly associated with each construction. Moreover, as learners produce language, verb choices may shape ASC patterns; not every verb fits into every construction, and conversely, not all ASCs accommodate every verb. This distinction in verb-construction combinations has been highlighted as a significant difference between novice and advanced language speakers (**L1**: Cameron-Faulkner et al., 2003; Goldberg, 2013, 2019; Ninio, 1999; **L2**: Ellis, 2002; Ellis and Larsen-Freeman, 2009; Ellis and Ferreira-Junior, 2009; Kyle and Crossley, 2017).
- **Semantic roles and ASCs**: Each ASC comprises "argument roles". These roles often align with traditional semantic roles, such as *agent, patient, theme, source, goal, path, etc.* Rather than delving into the more concrete and specific meanings that verbs introduce to the constructions (i.e., "participant roles" – which aligns closely with the concept of frame elements in the FrameNet project [1]), we focus on the argument roles of each ASC type. This is because they capture the generalizations across the unique participant roles that individual verbs present within constructions (For more theoretical explanation on "argument roles" and "participant roles", see Goldberg, 1995, 2013).
- **ASC Categories**: Goldberg (2013, p. 455) suggested that ASCs correspond to unique grammatical relations that are deeply rooted in semantics. These relations vary across active, passive, and middle voices because the same event can be represented differently depending on the speaker's focus. For example, *The car hit the man* can be recast as *The man was hit by a car*, a phenomenon referred to as 'allostructions' in previous studies (e.g., Bergen and Chang, 2005; Cappelle, 2006).

  The majority of previous works that we investigated primarily focused on ASCs in the active voice. Consequently, we initially included seven primary ASCs in our analysis: (1) Intransitive simple (active voice), (2) Intransitive motion, (3) Intransitive resultative, (4) Transitive simple, (5) Ditransitive, (6) Caused-motion, and (7) Transitive resultative. However, to capture a broader range of constructions appearing in English production datasets—especially in L2 datasets, which are of particular interest to our research—we expanded our categorization to include two additional constructions.

  The (8) Attributive construction correlates with an attribute, often represented by the subject predicate following the verb (e.g., *I am a student*). The (9) Passive construction places the subject as the undergoer/patient rather than the executor of the action denoted by the verb (e.g., *the phone was smashed*).

[1] https://framenet.icsi.berkeley.edu

Additionally, we classified the middle voice, which encapsulates reflexive or reciprocal actions wherein the subject plays both the actor and recipient roles (e.g., *the book reads easily*). We categorized this under (1) Intransitive simple, primarily due to similarities in form (nsubj-root [VERB]). In the constructions we introduced, the subjects may not always be cast in the agent role, distinguishing them from the initially introduced ASC types (1-7).

2.2 Data format: CoNLL-U

We used data structured in the CoNLL-U format[2]. The format represents each sentence vertically, organizing them into columns that convey morphological and syntactic information. The columns are presented in the sequence below, with each number corresponding to the column index:

1. ID: Word index
2. FORM: Word form or punctuation symbol
3. LEMMA: Lemma or stem of word form
4. UPOS: Universal part-of-speech tag
5. XPOS: Language-specific part-of-speech tag; underscore if not available
6. FEATS: List of morphological features from the universal feature inventory
7. HEAD: Head of the current word, which is either a value of ID or zero (0)
8. DEPREL: Universal dependency relation to the HEAD (root iff HEAD = 0)
9. DEPS: Enhanced dependency graph in the form of a list of head-deprel pairs
10. **MISC**: Any other annotation – **where the ASC tags will be annotated.**

```
# text = You called me!
1	You	you	PRON	PRP	Case=Nom	2	nsubj	_	_
2	called	call	VERB	VBD	Tense=Past	2	ROOT	_	TRAN_S
3	me	I	PRON	PRP	Case=Acc	2	dobj	_	_
4	!	!	PUNCT	.	Punct	2	punct	_	_
```

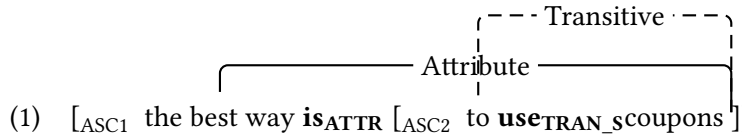
Figure 1: Example of data in the CoNLL-U format

2.3 The unit of analysis

- Each ASC is the unit of the analysis, with arguments bearing fixed semantic roles serving as its building blocks. These semantic roles include *agent, patient/theme, goal, path, result*, etc.
- The verb is a tagging place because it interacts with the surrounding arguments within the construction. In the CoNLL-U format, the columns 4 and 8 are helpful for annotators to identify the verb (if the data comes with the relevant tags). This is because the UPOS column (4th) shows the VERB tag, and in many instances, this aligns with the `root` tag in the DEPREL column (8th).

[2]https://universaldependencies.org/format.html

- Our annotations encompass both finite clauses (e.g., *the best way **is** to use coupons*) and non-finite clauses (e.g., *to **use** coupons*), allowing us to comprehensively capture the event structures delineated by ASCs. Consequently, a single sentence (which is identified as a distinct block with an individual ID in the CoNLL-U format) can comprise multiple layers of the ASCs. See the example (1) below. In our visual representations, a solid line represents the span of an ASC within a finite clause, whereas a dashed line signifies the span of an ASC in a non-finite clause.



Exceptional cases Annotators may encounter situations where (1) a different element (e.g., AUX) should be considered the basic unit of analysis instead of the VERB, (2) the verb itself is not suitable as the fundamental unit, or (3) there is no appropriate unit to tag.

  - Consider using an auxiliary (AUX) verb as the basic unit of analysis when it serves as a copular verb in an attributive construction (Figure 1).
  - In some instances, certain verbs function more as modifiers than as predicates. By examining the 8th column, annotators can identify tags indicative of modifiers, such as `amod` (adjectival modifier). Verbs fulfilling these roles should be bypassed (Figure 2).
  - Verbs in an `acl` dependency relationship should NOT be overlooked because they are predicates of adjectival clauses. In the subsequent example, 'fixed' will be tagged, but 'implied' will not (Figure 2).
  - There are instances where a sentence may have a root but lacks a verb[3], especially in data sourced from spoken discourse. In these cases, we do NOT tag any ASC to the case (Figure 3).

3 Annotation Process

3.1 Annotator training

- Undergraduate Linguistics majors who had taken upper-level courses in functional English syntax were recruited for the project. The annotators underwent three in-person training sessions, each lasting an hour.

[3]The root in UD is basically for marking the root of a sentence, in most cases, a main verb of the sentence. When there is no verb, the main content word or the word that carries the primary meaning of the sentence gets the root.

```
# text = They are currently using 9.5% fixed based on the 1 year implied volatility.
1	They	they	PRON	PRP	4	nsubj	4:nsubj	_
2	are	be	AUX	VBP	4	aux	4:aux	_
3	currently	currently	ADV	RB	4	advmod	4:advmod	_
4	using	use	VERB	VBG	0	root	0:root	TRAN_S
5	9.5	9.5	NUM	CD	6	nummod	6:nummod	_
6	%	%	SYM	NN	4	obj	4:obj	_
7	fixed	fix	VERB	VBN	6	acl	6:acl	_
8	based	base	VERB	VBN	14	case	14:case	PASSIVE
9	on	on	ADP	IN	14	case	14:case	_
10	the	the	DET	DT	14	det	14:det	_
11	1	1	NUM	CD	12	nummod	12:nummod	_
12	year	year	NOUN	NN	14	compound	14:compound	
13	implied	imply	VERB	VBN	14	amod	14:amod	_
14	volatility	volatility	NOUN	NN	4	obl	4:obl:on	_
15	.	.	PUNCT	.	4	punct	4:punct	_
```

Figure 2: Example of verbs functioning as modifiers (`amod`) and predicates of adjectival clauses (`acl`)

```
# text = All right .
1	All	_	ADV	RB	_	2	advmod	_	_
2	right	_	ADJ	JJ	_	0	root	_	_
3	.	_	PUNCT	.	_	2	punct	_	_
```

Figure 3: Example of a sentence that does not contain a verb

- **Session 1:** Annotators were acquainted with the theoretical background, the objectives of the project, their expected roles, the data format, tag categories, procedures for loading/saving data from the shared folder, and an overview of the entire project.
- **Sessions 2 & 3:** Annotators were tasked with tagging sample items independently. Feedback was then given based on their performance on these sentences.

3.2 Annotation work and discussions

- After the training sessions, the annotators began their work remotely. Consequently, all ongoing and completed files were directed through a shared folder. This folder was supervised by the researchers to ensure consistent performance monitoring.
- Throughout the annotation phase, annotators could refer to a detailed documentation web page made by the researchers. A Discord server was set up to facilitate the reporting and discussion of challenging cases asynchronously. Additionally, an extended tagging manual was available, regularly updated based on these discussions.

4 ASC Categories: Descriptions and examples

In this section, we provide comprehensive description of each ASC tag. We began each category with a broad overview of the construction, followed by its **syntactic frame**, which is presented as dependecy graphs in the form of head-deprel pairs, leveraging tags from the Universal Dependency project (De Marn-

effe et al., 2021; Nivre et al., 2016, 2017). Alongside, we discuss the **semantic frames**, which are derived from the Universal Propbank (For detailed information on how we extracted semantic roles, see Kyle and Sung, 2023, pp.53-54). To enhance understanding, we provide concrete examples from both L1 dataset (EWT) and L2 datasets (as mostly presented in further cases - including both written (ESL-WR) and spoken (ESL-SP) data). Note that only the examples sourced from the L1 dataset display semantic role tags beneath the arguments. Some of the sentences are streamlined for better readability.

