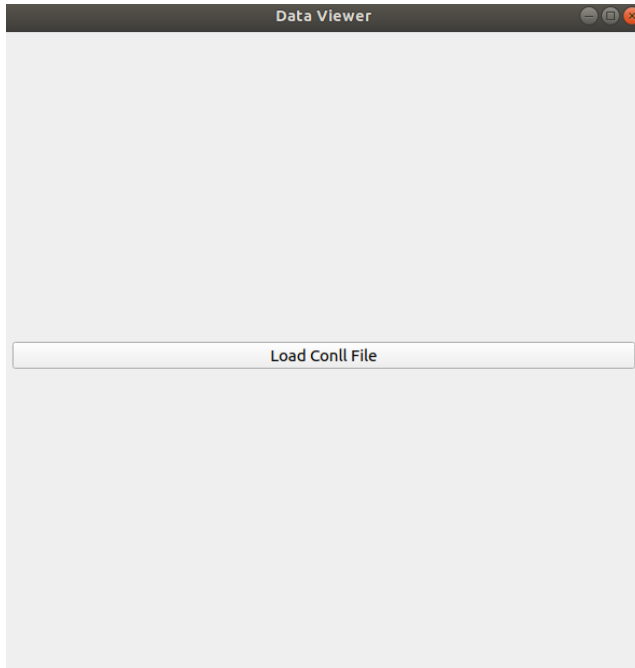


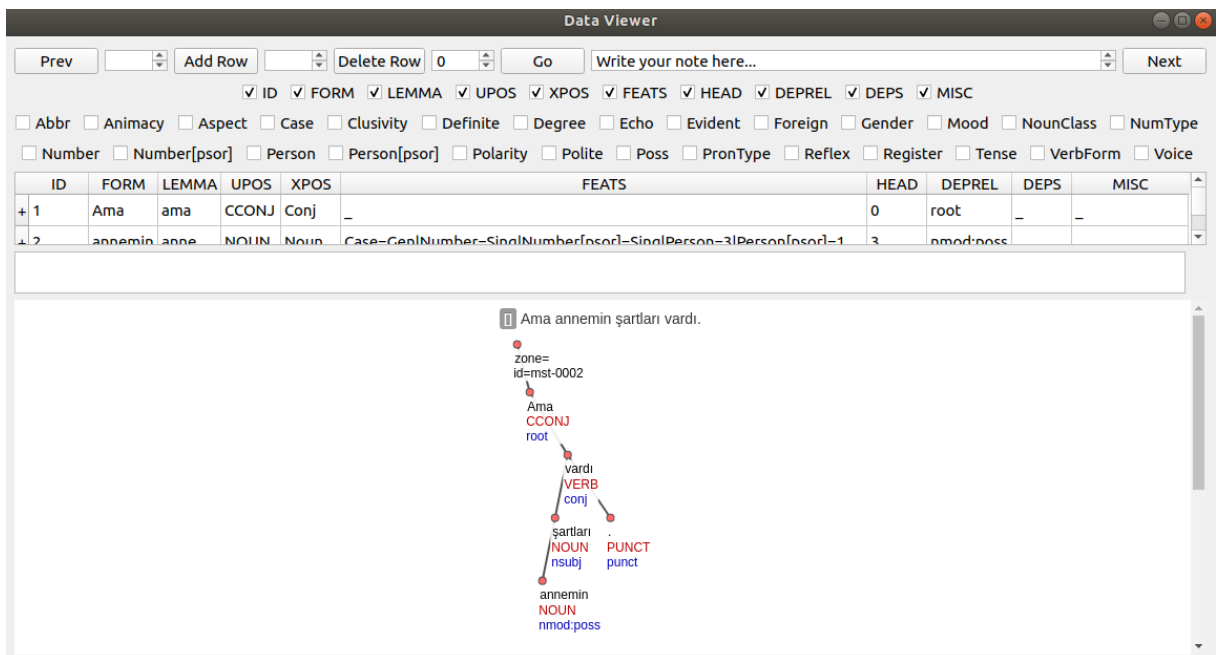
BoAT: BOUN Annotation Tool for Dependency Parsing

USER MANUAL

1. Load the Conll File by clicking the “Load Conll File” button. Then, chose your file from the file dialog and click “Open”.



