Document Overview

- This is a documentation overview of how to use Batchalign2 to do syntactic and morphological parsing with examples in Italian using the VoxCommunis spoken Italian corpus

Batchalign2 Overview

- Batchalign2 is a command-line pipeline developed under the TalkBank project for automatic morphological tagging and Universal Dependencies (UD) syntactic parsing across multiple languages, including Italian.
- It takes linguistic transcripts in the CHAT (.cha) format and produces two key annotation layers:
  - %mor: morphological analysis (lemmas, parts of speech, and inflectional features)
  - %gra: syntactic analysis following the Universal Dependencies (UD) standard.
- Batchalign2 integrates the **Stanza** NLP models and the **UD framework** into the TalkBank/CLAN ecosystem, enabling consistent morphosyntactic annotation of multilingual corpora.
- It automates tokenization, morphological tagging, lemmatization, and dependency parsing, storing all results back in CHAT format—making the output directly usable with CLAN tools for further linguistic or developmental analysis.
- Functions and Commands:
  - asr: ASR!
  - morphosyntax: PoS and dependency analysis
  - fa: Forced Alignment (requires utterance-level timings already)

Universal Dependencies Overview

- **Universal Dependencies** (UD) is an international framework for the consistent annotation of grammar—covering parts of speech, morphological features, and syntactic dependencies—across human languages.
- UD is an open, collaborative project involving over 600 contributors, maintaining more than 200 treebanks in 150+ languages.
- Its goal is to create a cross-linguistically consistent representation of grammatical relations, allowing comparable syntactic analysis across languages like Italian, Mandarin, English, and many others.

What we need to do in terms of using CLAN and Batchalign2 to do morphological and syntactic analysis and parsing:

1. **Download CLAN (MacBook or Windows, or for a more detailed overview of CLAN, please see → here)**
   https://dali.talkbank.org/clan/

2. **Download Batchalign/Environment Setup (Open MacBook Terminal and paste the following)**

   ```
   # Install UV (a lightweight Python environment manager)
   curl -LsSf https://astral.sh/uv/install.sh | sh

   # Install Batchalign2 for Python 3.11
   UV_PYTHON=3.11 uv tool install batchalign

   # Ensure ~/.local/bin is in your PATH
   echo 'export PATH="$HOME/.local/bin:$PATH"' >> ~/.zshrc
   source ~/.zshrc

   # Check installation
   batchalign --help
   ```

3. **Data Preprocessing (R Studio) → also see txt-to-cha.R**
   Aim data cleaning: Txt file → cha file
   - A valid .cha file must follow CHAT conventions: every utterance on its own line and starting with a speaker code (e.g., *ADU: ,*CHI: , *MOT:).

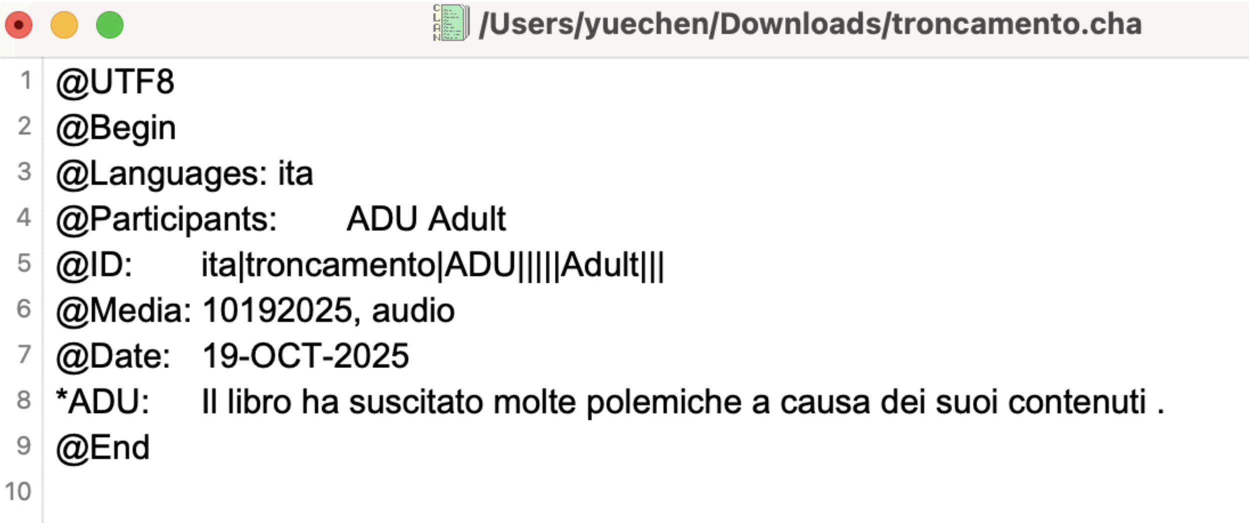
   Tips:
   - A Cha file can be automatically converted to a txt file and vice versa (just **rename** the a cha file to a txt file! **Example**: Sample.cha → Sample.txt)
   - @Languages: must include the language code (e.g., ita, zho, eng, spa).
     - The Italian model will be loaded automatically from Stanza/UD based on **@Languages: ita**.
   - Sentences must end with a space and punctuation (., !, or ?).
   - Delete empty or malformed tokens like <>, which cause lexer errors.