4.1 Intransitive simple

Description The intransitive simple (INTRAN_S) ASC is characterized by the presence of a singular argument that precedes the verb (when the ASC constitutes a clause and the argument functions as its subject). Semantically, this category can be specified two discrete types: (1) an action both initiated and carried out by an agent, often as in the active voice, or (2) a state or condition undergone by a theme, which aligns with the characteristics of the middle voice, in which the subject both acts and is acted upon[4].

Syntactic frame nsubj-root

Semantic frame 1 X_{agent} VERB

(2) I_{agent} am **working$_{INTRAN_S}$** from our Hong Kong office for this week only (EWT)

(3) Will you_{agent} **drive$_{INTRAN_S}$** on that day (EWT)

(4) Do $you_{co\text{-}agent}$ **concur$_{INTRAN_S}$** (EWT)

Semantic frame 2 X_{theme} VERB

(5) [Her cereal bowl]$_{theme}$ **came$_{INTRAN_S}$** with a lifeguard (EWT)

(6) [$_{ASC1}$ Her driver's license **say** [$_{ASC2}$ "Picture$_{theme}$ **continued$_{INTRAN_S}$** on the other side"]] (EWT)

[4]English intransitive verbs often combine with an affected subject, a role commonly associated with direct objects; e.g., "...mother dropped her pot of soup the other day and it **broke** on the floor" (Biber et al., 1999 p.124).

(7) [$_{ASC1}$ [$_{ASC2}$ You$_{theme}$ will have to **wait$_{INTRAN_S}$**] and **see$_{INTRAN_S}$**] (EWT)

Further cases

(8) We **agreed$_{INTRAN_S}$** with each other (ESL-SP)

(9) [$_{ASC1}$ I am **writing$_{INTRAN_S}$**] [$_{ASC2}$ to **give** you the information [$_{ASC3}$ you **requested**]] (ESL-WR)

(10) [$_{ASC1}$ I **knew** [$_{ASC2}$ I could not **make** any movement for my safety]] [$_{ASC3}$ but I **did$_{INTRAN_S}$**] (ESL-WR)

(11) [$_{ASC1}$ Some people don't **like** [$_{ASC2}$ [to go **shopping$_{INTRAN_S}$** with a friend]] (ESL-WR)

4.2 Intransitive motion

Description The intransitive motion (INTRAN_MOT) ASC involves two arguments including a mover and a path. The mover, usually the subject of the sentence, undertakes some motion or movement, while the path characterizes the trajectory or direction of this movement. This path is often denoted using adverbial particles (e.g., *up, down, away*) or more descriptive prepositional phrases (e.g., *into the room, towards the park*).

Syntactic frame nsubj-root-obl; nsubj-root-advmod

Semantic frame 1 X$_{theme}$ VERB Y$_{goal}$

(12) People$_{theme}$ do not **go$_{INTRAN_MOT}$** [to Ireland]$_{goal}$ on holiday for the weather (EWT)

(13) [$_{ASC1}$ It **wasn't** [$_{ASC2}$ until he **gave up** [$_{ASC3}$ and **walked$_{INTRAN_MOT}$** [out the door$_{goal}$]]]] (EWT)

(14) I won't **go$_{INTRAN_MOT}$** [out the door$_{goal}$] (EWT)

Semantic frame 2 X$_{patient}$ VERB Y$_{goal}$

(15) [$_{ASC1}$ **Check out** the kids' space] [$_{ASC2}$ as soon as you$_{patient}$ **get$_{INTRAN_MOT}$** onboard$_{goal}$] (EWT)

(16) He won't **get close$_{INTRAN_MOT}$** [to the crate$_{goal}$] (EWT)

Semantic frame 3 X_{agent} VERB Y_{goal}

(17) So, obviously I$_{agent}$ **switched back$_{INTRAN_MOT}$** [to these guys]$_{goal}$ (EWT)

Semantic frame 4 X_{theme} VERB $Y_{ARGM\text{-}DIR}$

(18) [$_{ASC1}$ One can **see** [$_{ASC2}$ [numerous options]$_{theme}$ **passing$_{INTRAN_MOT}$** [through their mind]$_{ARGM\text{-}DIR}$](EWT)

(19) [$_{ASC1}$ Two weeks ago he$_{theme}$ **came$_{INTRAN_MOT}$** out$_{ARGM\text{-}DIR}$] [$_{ASC2}$ and **did** my horses feet] (EWT)

(20) When [the next hailstorm]$_{theme}$ **blows$_{INTRAN_MOT}$** through$_{ARGM\text{-}DIR}$ ...(EWT)

Semantic frame 5 $X_{patient}$ VERB $Y_{ARGM\text{-}DIR}$

(21) [$_{ASC1}$ In 2005 you'll **find** yourself [$_{ASC2}$ **pushed$_{INTRAN_MOT}$** [out into the world$_{ARGM\text{-}DIR}$]] (EWT)

(22) [$_{ASC1}$ Kittens are **chasing** each other], [$_{ASC2}$ **rolling$_{INTRAN_MOT}$** [on the ground]$_{ARGM\text{-}DIR}$] (EWT)

Semantic frame 6 X_{agent} VERB $Y_{ARGM\text{-}DIR}$

(23) [$_{ASC1}$ It tasted like [$_{ASC2}$ I$_{agent}$ **flew$_{INTRAN_MOT}$** [back home]$_{ARGM\text{-}DIR}$]] (EWT)

(24) [$_{ASC1}$ I$_{agent}$ finally **drove$_{INTRAN_MOT}$** there$_{ARGM\text{-}DIR}$] [$_{ASC2}$ to **place** my order in person] (EWT)

(25) [$_{ASC1}$ We$_{agent}$ **stopped$_{INTRAN_MOT}$** in$_{ARGM\text{-}DIR}$] [$_{ASC2}$ and **got** some take out] (EWT)

Further cases

(26) [$_{ASC1}$ After the hard work we **took** cool drinks], [$_{ASC2}$ **sat down$_{INTRAN_MOT}$**] [on the ground], [$_{ASC1}$ and enjoyed the pop concert] (ESL-WR)

(27) [$_{ASC1}$ Just as Pat **left$_{INTRAN_MOT}$** [my house]] [$_{ASC2}$ he **went away$_{INTRAN_MOT}$** [to the hospital]] (ESL-WR)

(28) [$_{ASC1}$ it **was** her best friend 's family] [$_{ASC2}$ that she must **lie$_{INTRAN_MOT}$** to] (ESL-WR)
goal fronted

(29) [$_{ASC1}$ It **contributes$_{INTRAN_MOT}$** [to the world treasurehouse of literature]] [$_{ASC2}$ and **arouses** irresistible fascination] (ESL-WR)

(30) [$_{ASC1}$ (At) event this year some halls, [$_{ASC2}$ which I had **visited$_{INTRAN_MOT}$**] were **overcrowded**] (ESL-WR)
goal fronted

Notes The following are some notes for later discussions.

- A discussion is needed regarding whether the locative adjunct (26) should be considered a goal or not.
- A discussion is needed regarding on (28). A questions from the reviewer was ... 'Is 'She lied to me' truly a motion predicate while 'She left' is not?' The characterization labeling 'she' as a 'mover' in the lie-clause raises further discussion. It seems that the classification hinges more on the presence of a prepositional phrase (PP) or other adverbial in the clause, and whether it can be construed directionally. While some PPs, such as 'She went to Portland,' clearly denote motion, this isn't universally applicable, as seen in the case of the 'lie' clause (28).

4.3 Intransitive resultative

Description The intransitive resultative (INTRAN_RES) ASC involves two arguments: a doer (i.e., a subject in a clause who carries out an action) and the result of the corresponding action, which signifies the transformed state or condition of the doer. The results in this construction is denoted using noun or adjective phrases to convey the outcome.

Syntactic frame nsubj-root-xcomp; nsubj-root-advmod

Semantic frame 1 $X_{patient}$ VERB Y_{result}

(31) Money$_{\text{patient}}$ may suddenly **become$_{\textbf{INTRAN_RES}}$** tight$_{\text{result}}$ at a very inconvenient time (EWT)

(32) Often [a cold wet morning]$_{\text{patient}}$ will **[turn into]$_{\textbf{INTRAN\_RES}}$** [a lovely warm sunny afternoon]$_{\text{result}}$ (EWT)

(33) He$_{\text{patient}}$ **[ended up]$_{\textbf{INTRAN\_RES}}$** [leaving one nail out of my horses foot]$_{\text{result}}$ (EWT)

Semantic frame 2 X$_{\text{patient}}$ VERB Y$_{\text{goal}}$

(34) After a couple of weeks, I$_{\text{patient}}$ **got$_{\textbf{INTRAN_RES}}$** [tired of them]$_{\text{goal}}$ (EWT)

(35) [$_{\text{ASC1}}$ **Thank** you for [$_{\text{ASC2}}$ **helping** [$_{\text{ASC3}}$ me$_{\text{patient}}$ **get$_{\textbf{INTRAN_RES}}$** [more healthy]$_{\text{goal}}$]]] (EWT)

Further Cases

(36) ... but finally [every thing] **went$_{\textbf{INTRAN_RES}}$** well (ESL-WR)

(37) And then, piano **turned$_{\textbf{INTRAN_RES}}$** just [(into) my hobby] (ESL-SP)

(38) [$_{\text{ASC1}}$ The toilet should **remain$_{\textbf{INTRAN_RES}}$** [the same]] [$_{\text{ASC2}}$ as people need to **use** (it) daily] (ESL-WR)

Notes The following are some notes for later discussions.