2. The file is loaded successfully. Make the window full size for effective use.



- Sentences in the file is shown one by one along with the default fields (ID, FORM, LEMMA, UPOS, XPOS, FEATS, HEAD, DEPREL, DEPS and MISC) and the parse tree of the sentence. The cells in the first column of the table (written "+" or "-") are clickable and used for MWE manipulation.

The screenshot shows the Data Viewer interface. At the top, there are controls for navigating between rows (Prev, Add Row, Delete Row, Go) and a search bar. Below these are checkboxes for various fields: ID, FORM, LEMMA, UPOS, XPOS, FEATS, HEAD, DEPREL, DEPS, MISC, Abbr, Animacy, Aspect, Case, Clusivity, Definite, Degree, Echo, Evident, Foreign, Gender, Mood, NounClass, NumType, Number, Number[psor], Person, Person[psor], Polarity, Polite, Poss, PronType, Reflex, Register, Tense, VerbForm, and Voice. The main table displays the following data:

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC
+ 1	Ama	ama	CCONJ	Conj	-	0	root	-	-
+ 2	annemin	anne	NOUN	Noun	Case=Gen Number=Sing Number[psor]=Sing Person=3 Person[psor]=1	3	nmod:poss	-	-
+ 3	şartları	şart	NOUN	Noun	Case=Nom Number=Plur Number[psor]=Sing Person=3 Person[psor]=3	4	nsubj	-	-
+ 4	vardı	var	VERB	Verb	Aspect=Perf Mood=Ind Number=Sing Person=3 Polarity=Pos Tense=Past	1	conj	-	SpaceAfter=No
+ 5	.	.	PUNCT	Punc	-	4	punct	-	-

Below the table, there is a text input field containing the sentence "Ama annemin şartları vardı." and a parse tree diagram showing the hierarchical structure of the sentence.

- The remaining fields are the features which are obtained from the FEATS field. The features are obtained by parsing FEATS because they are not available in the Conll files separately. They are stored in the FEATS field together. To customize the screen, the checkboxes of the fields are used.

The screenshot shows the Data Viewer interface with additional fields checked: Aspect, Case, Mood, NumType, Number, Number[psor], Polarity, and Tense. The main table displays the following data:

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC	Aspect	Case	Mood	NumType	Number	Number[psor]	Polarity	Tense
+ 1	Ama	ama	CCONJ	Conj	-	0	root	-	-	-	-	-	-	-	-	-	-
+ 2	annemin	anne	NOUN	Noun	Case=Gen Number=Sing Number[psor]=Sing Person=3 Person[psor]=1	3	nmod:poss	-	-	-	Gen	-	-	Sing	Sing	-	-
+ 3	şartları	şart	NOUN	Noun	Case=Nom Number=Plur Number[psor]=Sing Person=3 Person[psor]=3	4	nsubj	-	-	-	Nom	-	-	Plur	Sing	-	-
+ 4	vardı	var	VERB	Verb	Aspect=Perf Mood=Ind Number=Sing Person=3 Polarity=Pos Tense=Past	1	conj	-	-	SpaceAfter=No	Perf	-	Ind	-	Sing	-	Past
+ 5	.	.	PUNCT	Punc	-	4	punct	-	-	-	-	-	-	-	-	-	-

Below the table, there is a text input field containing the sentence "Ama annemin şartları vardı." and a parse tree diagram showing the hierarchical structure of the sentence.

5. Arrow keys are used to move between cells in the table. “Prev” and “Next” buttons are used to move between sentences. “Prev” button has shortcut “Alt+O” and “Next” button has shortcut “Alt+P”. To directly go to a specific sentence, write the id of the sentence and click “Go” button.