   Things we need to do for the VoxCommunis dataset:
   1. Add speaker tier at the beginning of each sentence (e.g., *ADU:)
   2. Remove all other symbols (e.g., @, *, % or <>:)
   3. Ensure a space before punctuations (e.g., . ! or ?)
   4. Add a final period if missing
   5. Make sure to add the following at the beginning of the cha file

      ```
      @UTF8
      @Begin
      @Languages: ita
      @Participants: ADU Adult
      @ID: ita|troncamento|ADU|||||Adult|||
      ```

6. Make sure to add the following at the end of the cha file
   @End

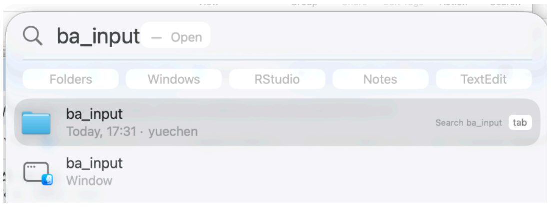
Successful Example:
@UTF8 → Must
@Begin → Must
@Languages: ita → Must
@Participants: ADU Adult → Must
@ID: ita|troncamento|ADU|||||Adult||| → Must
@Media: 10192025, audio → optional
@Date: 19-OCT-2025 → optional
*ADU:Il libro ha suscitato molte polemiche a causa dei suoi contenuti . → Must
@End → Must



/Users/yuechen/Downloads/troncamento.cha

```
@UTF8
@Begin
@Languages: ita
@Participants:	ADU Adult
@ID:	ita|troncamento|ADU|||||Adult|||
@Media: 10192025, audio
@Date:	19-OCT-2025
*ADU:	Il libro ha suscitato molte polemiche a causa dei suoi contenuti .
@End
```

Sample Input file (where you should put it? Search ba_input!):

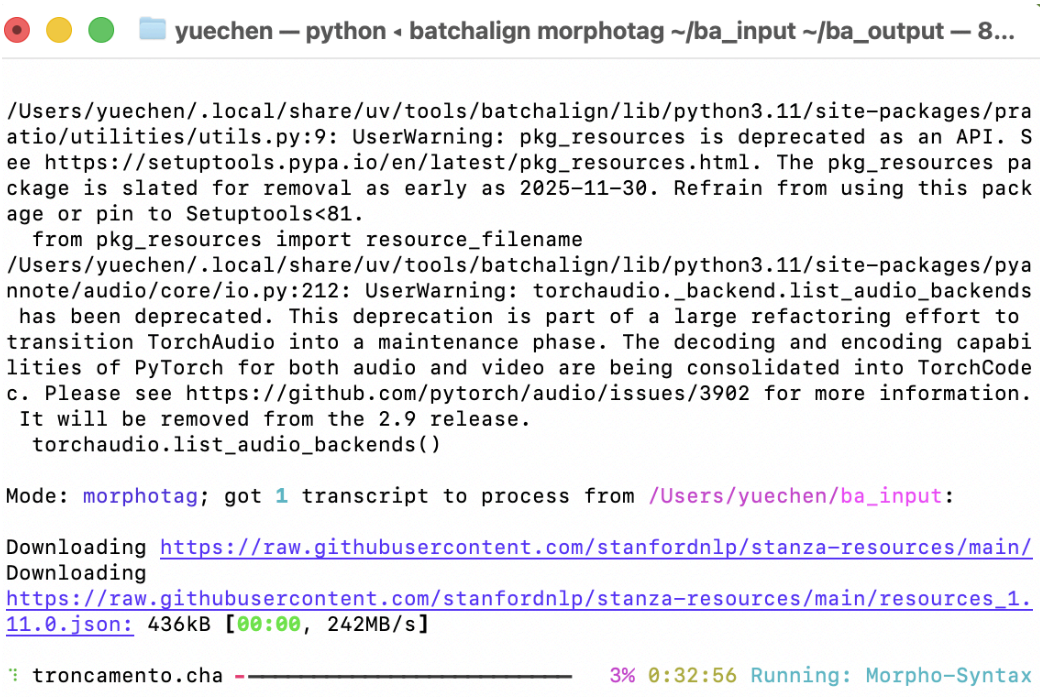


4. Parsing: Commands used for Batchalign2 (MacBook Terminal)

```
# check input file
sed -n '1,25p' ~/ba_input/troncamento.cha
```

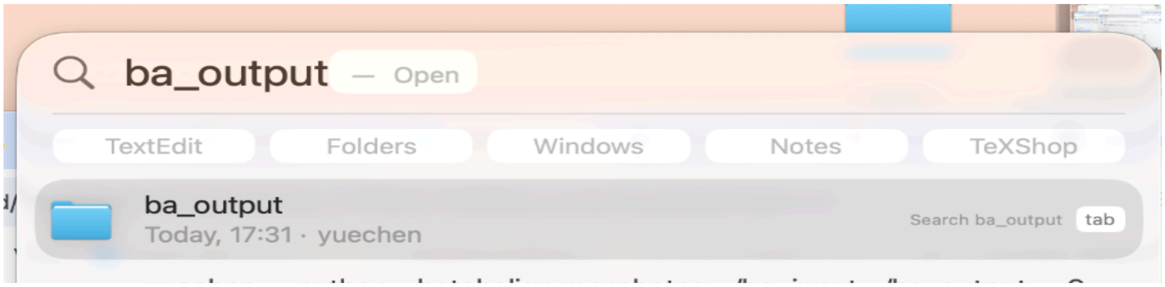
```
# run batch aligner
batchalign morphotag ~/ba_input ~/ba_output

# check output file %mor / %gra
grep -n "^%mor:\|^%gra:" ~/ba_output/troncamento.cha | head
```