- A discussion is needed to clarify the semantic role labels. Although we adapted the labels directly from previous semantic role labeling projects, there are some confusions between the tags (e.g., result, goal) – we need to clarify the usage of 'goal' in motion events and resultatives.
- A discussion is needed to check some cases included in the Further Cases section (36, 38). A reviewer commented that they are not resultative because the result does not come about as a result of the verb.

4.4 Transitive simple

Description The transitive simple (TRAN_S) ASC describes a direct relationship between a subject performing an action and an object receiving that action. Therefore, this construction includes an agent (typically the subject of the clause, the entity that performs/initiates the action) and a theme/patient[5] (the entity that is affected by, or undergoes, the action carried by the agent). The construction is compatible with a wide range of the verbs and can represent:

- Physical activities: In the sentence "The cat chased the mouse", the "cat" is the agent performing the action, while the "mouse" is the theme/patient being acted upon.
- Mental activities: These describe cognitive or emotional states or processes. In "She loves chocolate". "she" is the agent experiencing the emotion, and the "chocolate" is the theme of her affection.
- Communicative activities: For instance, in "The professor taught the students", "the professor" is the agent conveying knowledge, and "the students" are the theme/patient receiving it.

Syntactic frame nsubj-root-obj; nsubj-root-ccomp

Semantic frame 1 X_{agent} VERB Y_{theme}

(39) I_{agent} should **buy$_{TRAN_S}$** $[a\ new\ one]_{theme}$ (EWT)

(40) **See$_{TRAN_S}$** you_{theme} there! (EWT)

(41) $[_{ASC1}$ It **pressured**$_{TRAN\_S}$ $Musharraf_{theme}$ $[_{ASC2}$ to **[take off]$_{TRAN\_S}$** $[his\ uniform]_{theme}$ $]]$(EWT)

(42) $[_{ASC1}$ He_{agent} **confirmed$_{TRAN_S}$** $[_{ASC2}$ $they_{agent}$ had not **received$_{TRAN_S}$** $weapons_{theme}$ $]]_{theme}$ (EWT)

(43) $[_{ASC1}$ Please **let** us $[_{ASC2}$ **know$_{TRAN_S}$** $[_{ASC3}$ if you_{pivot} **have$_{TRAN_S}$** $[a\ new\ one]_{theme}$ $]$ $]]$ (EWT)

Semantic frame 2 X_{pivot} VERB Y_{theme}

(44) $Megan_{pivot}$ does **have$_{TRAN_S}$** $[a\ couple\ of\ big\ things]_{theme}$ on the weekend in January (EWT)

[5] In this project, we did not differentiate between a theme and a patient.

(45) **Have$_{\text{TRAN_S}}$** you$_{\text{pivot}}$ [any thoughts on draft or cover note]$_{\text{theme}}$? (EWT)
inverted

(46) I$_{\text{pivot}}$ will **need$_{\text{TRAN_S}}$** [a rest from formal affairs]$_{\text{theme}}$ (EWT)

Semantic frame 3 X$_{\text{agent}}$ VERB Y$_{\text{topic}}$

(47) [$_{\text{ASC1}}$ She$_{\text{agent}}$ **said$_{\text{TRAN_S}}$** [$_{\text{ASC2}}$ a shower would **be** grand]$_{\text{topic}}$] (EWT)

(48) [$_{\text{ASC1}}$ I$_{\text{agent}}$ **recommend$_{\text{TRAN_S}}$** [$_{\text{ASC2}}$ that NASA$_{\text{agent}}$ **build$_{\text{TRAN_S}}$** [new shuttles]$_{\text{theme}}$]$_{\text{topic}}$] (EWT)

(49) Have you$_{\text{agent}}$ **read$_{\text{TRAN_S}}$** [the Declaration of Independence and US Constitution]$_{\text{topic}}$]? (EWT)

Semantic frame 4 X$_{\text{agent}}$ VERB Y$_{\text{result}}$

(50) [Increased car usage in China]$_{\text{agent}}$ is **creating$_{\text{TRAN_S}}$** [a high demand for petrol]$_{\text{theme}}$] (EWT)

(51) [$_{\text{ASC1}}$ If you$_{\text{agent}}$ **want$_{\text{TRAN_S}}$** [$_{\text{ASC2}}$ to **start$_{\text{TRAN_S}}$** [a new project]$_{\text{theme}}$]]$_{\text{result}}$, [$_{\text{ASC3}}$ your imagination will be **working** overtime] (EWT)

(52) [$_{\text{ASC1}}$ If sites$_{\text{agent}}$ don't **have$_{\text{TRAN_S}}$** [$_{\text{ASC2}}$ what$_{\text{result}}$ you$_{\text{agent}}$ **want$_{\text{TRAN_S}}$**]]$_{\text{theme}}$, [$_{\text{ASC3}}$ **contact$_{\text{TRAN_S}}$** [the archive]$_{\text{theme}}$] (EWT)
result fronted

Semantic frame 5 X$_{\text{experiencer}}$ VERB Y$_{\text{stimulus}}$

(53) [$_{\text{ASC1}}$ We$_{\text{experiencer}}$ **felt$_{\text{TRAN_S}}$** [$_{\text{ASC2}}$ we **made** it better]$_{\text{stimulus}}$] [$_{\text{ASC3}}$ because it$_{\text{agent}}$ **introduced$_{\text{TRAN_S}}$** [a process for review]$_{\text{theme}}$] (EWT)

(54) [$_{\text{ASC1}}$ Abby$_{\text{pivot}}$ **has$_{\text{TRAN_S}}$** friends$_{\text{theme}}$ [$_{\text{ASC2}}$ who$_{\text{experiencer}}$ **love$_{\text{TRAN_S}}$** her$_{\text{stimulus}}$]] (EWT)

Further Cases

(55) [$_{ASC1}$ Some people [$_{ASC2}$ when they **are** tired] **relax**], [$_{ASC3}$ **sleeping**], [$_{ASC4}$ **reading$_{TRAN_S}$** [a book]], [$_{ASC5}$ or **reading$_{TRAN_S}$** a television] (ESL-WR)

(56) [$_{ASC1}$ We **tried$_{TRAN_S}$** [to complain]] [$_{ASC2}$ but the girl at the tickets desk **was** impolite] (ESL-WR)

(57) [$_{ASC1}$ ... policeman **came** to,] [$_{ASC2}$ and they **explained$_{TRAN_S}$** [what happened there]] (ESL-SP)

(58) With reference to the information [that you had **requested$_{TRAN_S}$**]... (ESL-WR)
theme fronted

(59) [$_{ASC1}$ I particularly **liked$_{TRAN_S}$** [$_{ASC2}$ **seeing$_{TRAN_S}$** [all this (those) people]]], [$_{ASC3}$ and I **meet$_{TRAN_S}$** [a lot of new friend(s)] there] (ESL-WR)

Notes The following is a note for later discussions.

- A discussion is needed to clarify the semantic role labels here as well (especially 'pivot', see Semantic frame 2). A reviewer commented on that the term is closely related to syntax rather than semantics: a pivot is the structural position that a particular construction targets. In English, the pivot (in this sense) is basically always the subject. (HS: Need to check the semantic role label guidelines!)

4.5 Ditransitive

Description The ditransitive (DITRAN) ASC typically involves three core arguments: a giver (typically a subject of the clause that initiates the transfer), a receiver (the entity that accepts or is intended to accept the theme), and the theme (the entity being transferred). This construction delineates a process in which there's a direct or figurative transfer between involved arguments, including the literal transfer of giving an object, the transfer of topics in conversations.

- Literal transfers: In the sentence "She gave John a book", "She" is the giver initiating the transfer, "John" is the receiver, and "a book" is the theme being transferred.
- Metaphorical transfers: In the phrase "She gave him a fright", "She" is metaphorically transferring an emotion, a fright, to "him", making her the giver and "him" the receiver.
- Communicative transfers: An instance can be found in "She told him a secret", where "She" is the giver of information, "a secret" is the theme, and "him" is the receiver of that information.