The screenshot shows the Data Viewer interface with a table of morphological and syntactic data. The table has columns for ID, FORM, LEMMA, UPOS, XPOS, FEATS, HEAD, DEPREL, DEPS, MISC, Aspect, Case, Mood, NumType, Number, Number[psor], Polarity, and Tense. The data is organized into rows, with some rows highlighted in blue. Below the table is a dependency tree diagram for the sentence "Sanal parçacıklar sa bunların hiçbirini yapamazlar." The tree shows the hierarchical structure of the sentence, with nodes representing words and their grammatical relationships.

6. The value in the cells are edited by directly typing when the focus is on that cell. To finish editing, press “Enter”. To save the file, use “Prev” or “Next” buttons. It is possible to make more than one edits at the same time.

The screenshot shows the Data Viewer interface with the same table as before. The 'Number' column for the last row (ID 7) is highlighted in blue. Below the table is a dependency tree diagram for the sentence "Sanal parçacıklar sa bunların hiçbirini yapamazlar." The tree shows the hierarchical structure of the sentence, with nodes representing words and their grammatical relationships.

7. If one of the features is edited, FEATS is updated accordingly. Keep in mind that FEATS may not contain some of the features since the features are not separate fields actually. In that case, these cells are marked by “_”. So, when the user wants to remove a feature, FEATS should be edited instead of clearing the cell of the field or writing “_”. However, “_” can be used for default fields.

Data Viewer

Prev Add Row Delete Row 1 Go Write your note here... Next

☐ Abbr ☐ Animacy ☒ Aspect ☒ Case ☐ Clusivity ☐ Definite ☐ Degree ☐ Echo ☐ Evident ☐ Foreign ☐ Gender ☐ Mood ☐ NounClass ☒ NumType ☒ Number ☒ Number[psor] ☐ Person ☐ Person[psor] ☒ Polarity ☐ Polite ☐ Poss ☐ PronType ☐ Reflex ☐ Register ☒ Tense ☐ VerbForm ☐ Voice

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC	Aspect	Case	Mood	NumType	Number	Number[psor]	Polarity	Tense
+ 1	Sanal	sanal	ADJ	Adj		2	amod										
- 2-3	parçacıklar																
+ 2	parçacıklar	parçacık	NOUN	Noun	Case=Nom Number=Plur Person=3	6	csubj				Nom			Plur			
+ 3	sa	i	AUX	Zero	Aspect=Perf Mood=Cnd Number=Sing Person=3 Tense=Pres	2	cop			Perf		Cnd		Sing			Pres
+ 4	bunların	bu	PRON	Demons	Case=Gen Number=Plur Person=3 PronType=Dem	5	nmod:poss				Gen			Plur			
+ 5	hiçbirini	hiçbiri	PRON	Quant	Case=Acc Number=Sing Number[psor]=Sing Person=3 Person[psor]=3 PronType=Ind	6	obj				Acc			Sing	Sing		
+ 6	yapamazlar	yap	VERB	Verb	Aspect=Hab Mood=Pot Number=Plur Person=3 Polarity=Neg Tense=Pres	0	root		SpaceAfter=No	Hab		Pot		Plur		Neg	Pres
+ 7	.	.	PUNCT	Punc		6	punct										

Sanal parçacıklar sa bunların hiçbirini yapamazlar.

zone= id=mst-0003

8. If your edit is not valid, an error message occurs under the table.

Data Viewer

Prev Add Row Delete Row 1 Go Write your note here... Next

☐ Abbr ☐ Animacy ☒ Aspect ☒ Case ☐ Clusivity ☐ Definite ☐ Degree ☐ Echo ☐ Evident ☐ Foreign ☐ Gender ☐ Mood ☐ NounClass ☒ NumType ☒ Number ☒ Number[psor] ☐ Person ☐ Person[psor] ☒ Polarity ☐ Polite ☐ Poss ☐ PronType ☐ Reflex ☐ Register ☒ Tense ☐ VerbForm ☐ Voice