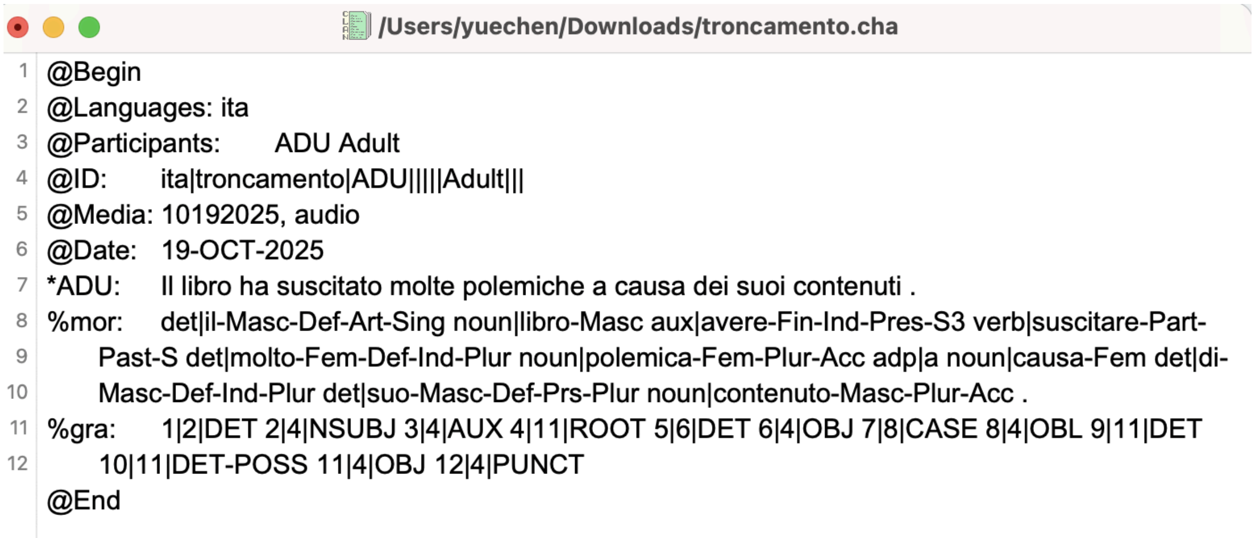
What you will see if it's successfully running:


yuechen — python ◂ batchalign morphotag ~/ba_input ~/ba_output — 8...

```
/Users/yuechen/.local/share/uv/tools/batchalign/lib/python3.11/site-packages/pra
atio/utilities/utils.py:9: UserWarning: pkg_resources is deprecated as an API. S
ee https://setuptools.pypa.io/en/latest/pkg_resources.html. The pkg_resources pa
ckage is slated for removal as early as 2025-11-30. Refrain from using this pack
age or pin to Setuptools<81.
  from pkg_resources import resource_filename
/Users/yuechen/.local/share/uv/tools/batchalign/lib/python3.11/site-packages/pya
nnote/audio/core/io.py:212: UserWarning: torchaudio._backend.list_audio_backends
 has been deprecated. This deprecation is part of a large refactoring effort to
transition TorchAudio into a maintenance phase. The decoding and encoding capabi
lities of PyTorch for both audio and video are being consolidated into TorchCode
c. Please see https://github.com/pytorch/audio/issues/3902 for more information.
 It will be removed from the 2.9 release.
  torchaudio.list_audio_backends()

Mode: morphotag; got 1 transcript to process from /Users/yuechen/ba_input:

Downloading https://raw.githubusercontent.com/stanfordnlp/stanza-resources/main/
Downloading
https://raw.githubusercontent.com/stanfordnlp/stanza-resources/main/resources_1.
11.0.json: 436kB [00:00, 242MB/s]

⠹ troncamento.cha ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━   3% 0:32:56 Running: Morpho-Syntax
```

Sample Output file (where you can find it? Search ba_output!):





```
@Begin
@Languages: ita
@Participants:	ADU Adult
@ID:	ita|troncamento|ADU|||||Adult|||
@Media: 10192025, audio
@Date:	19-OCT-2025
*ADU:	Il libro ha suscitato molte polemiche a causa dei suoi contenuti .
%mor:	det|il-Masc-Def-Art-Sing noun|libro-Masc aux|avere-Fin-Ind-Pres-S3 verb|suscitare-Part-
	Past-S det|molto-Fem-Def-Ind-Plur noun|polemica-Fem-Plur-Acc adp|a noun|causa-Fem det|di-
	Masc-Def-Ind-Plur det|suo-Masc-Def-Prs-Plur noun|contenuto-Masc-Plur-Acc .
%gra:	1|2|DET 2|4|NSUBJ 3|4|AUX 4|11|ROOT 5|6|DET 6|4|OBJ 7|8|CASE 8|4|OBL 9|11|DET
	10|11|DET-POSS 11|4|OBJ 12|4|PUNCT
@End
```

5. Understanding the Output:

Sample sentence: Il libro ha suscitato molte polemiche a causa dei suoi contenuti.
English Translation: The book has sparked much controversy because of its contents.

Morphological layer (%mor) (also see the column named **morphonological_analysis** in the it_vxc_spkr17_with_parsing.tsv file**):**

%mor: det|il-Masc-Def-Art-Sing noun|libro-Masc aux|avere-Fin-Ind-Pres-S3 verb|suscitare-Part-Past-S det|molto-Fem-Def-Ind-Plur noun|polemica-Fem-Plur-Acc adp|a noun|causa-Fem det|di-Masc-Def-Ind-Plur det|suo-Masc-Def-Prs-Plur noun|contenuto-Masc-Plur-Acc .

What does the Morphological layer (%mor) tire do?