Syntactic frame nsubj-root-iobj-obj; nsubj-root-obj-ccomp

Semantic frame 1 X_{agent} VERB $Z_{recipient}$ Y_{theme}

(60) [$_{ASC1}$ Murillo **twisted** my arm] [$_{ASC2}$ until I_{agent} **gave$_{\textbf{DITRAN}}$** him$_{recipient}$ [the address]$_{theme}$] (EWT)

(61) Can you$_{agent}$ **provide$_{\textbf{DITRAN}}$** me$_{recipient}$ **with$_{\textbf{particle}}$** [more details]$_{theme}$]? (EWT)
V-particle

(62) [$_{ASC1}$ I've **had** cats for 35 years] [$_{ASC2}$ and I_{agent} only **feed$_{\textbf{DITRAN}}$** them$_{recipient}$ [dry food]$_{theme}$] (EWT)

Semantic frame 2 X_{agent} VERB $Z_{recipient}$ Y_{topic}

(63) [$_{ASC1}$ **Let** me$_{agent}$ [$_{ASC2}$ **tell$_{\textbf{DITRAN}}$** you$_{recipient}$ [$_{ASC3}$ why it **works**]]$_{topic}$] (EWT)

(64) [$_{ASC1}$ I_{agent} **asked$_{\textbf{DITRAN}}$** him$_{recipient}$ [$_{ASC2}$ [$_{ASC3}$ what he **believed**] would **be** a solution]]$_{topic}$ (EWT)

(65) [$_{ASC1}$ You$_{agent}$ **informed$_{\textbf{DITRAN}}$** me$_{recipient}$ [$_{ASC2}$ that you would **discuss** this with your attorney]$_{topic}$] (EWT)

(66) [$_{ASC1}$ Winter$_{agent}$ **consulted$_{\textbf{DITRAN}}$** him$_{co\text{-}agent}$ **about$_{\textbf{particle}}$** [$_{ASC2}$ **eliminating** the price cap]$_{topic}$] (EWT)
V-particle

Further Cases

(67) [$_{ASC1}$ I **was** a little bit like childish], [$_{ASC2}$ but he **taught$_{\textbf{DITRAN}}$** me lots of things] (ESL-SP)

(68) [$_{ASC1}$ **Is** there a possibility [$_{ASC2}$ that you send **send$_{\textbf{DITRAN}}$** me [a list of material and clothes] [$_{ASC3}$ that I have to **take**]]]? (ESL-WR)

(69) [$_{ASC1}$ I **suggest$_{\textbf{DITRAN}}$** you [$_{ASC2}$ to **listen** to it]] [$_{ASC3}$ because it **is** very interesting] (ESL-WR)

(70) So please **sell$_{\textbf{DITRAN}}$** me [your ticket] (ESL-SP)

(71) [$_{ASC1}$ I am writing [$_{ASC2}$ to **reply** [$_{ASC3}$ your letter you **wrote$_{\textbf{DITRAN}}$** me on 10 June]]] (ESL-WR)
theme fronted

(72) It **cost$_{\text{DITRAN}}$** me [about two thousand yen] (ESL-SP)

4.6 Caused-motion

Description The caused-motion (CAUS_MOT) ASC describes a situation in which an agent either directly or indirectly influences another entity, known as the theme, to move or change location, with this movement usually clarified by a directional phrase. Two types of causation are inherent to this construction:

- Direct causation: It refers to scenarios where the agent's action immediately instigates the theme's motion, as seen in "She threw the ball into the room", where the direct action of "throwing" causes the ball's movement.
- Indirect causation: It implies that the agent's action sets off a series of events or conditions leading to the theme's movement, such as "She scared the cat into the house", where the act of scaring results in the cat's relocation.

Syntactic frame nsubj-root-obj-obl; nsubj-root-obj-advmod

Semantic frame 1 X$_{\text{agent}}$ VERB Y$_{\text{theme}}$ Z$_{\text{destination}}$

(73) I$_{\text{agent}}$ will **put$_{\text{CAUS_MOT}}$** this$_{\text{theme}}$ [on our calenders]$_{\text{destination}}$ (EWT)

(74) I$_{\text{agent}}$ will **take$_{\text{CAUS_MOT}}$** it$_{\text{theme}}$ [to Ed]$_{\text{destination}}$ for his approval (EWT)

(75) [$_{\text{ASC1}}$ Learning **combines** activity],[$_{\text{ASC2}}$ **bringing$_{\text{CAUS_MOT}}$** [mind]$_{\text{theme}}$ [into a new space]$_{\text{destination}}$] (EWT)

(76) ...ways of [$_{\text{ASC1}}$ **creating** job] [$_{\text{ASC2}}$ and **pumping$_{\text{CAUS_MOT}}$** money$_{\text{theme}}$ [into households]$_{\text{destination}}$] (EWT)

(77) After execution, we$_{\text{agent}}$ will **forward$_{\text{CAUS_MOT}}$** [a final copy]$_{\text{theme}}$ [to you]$_{\text{destination}}$ (EWT)

Semantic frame 2 X$_{\text{agent}}$ VERB Y$_{\text{theme}}$ Z$_{\text{ARGM-DIR}}$

(78) We$_{\text{agent}}$ **took$_{\text{CAUS_MOT}}$** [our vehicle]$_{\text{theme}}$ in$_{\text{ARGM-DIR}}$ for a repair (EWT)

(79) [$_{ASC1}$ Its **like** a pizza role], [$_{ASC2}$ but they$_{agent}$ **toss$_{CAUS_MOT}$** [ham and cheese]$_{theme}$ inside$_{ARGM-DIR}$] (EWT)

(80) [$_{ASC1}$ I **called**] [$_{ASC2}$ and **got** the same runaround on hold] [$_{ASC3}$ and no one$_{agent}$ just **calls$_{CAUS_MOT}$** you$_{theme}$ back$_{ARGM-DIR}$] (EWT)

Semantic frame 3 X$_{agent}$ VERB Y$_{theme}$ Z$_{goal}$

(81) [$_{ASC1}$ The TC$_{agent}$ will **cancel** it] [$_{ASC2}$ and **allot$_{CAUS_MOT}$** [the seat]$_{theme}$ [to others]$_{goal}$] (EWT)

(82) They$_{agent}$ will **extend$_{CAUS_MOT}$** [the same courtesy]$_{theme}$ [to the arriving gray wolf]$_{goal}$ (EWT)

(83) A Spanish reformer$_{agent}$ **includes$_{CAUS_MOT}$** chess$_{theme}$ [in her writings]$_{goal}$ (EWT)

Semantic frame 4 X$_{agent}$ VERB Y$_{theme}$ Z$_{source}$

(84) [the well-known Circus]$_{agent}$ **takes$_{CAUS_MOT}$** [its name]$_{theme}$ [from an old brothel]$_{SOURCE}$ (EWT)

(85) China$_{agent}$ will **purchase$_{CAUS_MOT}$** [natural gas]$_{theme}$ [from Iran]$_{SOURCE}$ (EWT)

(86) ...you$_{agent}$ might **get$_{CAUS_MOT}$** [a more accurate answer]$_{theme}$ [from someone else]$_{source}$ (EWT)

Semantic frame 5 X$_{agent}$ VERB Y$_{patient}$ Z$_{co\text{-}patient}$

(87) They$_{agent}$ have to **add$_{CAUS_MOT}$** [pure delicious extract]$_{patient}$ [to the mix]$_{co\text{-}patient}$ (EWT)

(88) [$_{ASC1}$ I$_{agent}$ am **comparing$_{CAUS_MOT}$** it$_{patient}$ [to the big results [$_{ASC2}$ I've **stayed** at]]$_{co\text{-}patient}$ (EWT)

(89) [$_{ASC1}$ ... **asking** the person [$_{ASC2}$ to **introduce$_{CAUS_MOT}$** [my friend]$_{patient}$ [to the chef]$_{co\text{-}patient}$]] (EWT)

Further Cases

(90) It is **causing**$_{\textbf{CAUS\_MOT}}$ some changes [to me] (ESL-WR)

(91) People will be able to **locate**$_{\textbf{CAUS\_MOT}}$ us [around the world] (ESL-WR)

(92) I **dropped**$_{\textbf{CAUS\_MOT}}$ my bag [on the rail] (ESL-SP)

(93) [$_{\text{ASC1}}$ I **took** some photographies] and [$_{\text{ASC2}}$ I'll **show**$_{\textbf{CAUS\_MOT}}$ them [to you]] [$_{\text{ASC3}}$ when you **come** next week] (ESL-WR)

(94) [$_{\text{ASC1}}$ There **are** so many people in a shopping centre] [$_{\text{ASC2}}$ that you can't **move**$_{\textbf{CAUS\_MOT}}$ yourself [forwards or backwards]] (ESL-WR)

(95) [$_{\text{ASC1}}$ I'd **like** [$_{\text{ASC2}}$ to **know** [$_{\text{ASC3}}$ if you can **give**$_{\textbf{CAUS\_MOT}}$ the money [back to me]]]] or [$_{\text{ASC4}}$ I'll **write**$_{\textbf{CAUS\_MOT}}$ letters [to all the media departments]] (ESL-WR)

Notes The following is a note for later discussions.

- In relation to motion constructions, there was a (theoretical) comment that this could be an mixture of syntactic and semantic criteria (e.g., how to treat prepositional datives). For example, the existence of a PP (or other directional adjunct) distinguishes transitive from caused-motion. But for intransitive motion, the semantic content of the adjunct was important (HS: I might not entirely agree with this comment because the existence of a PP was important on tagging intransitive motion as well; see the Notes on INTRAN_MOT (p.11). I am not also clear with the comment on 'the semantic content of the adjunct was important' in intransitive motion constructions.).