ID	FORM	LEMMA	UPOS	XPOS	FEATS	HEAD	DEPREL	DEPS	MISC	Aspect	Case	Mood	NumType	Number	Number[psor]	Polarity	Tense
+ 1	Sanal	sanal	ADJ	Adj		2	amod										
- 2-3	parçacıklar																
+ 2	parçacıklar	parçacık	Adj	Noun	Case=Nom Number=Plur Person=3	6	csubj				Nom			Plur			
+ 3	sa	i	AUX	Zero	Aspect=Perf Mood=Cnd Number=Sing Person=3 Tense=Pres	2	cop			Perf		Cnd		Sing			Pres
+ 4	bunların	bu	PRON	Demons	Case=Gen Number=Plur Person=3 PronType=Dem	5	nmod:poss				Gen			Plur			
+ 5	hiçbirini	hiçbiri	PRON	Quant	Case=Acc Number=Sing Number[psor]=Sing Person=3 Person[psor]=3 PronType=Ind	6	obj				Acc			Sing	Sing		
+ 6	yapamazlar	yap	VERB	Verb	Aspect=Hab Mood=Pot Number=Plur Person=3 Polarity=Neg Tense=Pres	0	root		SpaceAfter=No	Hab		Pot		Plur		Neg	Pres
+ 7	.	.	PUNCT	Punc		6	punct										

ERRORS:
[Line 5 Sent mst-0003]: Invalid UPOS value Adj

Sanal parçacıklar sa bunların hiçbirini yapamazlar.

zone= id=mst-0003

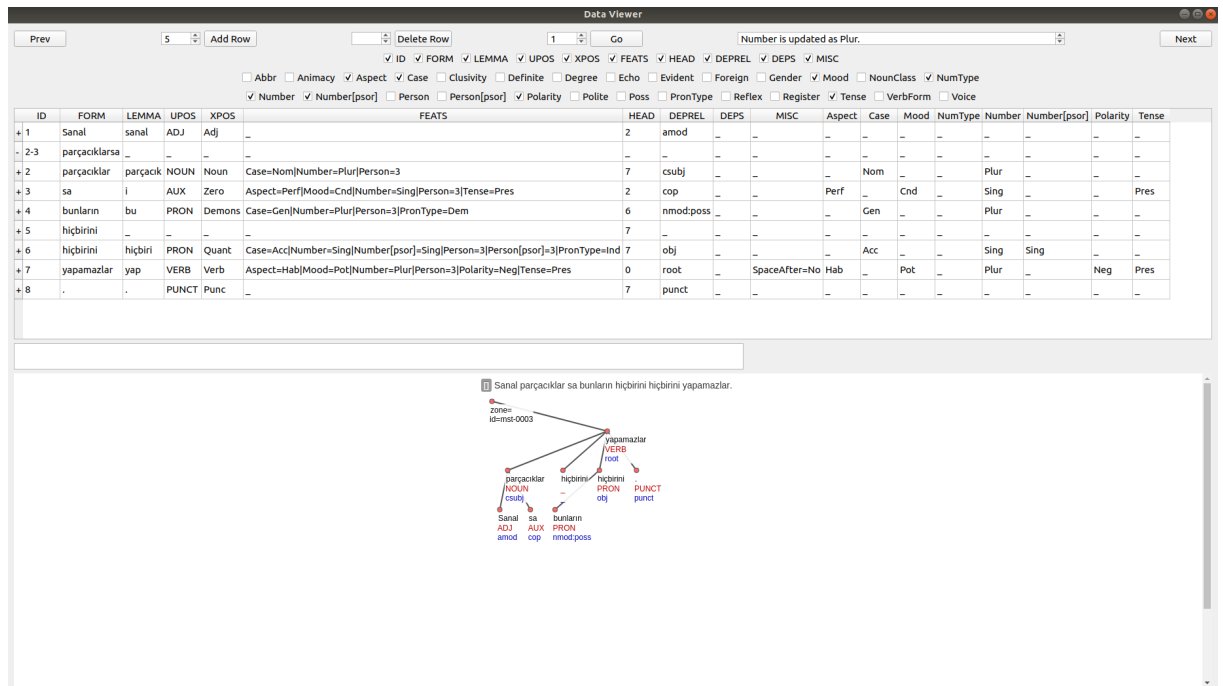
9. Each sentence may have a note attached. Shortcut for writing notes is “Alt+M”. The note is saved by using “Prev” or “Next” buttons.

