Gold Standard Annotations for Preposition and Verb Sense with Semantic Role Labels in Adult-Child Interactions

Department of Psychology, University of Illinois at Urbana-Champaign¹, Amazon Research², Intelligent Medical Objects³, Department of Computer Science, University of Pennsylvania⁴

PROJECT

We annotated adult-child dialogues with predicate and semantic role labels. We annotated the Adam files of the Brown (1973) corpus from the CHILDES corpora. This project extends the annotations of Connor et al. 2010.



AIMS

Syntactic bootstrapping refers to the ability that young children have to learn language in spite of ambiguous scenes. Structure*mapping* refers to children's hypothesized use of simple abstract structures to facilitate the mapping of scenes to utterances (Fisher et al. 2010).

We use these annotated corpora to test our BabySRL model (Connor et al. 2010, 2011, 2013), which learned to interpret sentences based on simple representations of syntactic structure, derived from a constrained distributional analysis of child-directed speech, amplified by simple built-in expectations about predicateargument structure.

METHODS

Semantic Role Labels (SRL) are a common means of representing semantic information about scenes in the world.

Sentence: Do <u>you</u> the bus ? see Arg0-viewer see.01 Arg1-thing viewed SRLs:

Our annotators used Propbank verb senses SRLs (Palmer et al. 2005) and preposition senses and SRLs from Srikumar & Roth 2011.

If the automatic parse interfered with labeling semantic roles, annotators were instructed to change it according to the Penn Treebank Guidelines as outlined in Bies et al. (1995), and the modified guidelines in Warner et al. (2012).

Lori A Moon¹, Christos Christodouloupolos², Cynthia Fisher¹, Sandra Franco³, & Dan Roth⁴



Inter-Annotator Agreement

ter-Annota	tor Agree	ement				File
All Files		sense ĸ	# args к	label id ĸ	span ĸ	Adam 1
Verb	Adult	74.7	86.7	92.8	94.5	Adam 1
	Child	68.8	83.6	90.2	96.1	Adam 1
	Total	71.8	85.1	91.5	95.3	Adam '
Preposition	Adult	72.9	75.5	91.4	93.3	In comparis
	Child	73.7	64.0	90.0	91.6	15 of the 13
	Total	73.3	69.8	90.7	92.5	Across all
Predicates	Total	72.6	77.4	91.1	93.9	average of and label.

Jubilee Annotation Tool

The annotation tool was based on the Jubilee tool by (Choi et al., 2010) and the modified version is available at <u>https://gitlab-beta.engr.illinois.edu/babysrl-group/jubilee</u>.