- Breaks the sentence down word-by-word and encodes part of speech + grammatical features (gender, number, tense, etc.).
  - Lemma (dictionary form): e.g. libro, suscitare
  - Part of speech: noun, verb, det, adp, etc.
  - Morphosyntactic features:
    - Gender: Masc/Fem
    - Number: Sing/Plur
    - Tense: Pres, Past, Fut
    - Mood: Ind (indicative), Subj (subjunctive)
    - Person: S1, S2, S3 (subject 1st, 2nd, 3rd person)
    - Case: Nom (nominative), Acc (accusative), etc.

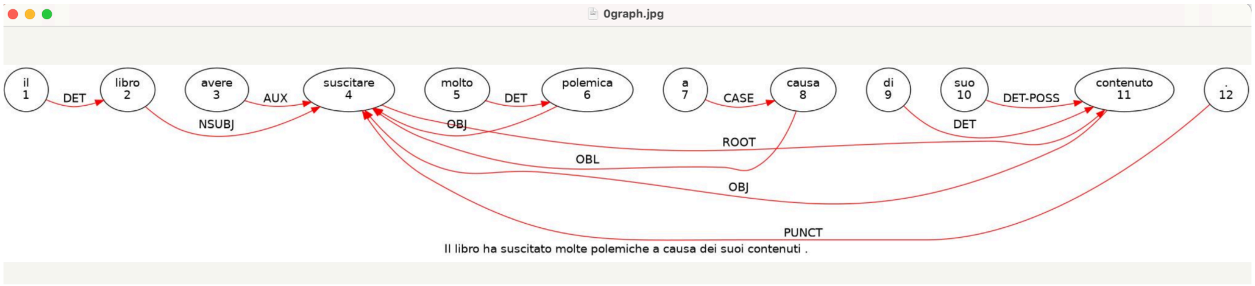
Syntactic layer (%gra) **(**also see the column named **syntactic_parsing** in the it_vxc_spkr17_with_parsing.tsv file**):**

%gra: 1|2|DET 2|4|NSUBJ 3|4|AUX 4|11|ROOT 5|6|DET 6|4|OBJ 7|8|CASE 8|4|OBL 9|11|DET 10|11|DET-POSS 11|4|OBJ 12|4|PUNCT

What does the Syntactic layer (%gra) tire do?

- These lines show token IDs (word order, such as word 1, word 2, word 3), dependency heads, and UD relations (e.g., NSUBJ = subject, OBL = oblique).

Example of the Universal Dependencies (UD) dependency tree:

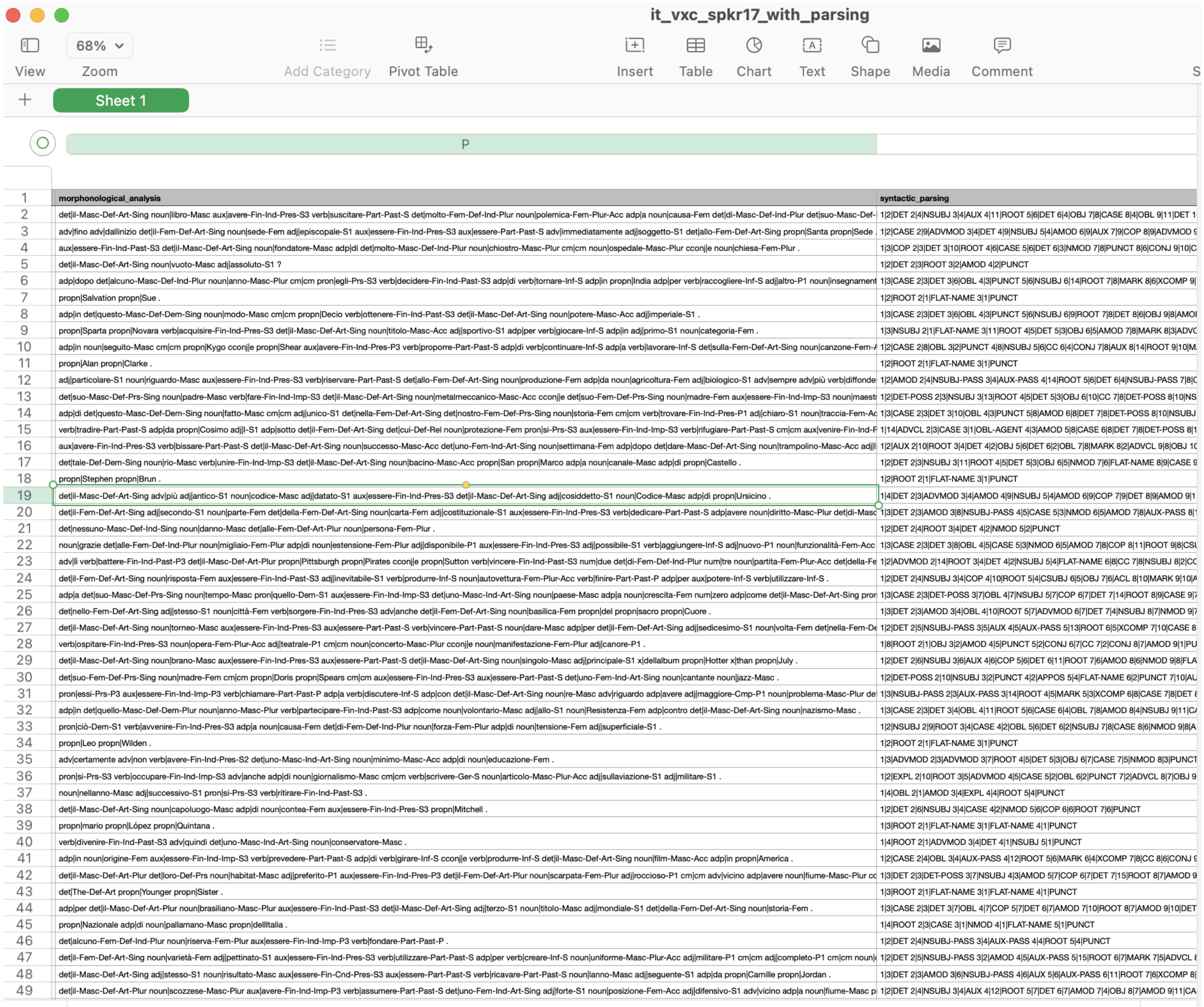


Il libro ha suscitato molte polemiche a causa dei suoi contenuti .