The screenshot shows the 'Data Viewer' application. At the top, there are buttons for 'Prev', 'Add Row', 'Delete Row', and 'Go'. Below these are various checkboxes for linguistic features like ID, FORM, LEMMA, UPOS, XPOS, FEATS, HEAD, DEPREL, DEPS, MISC, etc. The main part of the window is a table with columns: ID, FORM, LEMMA, UPOS, XPOS, FEATS, HEAD, DEPREL, DEPS, MISC, Aspect, Case, Mood, NumType, Number, Number[psor], Polarity, Tense. The table contains 7 rows of data. Below the table, there is a text area with the sentence 'Sanal parçacıklar sa bunların hiçbirini yapamazlar.' and a syntax tree diagram. The syntax tree shows the hierarchical structure of the sentence, with nodes labeled with linguistic features like NOUN, PRON, VERB, AUX, and PUNCT.

10. Notes are stored in a separate “notes-*TheNameOfTheConllFile.txt*” file in the working directory. A note is stored together with the id of the sentence.

The screenshot shows a text editor window titled 'notes-tr_imst-ud-dev.txt' with the path '~/Desktop/BoAT-master'. The editor contains the text '1 --- Number is updated as Plur.'.

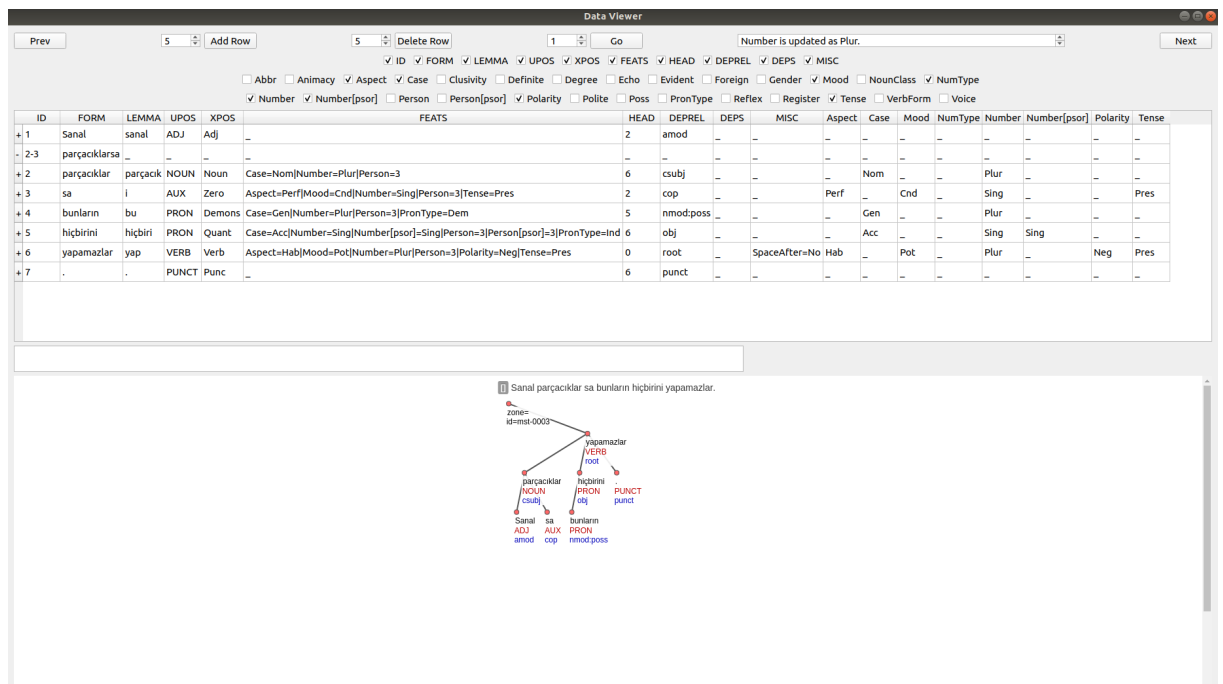
11. To add a new row, enter a row ID and click “Add Row” button. A new row is added on top of the row with entered ID. The entered row ID shouldn’t belong to a multiword expression.