4.7 Transitive resultative

Description The transitive resultative (TRAN_RES) ASC is another complex transitive constructions, emphasizing the transition of an entity, referred to as the theme, into a specific end state or result due to the direct or indirect influence of the agent. This construction is typified by a three-part structure: the agent, the theme, and the resulting state or condition the theme attains. For example, in the sentence "He painted the wall red", the agent "He" prompts a change in the theme "the wall", leading to the resultative state "red". Additionally, TRAN_RES constructions can incorporate verb-particle pairings where the accompanying particle carries a metaphorical representation of the theme's eventual state. For example, in "He laughed her into a good mood", where "into" conveys

the transition of her mood from a presumably neutral or negative state to a positive one due to his laughter.

Syntactic frame nsubj-root-obj-xcomp; nsubj-root-obj-advcl

Semantic frame 1 X_{agent} VERB Y_{result} Z_{result}

(96) How do you$_{agent}$ **make$_{\textbf{TRAN_RES}}$** [your cat]$_{result}$ [$_{ASC2}$ **adjust** to a new house]$_{result}$? (EWT)

(97) [$_{ASC1}$ BTW you$_{agent}$ don't **want$_{\textbf{TRAN_RES}}$** [your leg]$_{result}$ [locked in place]$_{result}$,][$_{ASC2}$ it will **move** with the horse a bit] (EWT)

Semantic frame 2 X_{agent} VERB Y_{theme} Z_{theme}

(98) [$_{ASC1}$ if you **drink** the soda from the fridge in your room], [$_{ASC2}$ you must **prove** it [$_{ASC3}$ by **leaving$_{\textbf{TRAN_RES}}$** [the can]$_{theme}$ [in the trash]$_{theme}$]] (EWT)

(99) [$_{ASC1}$ It **is** never ok [$_{ASC2}$ to **let$_{\textbf{TRAN_RES}}$** [a customer]$_{theme}$ [$_{ASC3}$ **walk** out unhappy]$_{theme}$]] (EWT)

(100) [$_{ASC1}$ [$_{ASC2}$ All he$_{agent}$ had to **do** for a complete victory] **was** [$_{ASC3}$ **allow$_{\textbf{TRAN_RES}}$** [the military]$_{theme}$ [$_{ASC3}$ **go** into Laos and Cambodia]$_{theme}$]] (EWT)

(101) [$_{ASC1}$ I$_{agent}$ don't **find$_{\textbf{TRAN_RES}}$** [the news [$_{ASC2}$ that the groups **penetrate** the government]]$_{theme}$ [very hopeful]$_{theme}$]] (EWT)

Semantic frame 3 X_{agent} VERB Y_{theme} $Z_{predicate}$

(102) [The team at Bradley Chevron]$_{agent}$ **kept$_{\textbf{TRAN_RES}}$** [my car]$_{theme}$ [running for well past its expected death]$_{predicate}$]] (EWT)

(103) [Secretive efforts]$_{agent}$ **keep$_{\textbf{TRAN_RES}}$** outsiders$_{theme}$ [in the dark about the program's goal]$_{predicate}$]] (EWT)

Semantic frame 4 X_{agent} VERB $Y_{[patient]}$ $Z_{patient}$

(104) [$_{ASC1}$ How can I$_{agent}$ **get$_{TRAN_RES}$** [my kitten]$_{patient}$ [$_{ASC2}$ to **stop** biting]$_{predicate}$] [$_{ASC3}$ when he **plays**]? (EWT)

(105) [$_{ASC1}$ We$_{agent}$ **hauled** the horse to Windsor][$_{ASC2}$ and **got$_{TRAN_RES}$** him$_{patient}$ [treated there]$_{predicate}$] (EWT)

Further Cases

(106) [$_{ASC1}$ It **caused$_{TRAN_RES}$** me [missed (to miss) the train]] [$_{ASC2}$ and I **went** to the hotel nearly in the mid-night] (ESL-WR)

(107) [$_{ASC1}$ It could **make$_{TRAN_RES}$** things worse that the visit to your theatre restaurant,] [$_{ASC2}$ which had already **been** closed] (ESL-WR)

Notes The following is a note for later discussions.

- A reviewer raised a concern regarding the ambiguity in identifying whether certain arguments qualify as 'results' (see the example (98)). In this case, an in-PP was classified as a "result", but it rather represents a static location rather than a result. Specifically, they observed that the "can" is already in the trash prior to the action of "leaving", implying that the action does not alter the location of the "can". This interpretation suggests that the action of the agent ("you") does not impact the can's placement. (HS: The previous classification of "leave the can in the trash" as a result was because that inaction — choosing not to move the 'can' — aligns with resultative constructions. This approach suggests that the can's continued presence in the trash, maintained through deliberate non-action; similar discussion in the section 5.2.2; example about the verb 'keep').

4.8 Attributive

Description The attributive (ATTR) ASC delineates a relationship between a theme and its respective attribute. In this structure, the attribute typically consists of a noun, an adjective, or a prepositional phrase, serving to further describe the theme it is associated with. Consider the example "The sky is blue": "The sky" stands as the theme, while "blue" serves as its defining attribute, depicting the sky's hue. The copular verb (e.g., *be*) is most frequently used in this construction.

Syntactic frame nsubj-cop-root; nsubj-root-xcomp[6]

[6]This applies when the verb is not a *be* verb.

Semantic frame 1 X_{theme} VERB $Y_{attribute}$

(108) $[_{ASC1}$ [One of them]$_{theme}$ **was$_{ATTR}$** [from the Jubur tribe]$_{attribute}$] $[_{ASC2}$ and **was$_{ATTR}$** [deputy commander of the Hawijah garrison]$_{attribute}$] (EWT)

(109) $[_{ASC1}$ Fedayeen$_{theme}$ **are$_{ATTR}$** [now visible on the street]$_{attribute}$ $[_{ASC2}$ and they have become bolder than ever (EWT)

(110) $[_{ASC1}$ This time we$_{theme}$ will **be$_{ATTR}$** [the first $[_{ASC2}$ to loot]$_{attribute}$] (EWT)

(111) $[_{ASC1}$ It would **be$_{ATTR}$** dangerous$_{attribute}$ $[_{ASC2}$ to **take** one's eyes off the former ball]$_{theme}$] (EWT)

Semantic frame 2 $X_{attribute}$ VERB $Y_{attribute}$

(112) $[_{ASC1}$ [The wolf's absence]$_{stimulus}$ **seemed$_{ATTR}$** $[_{ASC2}$ to **rehabilitate** its image]$_{attribute}$] (EWT)

(113) [The situation]$_{stimulus}$ **appears$_{ATTR}$** $[_{ASC2}$ to **be** even worse in Missisippi]$_{attribute}$] (EWT)

Further cases

(114) Sorry that [this response]$_{stimulus}$ **looks$_{ATTR}$** [so long] (EWT)

(115) $[_{ASC1}$ I **thought** $[_{ASC2}$ **holding** his mouth thing$_{stimulus}$ **sounded$_{ATTR}$** wrong] (EWT)

(116) $[_{ASC1}$ [Their store]$_{stimulus}$ **[feels like]$_{ATTR}$** $[_{ASC2}$ you're **stepping** into the 1970s]] (EWT)

(117) [The coffee]$_{stimulus}$ **tastes$_{ATTR}$** [burnt and very bitter] (EWT)

4.9 Passive

Description The passive (PASSIVE) ASC emphasizes the recipient of an action or the action itself, rather than the doer (i.e., a change in voice). This allows for constructions in previous active voice ASC categories, which have an agent and a theme arguments (i.e., TRAN_S, DITRAN, CAUS_MOT, TRAN_RES), mirrors their passive counterparts. For example:

- "The book was read" is a passive counterpart to "I read the book" (TRAN_S).
- "The student was given a book" is a passive counterpart to "I gave the student a book" (DITRAN).
- "The cart was pushed to the wall" is a passive counterpart to "I pushed the cart to the wall" (CAUS_MOT).
- "The wall was painted yellow" is a passive counterpart to "I painted the wall yellow" (TRAN_RES).

In our current categorization, we have grouped them all under a single PASSIVE ASC tag.

Note that the passive construction wasn't initially incorporated into the clause-level constructions that encapsulated scenes fundamental to human experience in the foundational studies of construction grammar (Goldberg, 1995). Instead, it was perceived more in line with pragmatic functions similar to the structures like clefts, questions, and topicalizations, which adds distinct layers to the ASCs altering dynamics of agency and focus within sentences (Goldberg, 2013).

When extracting sentences for passive constructions, we place greater emphasis on the syntactic structure (aux-$V_{PASSIVE}$) along with the semantic context.

(118) [These actors] were **recommended$_{\textbf{PASSIVE}}$** [by my friend] (ESL-WR)

(119) [$_{ASC1}$ It **made** things worse that [our visit to your theatre restaurant], [$_{ASC2}$ which had already been **closed$_{\textbf{PASSIVE}}$** because of lack of customers], was **cancelled$_{\textbf{PASSIVE}}$**] (ESL-WR)

(120) [$_{ASC1}$ there **are** few famous people] [$_{ASC2}$ who desire to be **watched$_{\textbf{PASSIVE}}$** [by media]] (ESL-WR)

(121) [$_{ASC1}$ Also we should **be** careful] [$_{ASC2}$ as we could be **watched$_{\textbf{PASSIVE}}$** [by security camera]] [$_{ASC3}$ which has been **combined$_{\textbf{PASSIVE}}$** with modern technology] (ESL-WR)

(122) I was **born$_{\textbf{PASSIVE}}$** in January (ESL-SP)

5 Annotation Challenges and Discussions

When categorizing ASCs, similar to other annotation projects, decision-making is important (Gerdes and Kahane, 2016). The previous section (§4) addressed the 'what choices' and 'why those choices' made by the annotators in constructing the ASC treebank. However, the process often encountered challenges due to inherent ambiguity of language, which sometimes lacks distinct boundaries for grammatical categorization. In this section, we delve into the difficulties faced during the annotation of ASC tags in our datasets. These challenges were summarized from asynchronous discussions among the annotators.