Overall, what this shows:

- Each oval = a token (word), labeled with
    - its lemma (dictionary form), and
    - its index (1–12 here).
- Each red arrow = a dependency relation between two words:
    - The arrow points from the head (governing word) to its dependent (modified word).
    - The label on the arrow (e.g., NSUBJ, AUX, OBJ, OBL, etc.) indicates the syntactic relation type.

Final sample TSV file output:



1	morphonological_analysis	syntactic_parsing
2	det|il-Masc-Def-Art-Sing noun|libro-Masc aux|avere-Fin-Ind-Pres-S3 verb|suscitare-Part-Past-S det|molto-Fem-Def-Ind-Plur noun|polemica-Fem-Plur-Acc adp|a noun|causa-Fem det|di-Masc-Def-Ind-Plur det|suo-Masc-Def-	1|2|DET 2|4|NSUBJ 3|4|AUX 4|11|ROOT 5|6|DET 6|4|OBJ 7|8|CASE 8|4|OBL 9|11|DET 1
3	adv|fino adv|dallinizio det|il-Fem-Def-Art-Sing noun|sede-Fem adj|episcopale-S1 aux|essere-Fin-Ind-Pres-S3 aux|essere-Part-Past-S adv|immediatamente adj|soggetto-S1 det|allo-Fem-Def-Art-Sing propn|Santa propn|Sede .	1|2|CASE 2|9|ADVMOD 3|4|DET 4|9|NSUBJ 5|4|AMOD 6|9|AUX 7|9|COP 8|9|ADVMOD 9
4	aux|essere-Fin-Ind-Past-S3 det|il-Masc-Def-Art-Sing noun|fondatore-Masc adp|di det|molto-Masc-Def-Ind-Plur noun|chiostro-Masc-Plur cm|cm noun|ospedale-Masc-Plur cconj|e noun|chiesa-Fem-Plur .	1|3|COP 2|3|DET 3|10|ROOT 4|6|CASE 5|6|DET 6|3|NMOD 7|8|PUNCT 8|6|CONJ 9|10|C
5	det|il-Masc-Def-Art-Sing noun|vuoto-Masc adj|assoluto-S1 ?	1|2|DET 2|3|ROOT 3|2|AMOD 4|2|PUNCT
6	adp|dopo det|alcuno-Masc-Def-Ind-Plur noun|anno-Masc-Plur cm|cm pron|egli-Prs-S3 verb|decidere-Fin-Ind-Past-S3 adp|di verb|tornare-Inf-S adp|in propn|India adp|per verb|raccogliere-Inf-S adj|altro-P1 noun|insegnament	1|3|CASE 2|3|DET 3|6|OBL 4|3|PUNCT 5|6|NSUBJ 6|14|ROOT 7|8|MARK 8|6|XCOMP 9|
7	propn|Salvation propn|Sue .	1|2|ROOT 2|1|FLAT-NAME 3|1|PUNCT
8	adp|in det|questo-Masc-Def-Dem-Sing noun|modo-Masc cm|cm propn|Decio verb|ottenere-Fin-Ind-Past-S3 det|il-Masc-Def-Art-Sing noun|potere-Masc-Acc adj|imperiale-S1 .	1|3|CASE 2|3|DET 3|6|OBL 4|3|PUNCT 5|6|NSUBJ 6|9|ROOT 7|8|DET 8|6|OBJ 9|8|AMOI
9	propn|Sparta propn|Novara verb|acquisire-Fin-Ind-Pres-S3 det|il-Masc-Def-Art-Sing noun|titolo-Masc-Acc adj|sportivo-S1 adp|per verb|giocare-Inf-S adp|in adj|primo-S1 noun|categoria-Fem .	1|3|NSUBJ 2|1|FLAT-NAME 3|11|ROOT 4|5|DET 5|3|OBJ 6|5|AMOD 7|8|MARK 8|3|ADVC
10	adp|in noun|seguito-Masc cm|cm propn|Kygo cconj|e propn|Shear aux|avere-Fin-Ind-Pres-P3 verb|proporre-Part-Past-S adp|di verb|continuare-Inf-S adp|a verb|lavorare-Inf-S det|sulla-Fem-Def-Art-Sing noun|canzone-Fem-A	1|2|CASE 2|8|OBL 3|2|PUNCT 4|8|NSUBJ 5|6|CC 6|4|CONJ 7|8|AUX 8|14|ROOT 9|10|M.
11	propn|Alan propn|Clarke .	1|2|ROOT 2|1|FLAT-NAME 3|1|PUNCT
12	adj|particolare-S1 noun|riguardo-Masc aux|essere-Fin-Ind-Pres-S3 verb|riservare-Part-Past-S det|allo-Fem-Def-Art-Sing noun|produzione-Fem adp|da noun|agricoltura-Fem adj|biologico-S1 adv|sempre adv|più verb|diffonde	1|2|AMOD 2|4|NSUBJ-PASS 3|4|AUX-PASS 4|14|ROOT 5|6|DET 6|4|NSUBJ-PASS 7|8|C
13	det|suo-Masc-Def-Prs-Sing noun|padre-Masc verb|fare-Fin-Ind-Imp-S3 det|il-Masc-Def-Art-Sing noun|metalmeccanico-Masc-Acc cconj|e det|suo-Fem-Def-Prs-Sing noun|madre-Fem aux|essere-Fin-Ind-Imp-S3 noun|maestr	1|2|DET-POSS 2|3|NSUBJ 3|13|ROOT 4|5|DET 5|3|OBJ 6|10|CC 7|8|DET-POSS 8|10|NS
