Text Simplification

Tom Wysocki
Peter Pribil
Margot Guettier
Morgan Ruiz-Huidobro

December 10th 2021

Advancement

- Continue syntactic simplification
- Begin lexicon simplification
- Interface creation
- Find a large dataset

Syntactic simplification

Relative clause:

Using tag instead of dependency relation:

She lives in New York, **which she likes/hates** ⟶ She **likes/hates to** live in New York.

Restrictive vs Non-restrictive

the cat **that is big** is mine ⟶ The **big** cat is mine.
my cat, **that has blue eyes** , sleeps in the garden ⟶ My cat sleeps in the garden

I bought a new car **that is very fast.**	I bought a **very fast** new car.
I'd like to paint the room a color **that is cheerful**.	I 'd like to paint the room a **cheerful** color.
I'm buying a gift for the neighbor **who helped me move.**	I ' m buying a gift for the neighbor. **The neighbor helped me move.**
The dog **that eats the cat** **that eats the mouse** **that is white** is big.	The dog is big. **The dog eats the cat**. **The cat eats the white mouse.**

Original sentence	passive to active	relative clause simplification
The apple **that is big** **is eaten by John.**	**John eats** the apple that is big.	John eats the **big** apple.
The sky, **which is blue**, **is observed by a group of children.**	**A group of children observes** the sky, which is blue,.	A group of children observes the sky.
Yesterday, an apple was **eaten by a fish that was red**	**A fish** that was red yesterday **ate** an apple,	A **red** yesterday fish ate an apple,.
The dog **was bitten by the fish**	**The fish bited** the dog.	**The fish bited** the dog.

Prepositional Phrases

Problematic parsing: pps might get different ancestors.
"are":
With: non-verbal
In/At/By: verbal
Not predictable: other inflections of "be" might have different ancestors.

So far: All pps removed except: "to", "of",
Testing: word.lemma_ != "be"

```
You, in the front row, are amazing. ---> You, in the front row, are amazing.
You, in the front row, are amazing.
I saw a monkey with a telescope. ---> I saw a monkey .
I saw a monkey .
I'm doing this with pleasure. ---> I'm doing this .
I'm doing this .
The guy with a knife threatened the shopkeeper. ---> The guy threatened the shopkeeper.
The guy threatened the shopkeeper.
Gus trash talks about his friends behind their backs. ---> Gus trash talks .
Gus trash talks .
The lady in the toilet is taking a lot of time. ---> The lady is taking a lot of time.
The lady is taking a lot of time.
A lady with a pillow hit a man in a wheelchair. ---> A lady hit a man .
A lady hit a man .
I was on time. ---> I was on time.
I was on time.
The car dealership told his customer with excitement to pay with coins behind the shop. ---> The car dealership told his customer to pay .
The car dealership told his customer to pay .
With all due respect, I disagree. ---> , I disagree.
, I disagree.
He shouted at his dog with a whistle. ---> He shouted .
He shouted .
We decided that we'll rent a car with enthusiasm. ---> We decided that we'll rent a car .
We decided that we'll rent a car .
Karen found her meaning in life in the shower. ---> Karen found her meaning .
Karen found her meaning .
Hold my hand with 6 fingers. ---> Hold my hand .
Hold my hand .
```