5.1 Ambiguous constructions

5.1.1 Caused-motion vs. Transitive resultative

Distinguishing between the caused-motion (CAUS_MOT) ASC and the transitive resultative (TRAN_RES) ASC can pose challenges, particularly from a syntactic perspective. Both constructions necessitate the presence of three arguments: the agent, the theme, and the change of location/state. Annotators often questioned this distinction, especially when the change of state is manifested through a prepositional phrase. To simplify and guide the annotation process, we adhered to the semantic characteristics of the two constructions and established the following guidelines:

- If the primary focus is on a change of location of the theme, the CAUS_MOT label should be applied (e.g., *Sam pushed the cart into the room*).
- Conversely, when a prepositional phrase is in play and it denotes a change of state of the theme, the TRAN_RES label is more fitting (e.g., *She froze the water into ice*).

We also guided the annotators with more exemplary studies from the previous studies (Boas, 2002; Goldberg, 1995; Talmy, 1985).

- **TRAN_RES**
  - *Claire painted the house red.*
  - *Pat broke the vase (in)to pieces.*
  - *Chris drank himself silly.*
  - *Carol danced Jim tired.*
  - *Rachel swept the floor clean.*
  - *He wiped the table clean.*
  - *He talked himself blue in the face.*
  - *She coughed herself sick.*
  - *She slept herself sober.*
  - *He made the metal safe.*

- *He made her a queen.*

- **CAUS_MOT**
  - *Frank pushed it into the box.*
  - *Frank kicked the dog into the bathroom.*
  - *Frank sneezed the tissue off the nightstand.*
  - *Sam shoved it into the carton.*
  - *Pat crammed the pennies into the jar.*
  - **X COMMUNICATE Y to MOVE Z:** Sometimes the construction can be related to force-dynamic verbs (Talmy, 1985) that encode a communicative verb. In this case, the motion is not strictly entailed.
    - *Sam ordered him out of the house.*
    - *Sam asked him into the room.*
    - *Sam invited him out to her cabin.*
    - *Sam beckoned him into the room.*
    - *Sam urged him into the room.*
    - *Sam sent him to the market.*
  - **X ENABLES Y to MOVE Z:** This class includes force-dynamic verbs that encode the removal of a barrier.
    - *Sam allowed Bob out of the room.*
    - *Sam let Bill into the room.*
    - *Sam freed the bird out of the cage.*
  - **X PREVENTS Y from MOVING Z:** This class imposes a barrier, causing the patient to stay in a location.
    - *Harry locked Joe into the bathroom.*
    - *He kept her at arm's length.*
    - *Sam barricaded him out of the room.*
  - **X HELPS Y to MOVE Z:** This case involves ongoing assistance to move in a certain direction.
    - *Sam helped him into a car.*
    - *Sam assisted her out of the room.*
    - *Sam guided him through the terrain.*
    - *Sam showed him into the living room.*

* *Sam walked him to the car.*

5.1.2 Ditransitive vs. Caused-motion

One recurring question concerned cases involving dative alternation. This alternation occurs between a prepositional indirect-object construction (e.g., *I gave the book to him*) and a double-object construction (e.g., *I gave him the book*). To distinguish between the ditransitive (DITRAN) ASC and the caused-motion (CAUS$_M$OT) ASC, one should assess based on the syntactic frame. This means that DITRAN features an indirect object preceding the direct object, whereas CAUS$_M$OT typically presents the direct object first, followed by the indirect object, often accompanied by a preposition.

Dative Alternation and its Ambiguity

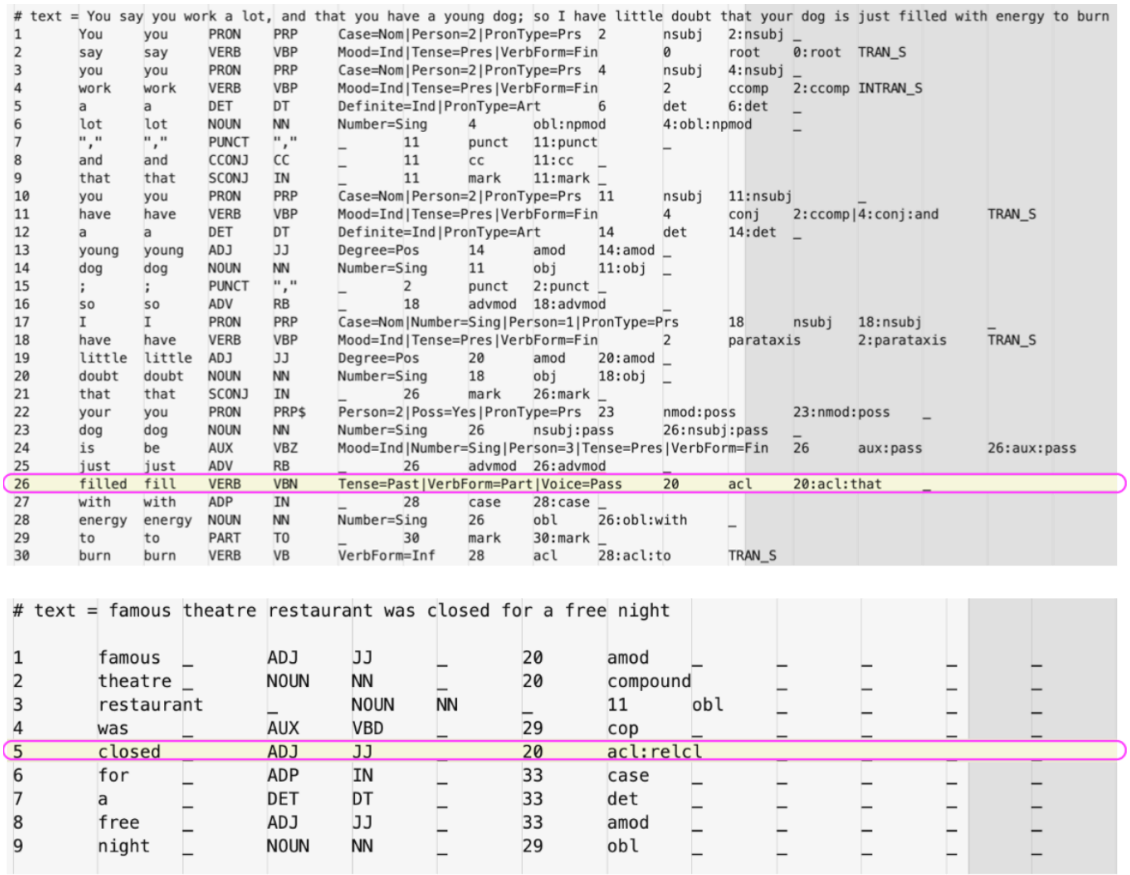
- The object of the *to* preposition encompasses a broad spectrum of semantic argument types. Examples include:
  - Recipients or possessonal goals: e.g., *She handed a letter to me* (Haspelmath, 2003; Hovav & Levin, 2008, Newman, 1996).
  - Spatial goals: e.g., *He kicked a ball to the target.*
- Occasionally, this can also embrace arguments that are not discernibly goals. This is observed in instances like *conform to* or *submit to*, which are more likely to be phrasal verbs.
- Within the scope of this project:
  - Every prepositional indirect-object construction is included when the argument succeeding the *to* preposition qualifies as either a recipient (possessional goal) or a spatial goal.
  - Such constructions will be tagged as **CAUS_MOT**.

5.1.3 Attributive vs. Passive

In distinguishing between the attributive and passive constructions, annotators frequently face challenges when attempting to differentiate them based solely on surface structure. This is particularly true when a verb exists on a spectrum spanning the past participle form and an adjective. For instance, consider the two examples below, which, despite their similarities in surface structures, differ in their underlying grammatical constructions (i.e., 4th column UPOS tag – VERB vs. ADJ) (Figure 4).