14	adp|di det|questo-Masc-Def-Dem-Sing noun|fatto-Masc cm|cm adj|unico-S1 det|nella-Fem-Def-Art-Sing det|nostro-Fem-Def-Prs-Sing noun|storia-Fem cm|cm verb|trovare-Fin-Ind-Pres-P1 adj|chiaro-S1 noun|traccia-Fem-Ac	1|3|CASE 2|3|DET 3|10|OBL 4|3|PUNCT 5|8|AMOD 6|8|DET 7|8|DET-POSS 8|10|NSUBJ
15	verb|tradire-Part-Past-S adp|da propn|Cosimo adj|I-S1 adp|sotto det|il-Fem-Def-Art-Sing det|cui-Def-Rel noun|protezione-Fem pron|si-Prs-S3 aux|essere-Fin-Ind-Imp-S3 verb|rifugiare-Part-Past-S cm|cm aux|venire-Fin-Ind-P	1|14|ADVCL 2|3|CASE 3|1|OBL-AGENT 4|3|AMOD 5|8|CASE 6|8|DET 7|8|DET-POSS 8|1
16	aux|avere-Fin-Ind-Pres-S3 verb|bissare-Part-Past-S det|il-Masc-Def-Art-Sing noun|successo-Masc-Acc det|uno-Fem-Ind-Art-Sing noun|settimana-Fem adp|dopo det|dare-Masc-Def-Art-Sing noun|trampolino-Masc-Acc adj|l	1|2|AUX 2|10|ROOT 3|4|DET 4|2|OBJ 5|6|DET 6|2|OBL 7|8|MARK 8|2|ADVCL 9|8|OBJ 1C
17	det|tale-Def-Dem-Sing noun|rio-Masc verb|unire-Fin-Ind-Imp-S3 det|il-Masc-Def-Art-Sing noun|bacino-Masc-Acc propn|San propn|Marco adp|a noun|canale-Masc adp|di propn|Castello .	1|2|DET 2|3|NSUBJ 3|11|ROOT 4|5|DET 5|3|OBJ 6|5|NMOD 7|6|FLAT-NAME 8|9|CASE 9
18	propn|Stephen propn|Brun .	1|2|ROOT 2|1|FLAT-NAME 3|1|PUNCT
19	det|il-Masc-Def-Art-Sing adv|più adj|antico-S1 noun|codice-Masc adj|datato-S1 aux|essere-Fin-Ind-Pres-S3 det|il-Masc-Def-Art-Sing adj|cosiddetto-S1 noun|Codice-Masc adp|di propn|Ursicino .	1|4|DET 2|3|ADVMOD 3|4|AMOD 4|9|NSUBJ 5|4|AMOD 6|9|COP 7|9|DET 8|9|AMOD 9|1
20	det|il-Fem-Def-Art-Sing adj|secondo-S1 noun|parte-Fem det|della-Fem-Def-Art-Sing noun|carta-Fem adj|costituzionale-S1 aux|essere-Fin-Ind-Pres-S3 verb|dedicare-Part-Past-S adp|avere noun|diritto-Masc-Plur det|di-Masc	1|3|DET 2|3|AMOD 3|8|NSUBJ-PASS 4|5|CASE 5|3|NMOD 6|5|AMOD 7|8|AUX-PASS 8|1
21	det|nessuno-Masc-Def-Ind-Sing noun|danno-Masc det|alle-Fem-Def-Art-Plur noun|persona-Fem-Plur .	1|2|DET 2|4|ROOT 3|4|DET 4|2|NMOD 5|2|PUNCT
22	noun|grazie det|alle-Fem-Def-Ind-Plur noun|migliaio-Fem-Plur adp|di noun|estensione-Fem-Plur adj|disponibile-P1 aux|essere-Fin-Ind-Pres-S3 adj|possibile-S1 verb|aggiungere-Inf-S adj|nuovo-P1 noun|funzionalità-Fem-Acc	1|3|CASE 2|3|DET 3|8|OBL 4|5|CASE 5|3|NMOD 6|5|AMOD 7|8|COP 8|11|ROOT 9|8|CSU
23	adv|li verb|battere-Fin-Ind-Past-P3 det|il-Masc-Def-Art-Plur propn|Pittsburgh propn|Pirates cconj|e propn|Sutton verb|vincere-Fin-Ind-Past-S3 num|due det|di-Fem-Def-Ind-Plur num|tre noun|partita-Fem-Plur-Acc det|della-Fe	1|2|ADVMOD 2|14|ROOT 3|4|DET 4|2|NSUBJ 5|4|FLAT-NAME 6|8|CC 7|8|NSUBJ 8|2|CC
24	det|il-Fem-Def-Art-Sing noun|risposta-Fem aux|essere-Fin-Ind-Past-S3 adj|inevitabile-S1 verb|produrre-Inf-S noun|autovettura-Fem-Plur-Acc verb|finire-Part-Past-P adp|per aux|potere-Inf-S verb|utilizzare-Inf-S .	1|2|DET 2|4|NSUBJ 3|4|COP 4|10|ROOT 5|4|CSUBJ 6|5|OBJ 7|6|ACL 8|10|MARK 9|10|A
25	adp|a det|suo-Masc-Def-Prs-Sing noun|tempo-Masc pron|quello-Dem-S1 aux|essere-Fin-Ind-Imp-S3 det|uno-Masc-Ind-Art-Sing noun|paese-Masc adp|a noun|crescita-Fem num|zero adp|come det|il-Masc-Def-Art-Sing pror	1|3|CASE 2|3|DET-POSS 3|7|OBL 4|7|NSUBJ 5|7|COP 6|7|DET 7|14|ROOT 8|9|CASE 9|7