Freebank View				Frameset V	iew		
Prev Next Approve auto (2	2) 6 *MOT – to [1/2]	\$	☆			Exa	ample
TOP S1				to	to ᅌ	View in	Browser
▼ SBARQ				- Roleset Int	formation —		
▼ WHNP				ID : to Des	tination		
WP what				Name : De	stination		
▼ SQ				Arg-P : go	vernor		
AUX are				Arg-P : obj	ject		
▼ NP							
PRP vou							
VP VP							
VBG doing 3:0-ARGP-GOV							
TO to 4:0 rol							
* NP							
DIthe							
NN paper 6:0-AKGP-OBJ							
. ?							
. ?							
. ?				Argument	/iew	0	
. ?				Argument	/iew	2	3
. ? . ?) *CHI: finished .				Argument V	/iew	2	3
. ? . *CHI: finished . L *URS: you finished ?				Argument V 0	/iew1	2 P-GOV (X)	3 P-OB1 (Y
. ?) *CHI: finished . L *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine				Argument V 0 4	/iew1 5	2 P-GOV (X)	3 P-OBJ (Y
. ?) *CHI: finished . L *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper				Argument V 0 4	/iew 1 5 M.CALL(C)	2 P-GOV (X)	3 P-OBJ (Y
. ?) *CHI: finished . L *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper . ## 5 *CHI: mine				Argument V 0 4 M-ADV (V)	/iew 1 5 M-CAU (C)	2 P-GOV (X) M-DIR (D)	3 P-OBJ (Y M-DIS (I)
. ?) *CHI: finished . 1 *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper . ## 5 *CHI: mine .				Argument V 0 4 M-ADV (V)	/iew 1 5 M-CAU (C)	2 P-GOV (X) M-DIR (D)	3 P-OBJ (Y M-DIS (I)
. ? . ? . ? . ? . ? . ? . ? . ?				Argument V 0 4 M-ADV (V) M-DSP (B)	/iew 1 5 M-CAU (C) M-EXT (E)	2 P-GOV (X) M-DIR (D) M-LOC (L)	3 P-OBJ (Y M-DIS (I) M-MNR (M
. ?) *CHI: finished . L *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper . ## 5 *CHI: mine . ## 7 *CHI: where my pen # Urs(u)la ? 8 *URS: who's that # Adam ?				Argument V 0 4 M-ADV (V) M-DSP (B)	/iew 1 5 M-CAU (C) M-EXT (E)	2 P-GOV (X) M-DIR (D) M-LOC (L)	3 P-OBJ (Y M-DIS (I) M-MNR (M
. ?) *CHI: finished . 1 *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper . ## 5 *CHI: mine . ## 5 *CHI: mine . ## 7 *CHI: where my pen # Urs(u)la ? 8 *URS: who's that # Adam ? ## 9 *CHI: Cromer .				Argument V 0 4 M-ADV (V) M-DSP (B) M-MOD (O)	/iew 1 5 M-CAU (C) M-EXT (E) M-NEG (N)	2 P-GOV (X) M-DIR (D) M-LOC (L) M-SLC (6)	3 P-OBJ (Y M-DIS (I) M-MNR (M M-PRD (7
. ?) *CHI: finished . L *URS: you finished ? ## 2 *URS: all set ? ## 3 *CHI: mine . ## 4 *CHI: my paper . ## 5 *CHI: mine . 5 *MOT: what are you doing to the paper ? ## 7 *CHI: where my pen # Urs(u)la ? 8 *URS: who's that # Adam ? ## 9 *CHI: Cromer . L0 *CHI: let me slip down .				Argument V 0 4 M-ADV (V) M-DSP (B) M-MOD (O)	/iew 1 5 M-CAU (C) M-EXT (E) M-NEG (N)	2 P-GOV (X) M-DIR (D) M-LOC (L) M-SLC (6)	3 P-OBJ (Y M-DIS (I) M-MNR (M M-PRD (7
. ? . ? . ? . ? . ? . ? . ? . ?				Argument V 0 4 M-ADV (V) M-DSP (B) M-MOD (O) M-PRP (P)	/iew 1 5 M-CAU (C) M-EXT (E) M-NEG (N) M-RCL (R)	2 P-GOV (X) M-DIR (D) M-LOC (L) M-SLC (6) M-REC (9)	3 P-OBJ (Y M-DIS (I) M-MNR (M M-PRD (7 M-TMP (1
. ? . ? . ? . ? . ? . ? . ? . ?				Argument V 0 4 M-ADV (V) M-DSP (B) M-MOD (O) M-PRP (P)	/iew 1 5 M-CAU (C) M-EXT (E) M-NEG (N) M-RCL (R)	2 P-GOV (X) M-DIR (D) M-LOC (L) M-SLC (6) M-REC (9)	3 P-OBJ (Y M-DIS (I) M-MNR (M M-PRD (7 M-TMP (1

Screenshot of the annotation tool with an example from the output of the NLP pipeline parse and semantic role labeler. The label 'auto', is in red, indicating that the labels are the output of the NLP pipeline. The context window at the bottom shows the surrounding discourse.

Annotation Issues

Go

go.01

go.01

The most common issues that came up in annotation involved labeling partial expressions and ambiguous arguments.

train.	
Arg1-en	tity in motion,
(means,	unlabeled, e.

Annotators were instructed to only give a semantic role if it was clear what it was from the dialogue context, limited the the context window (lower right) in Jubilee.

goer .g., "I go by train")

CORPORA

The resulting corpora are freely available. The xml version is available via the CogComp Group. https://cogcomp.org/page/resource_view/115

The CHAT (MacWhinney 2000) version is available through TalkBank at https://childes.talkbank.org/derived/

APPLICATIONS

In addition to supporting models of human language development, annotated corpora of adult-child dialogue can provide a useful context in which to evaluate the robustness of NLP learning models.

Toddlers producing their early word combinations often omit the function words that support high-accuracy part-ofspeech tagging, parsing and semantic-role assignment (e.g., Brown, 1973). Despite these omissions, however, they are often understood by their adult interlocutors. Understanding speech with missing elements requires a flexible knowledge of language.

ACKNOWLEDGEMENTS

The BabySRL project is funded by NIH grant R01-HD054448 and NSF grant BCS-060257. We would like to thank the anonymous mentor who gave us excellent feedback on our earlier draft, and three anonymous reviewers who gave helpful comments.



COLING2018 Santa Fe

	Annotator X	Annotator Y	Original	
1	1159	1133	1403	
5	909	884	1187	
17	1402	1379	1453	
9	1704	1695	1761	

parison to previous work of Connor et al. (2010), ng to http://cogcomp.org/Data/BabySRL.html e 133 files were held-out for measuring IAA. all of the files, annotators agreed on an of 96.57% of the annotated arguments for span el. Our span and label average is 93.8%.







JCILLINOIS