To delineate between the two, we recommend leveraging UPOS tags. A simple guideline is proposed:

- Should the element following the 'be' verb be labeled as VERB (4th column), then the 'verb' should be tagged as PASSIVE. Also, we mark the preceding *be* verb in this case as `aux:pass` (8th column) (example above).
- Conversely, if this element is tagged as ADJ, then the 'be' verb should be classified under ATTR. Also, the preceding *be* verb in this case is marked as `cop` (example below).



```
# text = You say you work a lot, and that you have a young dog; so I have little doubt that your dog is just filled with energy to burn
1       You     you     PRON    PRP     Case=Nom|Person=2|PronType=Prs  2       nsubj   2:nsubj _
2       say     say     VERB    VBP     Mood=Ind|Tense=Pres|VerbForm=Fin        0       root    0:root  TRAN_S
3       you     you     PRON    PRP     Case=Nom|Person=2|PronType=Prs  4       nsubj   4:nsubj _
4       work    work    VERB    VBP     Mood=Ind|Tense=Pres|VerbForm=Fin        2       ccomp   2:ccomp INTRAN_S
5       a       a       DET     DT      Definite=Ind|PronType=Art       6       det     6:det   _
6       lot     lot     NOUN    NN      Number=Sing     4       obl:npmod       4:obl:npmod     _
7       ","     ","     PUNCT   ","     _       11      punct   11:punct        _
8       and     and     CCONJ   CC      _       11      cc      11:cc   _
9       that    that    SCONJ   IN      _       11      mark    11:mark _
10      you     you     PRON    PRP     Case=Nom|Person=2|PronType=Prs  11      nsubj   11:nsubj        _
11      have    have    VERB    VBP     Mood=Ind|Tense=Pres|VerbForm=Fin        4       conj    2:ccomp|4:conj:and      TRAN_S
12      a       a       DET     DT      Definite=Ind|PronType=Art       14      det     14:det  _
13      young   young   ADJ     JJ      Degree=Pos      14      amod    14:amod _
14      dog     dog     NOUN    NN      Number=Sing     11      obj     11:obj  _
15      ;       ;       PUNCT   ";"     _       2       punct   2:punct _
16      so      so      ADV     RB      _       18      advmod  18:advmod       _
17      I       I       PRON    PRP     Case=Nom|Number=Sing|Person=1|PronType=Prs      18      nsubj   18:nsubj        _
18      have    have    VERB    VBP     Mood=Ind|Tense=Pres|VerbForm=Fin        2       parataxis       2:parataxis     TRAN_S
19      little  little  ADJ     JJ      Degree=Pos      20      amod    20:amod _
20      doubt   doubt   NOUN    NN      Number=Sing     18      obj     18:obj  _
21      that    that    SCONJ   IN      _       26      mark    26:mark _
22      your    you     PRON    PRP$    Person=2|Poss=Yes|PronType=Prs  23      nmod:poss       23:nmod:poss    _
23      dog     dog     NOUN    NN      Number=Sing     26      nsubj:pass      26:nsubj:pass   _
24      is      be      AUX     VBZ     Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin   26      aux:pass        26:aux:pass
25      just    just    ADV     RB      _       26      advmod  26:advmod       _
26      filled  fill    VERB    VBN     Tense=Past|VerbForm=Part|Voice=Pass     20      acl     20:acl:that     _
27      with    with    ADP     IN      _       28      case    28:case _
28      energy  energy  NOUN    NN      Number=Sing     26      obl     26:obl:with     _
29      to      to      PART    TO      _       30      mark    30:mark _
30      burn    burn    VERB    VB      VerbForm=Inf    28      acl     28:acl:to       TRAN_S
```

```
# text = famous theatre restaurant was closed for a free night

1       famous  _       ADJ     JJ      _       20      amod    _       _       _       _       _
2       theatre _       NOUN    NN      _       20      compound        _       _       _       _       _
3       restaurant      _       NOUN    NN      _       11      obl     _       _       _       _       _
4       was     _       AUX     VBD     _       29      cop     _       _       _       _       _
5       closed  _       ADJ     JJ      _       20      acl:relcl       _       _       _       _       _
6       for     _       ADP     IN      _       33      case    _       _       _       _       _
7       a       _       DET     DT      _       33      det     _       _       _       _       _
8       free    _       ADJ     JJ      _       33      amod    _       _       _       _       _
9       night   _       NOUN    NN      _       29      obl     _       _       _       _       _
```

Figure 4: Examples comparing PASSIVE (top) vs. texttttATTR (bottom)

5.2 Ambiguous ASC combinations

5.2.1 *enter*-type verbs

When *enter*-type verbs appear in simple transitive constructions, they are assigned the **INTRAN_MOT** tag. This is because they typically present a theme in the subject position followed by a destination. Examples of this include:

- ***enter*** *the classroom*
- ***reach*** *the summit*
- ***approach*** *the city*
- ***visit*** *the museum*

5.2.2 *keep/prevent*-type verbs

Verbs such as *keep* and *prevent*, despite not overtly evoking any alteration in direction or state, are categorized under caused-motion or transitive resultative constructions. This categorization stems from the interpretation that causing an object to maintain a steady position (as in **CAUS_MOT**) or to retain its

current state (as in **TRAN_RES**) is a form of induced change. Illustrative examples include:

- ***keep*** *someone happy*
- ***keep*** *her staying at their house*
- ***keep*** *the body youthful*
- ***prevent*** *someone from feeling sad*
- ***prevent*** *her from leaving their house*
- ***prevent*** *the body from aging*

5.2.3 Reflexive object

In the context of the transitive simple construction, reflexive pronouns play a unique role. These pronouns, such as "myself", "yourself", and "ourselves", are employed when an entity acts upon itself. Though they primarily serve to establish co-reference with the subject, their use can sometimes introduce a degree of semantic ambiguity. Despite this potential confusion, we still opted to assign the TRAN_S tag to those cases (Figure 5).

5.2.4 Figurative particle

In English, particles play an important role, particularly those that primarily denote motion and result. Such particles encompass forms like: *across*, *along*, *around*, *aside*, *away*, *by*, *down*, *forth*, *home*, *in*, *off*, *on*, *out*, *over*, *past*, *round*, *through*, *under*, and *up* (Biber et al., 1999). When these particles convey a resultant state of the object caused by the agent's action, they are tagged as TRAN_RES. Examples illustrating this include[7]:

- *he called his friend* ***up***
- *if you could shoot this* ***off*** *over night*
- *blow things* ***up***
- *blast something* ***away***
- *set something* ***off***

[7] In relation to the discussion here, we plan to incorporate some insights regarding 'lexical items'. A reviewer made an observation stating, "If you set off an alarm, there's no 'away from the deictic center' meaning to 'off', and it certainly does not reflect the alarm's resultant state—contrarily, the alarm is activated, not deactivated. I believe 'set off' is merely a lexicalized phrasal verb." It would be important guideline for annotators on how to approach particles—specifically, on distinguishing between "real" directional particles that convey a resultant state and particles that are integral to a complex verb.

```
# text = The life at home will be more easy because a refrigerator which fill up itself or a bathroom which clean itself .
1       The     _       DET     DT      _       2       det     _       _
2       life    _       NOUN    NN      _       8       nsubj   _       _
3       at      _       ADP     IN      _       4       case    _       _
4       home    _       NOUN    NN      _       2       nmod    _       _
5       will    _       AUX     MD      _       8       aux     _       _
6       be      _       AUX     VB      _       8       cop     _       ATTR
7       more    _       ADV     RBR     _       8       advmod  _       _
8       easy    _       ADJ     JJ      _       0       root    _       _
9       because _       SCONJ   IN      _       11      case    _       _
10      a       _       DET     DT      _       11      det     _       _
11      refrigerator    _       NOUN    NN      _       8       obl     _       _
12      which   _       DET     WDT     _       13      nsubj   _       _
13      fill    _       VERB    VB      _       11      acl:relcl       _       TRAN_S
14      up      _       ADP     RB      _       13      compound:prt    _       _
15      itself  _       PRON    PRP     _       13      obj     _       _
16      or      _       CCONJ   IN      _       18      cc      _       _
17      a       _       DET     DT      _       18      det     _       _
18      bathroom        _       NOUN    NN      _       11      conj    _       _
19      which   _       DET     WDT     _       20      nsubj   _       _
20      clean   _       VERB    VB      _       18      acl:relcl       _       TRAN_S
21      itself  _       PRON    PRP     _       20      obj     _       _
22      .       _       PUNCT   .       _       8       punct   _       _
```

Figure 5: Example of reflexive objects

5.3 Difficult grammatical structures

5.3.1 Inverted/Fronted word order

At this stage, we refrained from distinguishing between ASCs that exhibit varying argument orders, primarily attributing this decision to pragmatic considerations[8]. Different word orders can arise as a result of various pragmatic situations, as they often serve specific communication functions like emphasis, topicalization, and information structure. While we acknowledge that 'function' may significantly impact grammatical constructions (e.g., Goldberg, 2006), we believe that this aspect merits more in-depth exploration in future research.

- *I am writing to reply **your letter you wrote me** on 10 June* – DITRAN
- *You say you work a lot, and that you have a young dog; so I have little doubt that your dog is just filled with **energy to burn*** – TRAN_S
- ***"The Prophet's guidance," says Michael Scheuer, an al-Qaeda analyst** who recently retired from the CIA and once headed its Bin Laden unit, "was always, Before you attack someone, warn them very [...]* – TRAN_S

5.3.2 Dealing with L2 usage

Sentences/utterances in L2 datasets sometimes contain syntactical errors, yet remain semantically somewhat interpretable. Drawing from the two preceding projects related to the L2 English UD treebank (Berzak et al., 2016; Kyle et al., 2022), we aim to annotate the ASCs based on their realized syntactic form rather than their intended meaning. This approach was to avoid any subjective interpretation by the annotators and to ensure annotation consistency. Below is an illustrated example (Figure 6).