26	det|nello-Fem-Def-Art-Sing adj|stesso-S1 noun|città-Fem verb|sorgere-Fin-Ind-Pres-S3 adv|anche det|il-Fem-Def-Art-Sing noun|basilica-Fem propn|del propn|sacro propn|Cuore .	1|3|DET 2|3|AMOD 3|4|OBL 4|10|ROOT 5|7|ADVMOD 6|7|DET 7|4|NSUBJ 8|7|NMOD 9|7
27	det|il-Masc-Def-Art-Sing noun|torneo-Masc aux|essere-Fin-Ind-Pres-S3 aux|essere-Part-Past-S verb|vincere-Part-Past-S noun|dare-Masc adp|per det|il-Fem-Def-Art-Sing adj|sedicesimo-S1 noun|volta-Fem det|nella-Fem-De	1|2|DET 2|5|NSUBJ-PASS 3|5|AUX 4|5|AUX-PASS 5|13|ROOT 6|5|XCOMP 7|10|CASE 8
28	verb|ospitare-Fin-Ind-Pres-S3 noun|opera-Fem-Plur-Acc adj|teatrale-P1 cm|cm noun|concerto-Masc-Plur cconj|e noun|manifestazione-Fem-Plur adj|canore-P1 .	1|8|ROOT 2|1|OBJ 3|2|AMOD 4|5|PUNCT 5|2|CONJ 6|7|CC 7|2|CONJ 8|7|AMOD 9|1|PU
29	det|il-Masc-Def-Art-Sing noun|brano-Masc aux|essere-Fin-Ind-Pres-S3 aux|essere-Part-Past-S det|il-Masc-Def-Art-Sing noun|singolo-Masc adj|principale-S1 x|dellalbum propn|Hotter x|than propn|July .	1|2|DET 2|6|NSUBJ 3|6|AUX 4|6|COP 5|6|DET 6|11|ROOT 7|6|AMOD 8|6|NMOD 9|8|FLA
30	det|suo-Fem-Def-Prs-Sing noun|madre-Fem cm|cm propn|Doris propn|Spears cm|cm aux|essere-Fin-Ind-Pres-S3 aux|essere-Part-Past-S det|uno-Fem-Ind-Art-Sing noun|cantante noun|jazz-Masc .	1|2|DET-POSS 2|10|NSUBJ 3|2|PUNCT 4|2|APPOS 5|4|FLAT-NAME 6|2|PUNCT 7|10|AU
31	pron|essi-Prs-P3 aux|essere-Fin-Ind-Imp-P3 verb|chiamare-Part-Past-P adp|a verb|discutere-Inf-S adp|con det|il-Masc-Def-Art-Sing noun|re-Masc adv|riguardo adp|avere adj|maggiore-Cmp-P1 noun|problema-Masc-Plur det	1|3|NSUBJ-PASS 2|3|AUX-PASS 3|14|ROOT 4|5|MARK 5|3|XCOMP 6|8|CASE 7|8|DET 8
32	adp|in det|quello-Masc-Def-Dem-Plur noun|anno-Masc-Plur verb|partecipare-Fin-Ind-Past-S3 adp|come noun|volontario-Masc adj|allo-S1 noun|Resistenza-Fem adp|contro det|il-Masc-Def-Art-Sing noun|nazismo-Masc .	1|3|CASE 2|3|DET 3|4|OBL 4|11|ROOT 5|6|CASE 6|4|OBL 7|8|AMOD 8|4|NSUBJ 9|11|CA
33	pron|ciò-Dem-S1 verb|avvenire-Fin-Ind-Pres-S3 adp|a noun|causa-Fem det|di-Fem-Def-Ind-Plur noun|forza-Fem-Plur adp|di noun|tensione-Fem adj|superficiale-S1 .	1|2|NSUBJ 2|9|ROOT 3|4|CASE 4|2|OBL 5|6|DET 6|2|NSUBJ 7|8|CASE 8|6|NMOD 9|8|A
34	propn|Leo propn|Wilden .	1|2|ROOT 2|1|FLAT-NAME 3|1|PUNCT
35	adv|certamente adv|non verb|avere-Fin-Ind-Pres-S2 det|uno-Masc-Ind-Art-Sing noun|minimo-Masc-Acc adp|di noun|educazione-Fem .	1|3|ADVMOD 2|3|ADVMOD 3|7|ROOT 4|5|DET 5|3|OBJ 6|7|CASE 7|5|NMOD 8|3|PUNCT
36	pron|si-Prs-S3 verb|occupare-Fin-Ind-Imp-S3 adv|anche adp|di noun|giornalismo-Masc cm|cm verb|scrivere-Ger-S noun|articolo-Masc-Plur-Acc adj|sullaviazione-S1 adj|militare-S1 .	1|2|EXPL 2|10|ROOT 3|5|ADVMOD 4|5|CASE 5|2|OBL 6|2|PUNCT 7|2|ADVCL 8|7|OBJ 9
37	noun|nellanno-Masc adj|successivo-S1 pron|si-Prs-S3 verb|ritirare-Fin-Ind-Past-S3 .	1|4|OBL 2|1|AMOD 3|4|EXPL 4|4|ROOT 5|4|PUNCT
38	det|il-Masc-Def-Art-Sing noun|capoluogo-Masc adp|di noun|contea-Fem aux|essere-Fin-Ind-Pres-S3 propn|Mitchell .	1|2|DET 2|6|NSUBJ 3|4|CASE 4|2|NMOD 5|6|COP 6|6|ROOT 7|6|PUNCT
39	propn|mario propn|López propn|Quintana .	1|3|ROOT 2|1|FLAT-NAME 3|1|FLAT-NAME 4|1|PUNCT