The Data Viewer interface shows a table with 8 rows. The first row has ID 1, FORM 'Sanal', LEMMA 'sanal', UPOS 'ADJ', and XPOS 'Adj'. The second row has ID 2, FORM 'parçacıklar', LEMMA 'parçacık', UPOS 'NOUN', and XPOS 'Noun'. The third row has ID 3, FORM 'sa', LEMMA 'i', UPOS 'AUX', and XPOS 'Zero'. The fourth row has ID 4, FORM 'bunların', LEMMA 'bu', UPOS 'PRON', and XPOS 'Demos'. The fifth row has ID 5, FORM 'hiçbirini', LEMMA 'hiçbiri', UPOS 'PRON', and XPOS 'Quant'. The sixth row has ID 6, FORM 'yapamazlar', LEMMA 'yap', UPOS 'VERB', and XPOS 'Verb'. The seventh row has ID 7, FORM 'yapamazlar', LEMMA 'yap', UPOS 'VERB', and XPOS 'Verb'. The eighth row has ID 8, FORM 'yapamazlar', LEMMA 'yap', UPOS 'VERB', and XPOS 'Verb'. The table also includes columns for FEATS, HEAD, DEPREL, DEPS, MISC, Aspect, Case, Mood, NumType, Number, Number[psor], Polarity, and Tense.

Below the table, a dependency tree is shown for the sentence "Sanal parçacıklar sa bunların hiçbirini yapamazlar." The root node is "yapamazlar" (VERB, root). It has children "parçacıklar" (NOUN, csobj), "sa" (AUX, cop), and "bunların hiçbirini" (PRON, nmod:poss). The "parçacıklar" node has children "Sanal" (ADJ, amod) and "hiçbirini" (PRON, nmod:poss). The "bunların hiçbirini" node has children "bunların" (PRON, nmod:poss) and "hiçbirini" (PRON, nmod:poss).

12. To delete an existing row, enter a row ID and click “Delete Row” button. The row with entered ID is deleted. The entered row ID shouldn’t belong to a multiword expression and a HEAD of another token.



The Data Viewer interface shows a table with 7 rows. The first row has ID 1, FORM 'Sanal', LEMMA 'sanal', UPOS 'ADJ', and XPOS 'Adj'. The second row has ID 2, FORM 'parçacıklar', LEMMA 'parçacık', UPOS 'NOUN', and XPOS 'Noun'. The third row has ID 3, FORM 'sa', LEMMA 'i', UPOS 'AUX', and XPOS 'Zero'. The fourth row has ID 4, FORM 'bunların', LEMMA 'bu', UPOS 'PRON', and XPOS 'Demos'. The fifth row has ID 5, FORM 'hiçbirini', LEMMA 'hiçbiri', UPOS 'PRON', and XPOS 'Quant'. The sixth row has ID 6, FORM 'yapamazlar', LEMMA 'yap', UPOS 'VERB', and XPOS 'Verb'. The seventh row has ID 7, FORM 'yapamazlar', LEMMA 'yap', UPOS 'VERB', and XPOS 'Verb'. The table also includes columns for FEATS, HEAD, DEPREL, DEPS, MISC, Aspect, Case, Mood, NumType, Number, Number[psor], Polarity, and Tense.

Below the table, a dependency tree is shown for the sentence "Sanal parçacıklar sa bunların hiçbirini yapamazlar." The root node is "yapamazlar" (VERB, root). It has children "parçacıklar" (NOUN, csobj), "sa" (AUX, cop), and "bunların hiçbirini" (PRON, nmod:poss). The "parçacıklar" node has children "Sanal" (ADJ, amod) and "hiçbirini" (PRON, nmod:poss). The "bunların hiçbirini" node has children "bunların" (PRON, nmod:poss) and "hiçbirini" (PRON, nmod:poss).

13. Click exit button to quit.