The issue lies with the use of the preposition "about" in this context, which is not a typical usage

[8] In this context, pragmatics refers to how communicative context influences form and meaning.

```
# text = I explained about the total stories ?
1       I           _       PRON    PRP     _       2       nsubj   _       _
2       explained   _       VERB    VBD     _       0       root    _       INTRAN_S
3       about       _       ADP     IN      _       6       case    _       _
4       the         _       DET     DT      _       6       det     _       _
5       total       _       ADJ     JJ      _       6       amod    _       _
6       stories     _       NOUN    NNS     _       2       obl     _       _
7       ?           _       PUNCT   .       _       2       punct   _       _
```

Figure 6: Example of annotating an ungrammatical sentence

of the verb *explain* (e.g., *the teacher explained this to me*). By tagging this deviation with INTRAN_S, we aim to distinguish the varied usages of the verb *explain*. We anticipate that the verb will more likely be used in a caused-motion construction (as in the provided example) by advanced English speakers.

5.3.3 Multi-word expressions

Previous studies In the earlier PropBank project, multi-word expressions (MWEs) (including idiomatic expressions) received a focus regarding annotation methodology and approach (Bonial et al., 2014). In addition, in several syntactic annotation projects, they have been discussed in different approaches like (1) marking the extension of MWEs on the top of syntactic annotation or (2) distinguishing syntactically irregular structures (e.g., *heaven knows who, by and large*) from semantically non-compositional units (but syntactically transparent structures) (e.g., *kick the bucket* and suggesting different annotation strategies for these two MWE types (see Savary et al., 2017).

In the current phase of our project, our guiding principle is rooted in the belief that the meaning of MWEs arises from the collective semantics of its constituent words, rather than their individual meanings (Biber et al., 1999, p.58). Consequently, once we discern a unified meaning for the MWE, we assign it a singular ASC tag based on that semantic understanding.

Types We have categorized MWEs into four types, each of which "comprises relatively idiomatic units and functions as a single verb" (Biber et al., 1999, p.403). These includes:

- verb + adverbial particle (e.g., *pick up*)
- verb + preposition (e.g., *look at*)
- verb + prepositional phrase (e.g., *take into account*)
- verb + particle + preposition (e.g., *get away with*)
- verb + noun phrase (e.g., *take a look (at)* (related to the discussion about **light verb construction** below)
- verb + verb (e.g., *go get*)

Light verb constructions Light verb constructions are considered as a subset of MWEs. As illustrated in Hwang et al. (2010), previous linguistic studies offer varied interpretations of these structure. In this project, we follow Goldberg's viewpoint (1995) that taking them as a single lexical item. For example, in

sentence like "Henry took a walk", the agent ("Henry") isn't exerting any influence on an object (syntactically, an object in this sentence is "a walk"). Instead, "take a walk" conveys the same event as the verb "walk". Consequently, we assign the INTRAN_S tag. This approach is also evident in examples from our datasets:

- *hitch a ride back – ride back*
- *have a conference call – call*
- *have/take a look – look*

5.4 Skipped items

5.4.1 Modals and semi-modal verbs

In this project, modal auxiliary verbs are not tagged. These verbs primarily fulfill grammatical functions. The following are examples (Biber et al., 1999):

- Nine central modal auxiliary verbs for modality: *can, could, may, might, shall, should, will, would, must*
- Marginal auxiliary verbs (which behave like modals in taking auxiliary negation and yes-no question inversion): *need (to), ought to, dare (to), used to*
- Semi-modals (i.e., a number of fixed idiomatic phrases with functions similar to those of modals): *(had) better, have to, (have) got to, be supposed to, be going to*
- Some of the verb phrases with the force of a hedge (in certain contexts): *happen to, tend to be, appear to, (there) is said to be*

5.4.2 Some transitional phrases

We have excluded certain transitional phrases that indicate transitions between ideas within discourse[9]. Below are the examples we omitted during the annotations: *according to*, *given that*, *regarding*.

[9] While discourse markers can fall under our definition of MWEs, we have separated them into this section specifically because they were excluded from annotations.

References

Akbik, A., Chiticariu, L., Danilevsky, M., Li, Y., Vaithyanathan, S., and Zhu, H. (2015). Generating high quality proposition banks for multilingual semantic role labeling. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 397–407.

Bencini, G. M. and Goldberg, A. E. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language*, 43(4):640–651.

Bergen, B. K. and Chang, N. (2005). Embodied construction grammar in simulation-based language understanding. *Construction Grammars: Cognitive Grounding and Theoretical Extensions*.

Berzak, Y., Kenney, J., Spadine, C., Wang, J. X., Lam, L., Mori, K. S., Garza, S., and Katz, B. (2016). Universal dependencies for learner english. *arXiv preprint arXiv:1605.04278*.

Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). Longman grammar of spoken and written english.

Bies, A., Mott, J., Warner, C., and Kulick, S. (2012). English web treebank. *Linguistic Data Consortium, Philadelphia, PA*.

Boas, H. C. (2002). On constructional polysemy and verbal polysemy in construction grammar. In *Proceedings of the 2000 Western Conference on Linguistics*, volume 12.

Bonial, C., Green, M., Preciado, J., and Palmer, M. (2014). An approach to take multi-word expressions. In *Proceedings of the 10th Workshop on Multiword Expressions (MWE)*, pages 94–98.

Bowerman, M. (1996). The origins of children's spatial semantic categories: Cognitive vs. linguistic determinants. In *Rethinking linguistic relativity*, pages 145–176. Cambridge University Press.

Cameron-Faulkner, T., Lieven, E., and Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6):843–873.

Cappelle, B. (2006). Particle placement and the case for'allostructions'. *Constructions*.

Clark, E. V. (1978). Awareness of language: Some evidence from what children say and do. In *The child's conception of language*, pages 17–43. Springer.

De Marneffe, M.-C., Manning, C. D., Nivre, J., and Zeman, D. (2021). Universal dependencies. *Computational Linguistics*, 47(2):255–308.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2):143–188.

Ellis, N. C. and Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, 93(3):370–385.

Ellis, N. C. and Larsen-Freeman, D. (2009). Constructing a second language: Analyses and computational simulations of the emergence of linguistic constructions from usage. *Language Learning*, 59:90–125.

Fillmore, C. J. (1975). An alternative to checklist theories of meaning. In *Annual Meeting of the Berkeley Linguistics Society*, volume 1, pages 123–131.

Fillmore, C. J., Johnson, C. R., and Petruck, M. R. (2003). Background to framenet. *International Journal of Lexicography*, 16(3):235–250.

Gerdes, K. and Kahane, S. (2016). Dependency annotation choices: Assessing theoretical and practical issues of universal dependencies. In *LAW X (2016) The 10th Linguistic Annotation Workshop: 131*.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5):219–224.

Goldberg, A. E. (2006). Constructions work.

Goldberg, A. E. (2013). Argument structure constructions versus lexical rules or derivational verb templates. *Mind & Language*, 28(4):435–465.

Goldberg, A. E. (2019). *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton University Press.

Hwang, J. D., Bhatia, A., Bonial, C., Mansouri, A., Vaidya, A., Xue, N., and Palmer, M. (2010). Propbank annotation of multilingual light verb constructions. In *Proceedings of the Fourth Linguistic Annotation Workshop*, pages 82–90.

Kyle, K. and Crossley, S. (2017). Assessing syntactic sophistication in l2 writing: A usage-based approach. *Language Testing*, 34(4):513–535.

Kyle, K., Eguchi, M., Miller, A., and Sither, T. (2022). A dependency treebank of spoken second language english. In *Proceedings of the 17th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2022)*, pages 39–45.

Kyle, K. and Sung, H. (2023). An argument structure construction treebank. *Proceedings of the First International Workshop on Construction Grammars and NLP (CxGs+NLP, GURT/SyntaxFest 2023)*, pages 51–62.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites*, volume 1. Stanford university press.

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, 26(3):619–653.

Nivre, J., Agić, Ž., Ahrenberg, L., Antonsen, L., Aranzabe, M. J., Asahara, M., Ateyah, L., Attia, M., Atutxa, A., Augustinus, L., et al. (2017). Universal dependencies 2.1.

Nivre, J., De Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., et al. (2016). Universal dependencies v1: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1659–1666.

O'Connor, M. C. and Kay, P. (2003). Regularity and idiomaticity in grammatical constructions: The case of let alone. In *The New Psychology of Language*, pages 243–270. Psychology Press.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106.

Rappaport Hovav, M. and Levin, B. (1998). Building verb meanings. *The Projection of Arguments: Lexical and Compositional Factors*, pages 97–134.

Savary, A., Ramisch, C., Cordeiro, S. R., Sangati, F., Vincze, V., Qasemi Zadeh, B., Candito, M., Cap, F., Giouli, V., Stoyanova, I., et al. (2017). The parseme shared task on automatic identification of verbal multiword expressions. In *The 13th Workshop on Multiword Expression at EACL*, pages 31–47.

Schuler, K. K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon.* University of Pennsylvania.

Silveira, N., Dozat, T., De Marneffe, M.-C., Bowman, S. R., Connor, M., Bauer, J., and Manning, C. D. (2014). A gold standard dependency corpus for english. In *LREC*, pages 2897–2904. Citeseer.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. *The Crosslinguistic Study of Language Acquisition*, 2:1157–249.

Sung, H. and Kyle, K. (2024). Annotation scheme for english argument structure constructions treebank. *Proceedings of the 18th Linguistic Annotation Workshop (LAW-XVIII)*.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language Typology and Syntactic Description*, 3(99):36–149.