40	verb|divenire-Fin-Ind-Past-S3 adv|quindi det|uno-Masc-Ind-Art-Sing noun|conservatore-Masc .	1|4|ROOT 2|1|ADVMOD 3|4|DET 4|1|NSUBJ 5|1|PUNCT
41	adp|in noun|origine-Fem aux|essere-Fin-Ind-Imp-S3 verb|prevedere-Part-Past-S adp|di verb|girare-Inf-S cconj|e verb|produrre-Inf-S det|il-Masc-Def-Art-Sing noun|film-Masc-Acc adp|in propn|America .	1|2|CASE 2|4|OBL 3|4|AUX-PASS 4|12|ROOT 5|6|MARK 6|4|XCOMP 7|8|CC 8|6|CONJ 9
42	det|il-Masc-Def-Art-Plur det|loro-Def-Prs noun|habitat-Masc adj|preferito-P1 aux|essere-Fin-Ind-Pres-P3 det|il-Fem-Def-Art-Plur noun|scarpata-Fem-Plur adj|roccioso-P1 cm|cm adv|vicino adp|avere noun|fiume-Masc-Plur cc	1|3|DET 2|3|DET-POSS 3|7|NSUBJ 4|3|AMOD 5|7|COP 6|7|DET 7|15|ROOT 8|7|AMOD 9
43	det|The-Def-Art propn|Younger propn|Sister .	1|3|ROOT 2|1|FLAT-NAME 3|1|FLAT-NAME 4|1|PUNCT
44	adp|per det|il-Masc-Def-Art-Plur noun|brasiliano-Masc-Plur aux|essere-Fin-Ind-Past-S3 det|il-Masc-Def-Art-Sing adj|terzo-S1 noun|titolo-Masc adj|mondiale-S1 det|della-Fem-Def-Art-Sing noun|storia-Fem .	1|3|CASE 2|3|DET 3|7|OBL 4|7|COP 5|7|DET 6|7|AMOD 7|10|ROOT 8|7|AMOD 9|10|DET
45	propn|Nazionale adp|di noun|pallamano-Masc propn|dellItalia .	1|4|ROOT 2|3|CASE 3|1|NMOD 4|1|FLAT-NAME 5|1|PUNCT
46	det|alcuno-Fem-Def-Ind-Plur noun|riserva-Fem-Plur aux|essere-Fin-Ind-Imp-P3 verb|fondare-Part-Past-P .	1|2|DET 2|4|NSUBJ-PASS 3|4|AUX-PASS 4|4|ROOT 5|4|PUNCT
47	det|il-Fem-Def-Art-Sing noun|varietà-Fem adj|pettinato-S1 aux|essere-Fin-Ind-Pres-S3 verb|utilizzare-Part-Past-S adp|per verb|creare-Inf-S noun|uniforme-Masc-Plur-Acc adj|militare-P1 cm|cm adj|completo-P1 cm|cm noun|	1|2|DET 2|5|NSUBJ-PASS 3|2|AMOD 4|5|AUX-PASS 5|15|ROOT 6|7|MARK 7|5|ADVCL 8
48	det|il-Masc-Def-Art-Sing adj|stesso-S1 noun|risultato-Masc aux|essere-Fin-Cnd-Pres-S3 aux|essere-Part-Past-S verb|ricavare-Part-Past-S noun|lanno-Masc adj|seguente-S1 adp|da propn|Camille propn|Jordan .	1|3|DET 2|3|AMOD 3|6|NSUBJ-PASS 4|6|AUX 5|6|AUX-PASS 6|11|ROOT 7|6|XCOMP 8|
49	det|il-Masc-Def-Art-Plur noun|scozzese-Masc-Plur aux|avere-Fin-Ind-Imp-P3 verb|assumere-Part-Past-S det|uno-Fem-Ind-Art-Sing adj|forte-S1 noun|posizione-Fem-Acc adj|difensivo-S1 adv|vicino adp|a noun|fiume-Masc p	1|2|DET 2|4|NSUBJ 3|4|AUX 4|12|ROOT 5|7|DET 6|7|AMOD 7|4|OBJ 8|7|AMOD 9|11|CA

R code of how to merge the CHA file to the TSV file:

- See Cha-to-TSV.R

6. Some CLAN Post-processing (optional, if you want to look at a different analysis)

- Once %mor and %gra tiers exist, you can analyze the Italian corpus with CLAN tools:
    - **kwal +t%mor ~/ba_output/troncamento.cha**   # view morphology
        - **Command example:**
        - kwal +t%mor /Users/yuechen/ba_output/troncamento.cha
    - **kwal +t%gra +s"OBL" ~/ba_output/troncamento.cha** # search dependencies
    - **freq +t%mor -t* ~/ba_output/troncamento.cha**   # POS frequency

7. Citation and Documentation

- Batchalign2: https://talkbank.github.io/batchalign2/
- CLAN Manual, section on Batchalign and role conversion

- **Reference article**: Liu, H. & MacWhinney, B. (2024). Morphosyntactic Analysis for CHILDES. Language Development Research, 4(1), 233–258.