You, in the front row, are amazing. ---> You, , are amazing.
You, , are amazing.
I saw a monkey with a telescope. ---> I saw a monkey .
I saw a monkey .
I'm doing this with pleasure. ---> I'm doing this .
I'm doing this .
The guy with a knife threatened the shopkeeper. ---> The guy threatened the shopkeeper.
The guy threatened the shopkeeper.
Gus trash talks about his friends behind their backs. ---> Gus trash talks .
Gus trash talks .
The lady in the toilet is taking a lot of time. ---> The lady is taking a lot of time.
The lady is taking a lot of time.
A lady with a pillow hit a man in a wheelchair. ---> A lady hit a man .
A lady hit a man .
I was on time. ---> I was .
I was .
The car dealership told his customer with excitement to pay with coins behind the shop. ---> The car dealership told his customer to pay .
The car dealership told his customer to pay .
With all due respect, I disagree. ---> , I disagree.
, I disagree.
He shouted at his dog with a whistle. ---> He shouted .
He shouted .
We decided that we'll rent a car with enthusiasm. ---> We decided that we'll rent a car .
We decided that we'll rent a car .
Karen found her meaning in life in the shower. ---> Karen found her meaning .
Karen found her meaning .
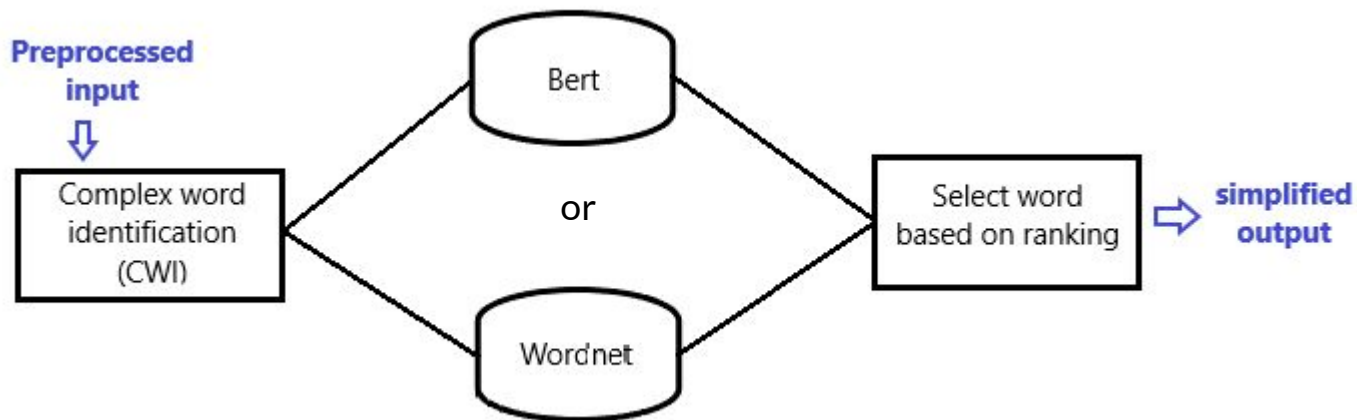Hold my hand with 6 fingers. ---> Hold my hand .
Hold my hand .

Further to-do with PP removal

- Search for in/mono/bi-transitivity to avoid unnecessary reduction (only clue: verbnet)
- Deal with PP ancestor problem (see stanza's parser)
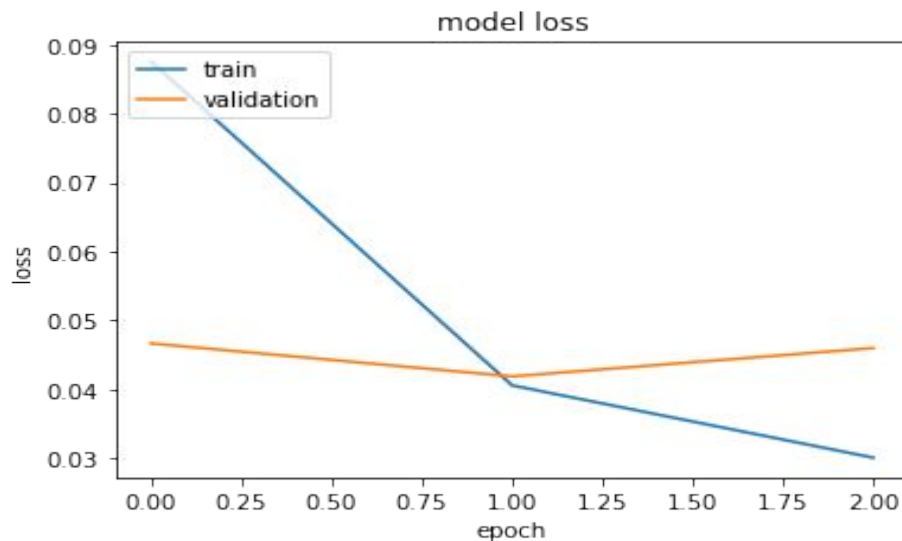- Assign importance to PPs and do removal based on them

Lexical simplification

Pipeline

CWI model

The model is trained but we are not sure how to use it for the next step

With Bert

Already existing system found

We can expect result of the genre:

Original text	Simplified text
The Risk That Students Could Arrive at School With the **Coronavirus** As schools **grapple** with how to **reopen**, new **estimates** show that large parts of the country would **probably** see infected students if **classrooms** opened now.	The Risk That Students Could Arrive at School With the **disease** As schools **deal** with how to **open** new **numbers** show that large parts of the country would **maybe** see infected students if **they** opened now.
How a **photograph** of a young man **cradling** his **dying** friend sent me on a **journey** across India.	How a **shot** of a young man **and** his **dead** friend sent me on a **tour** across India.

(https://medium.com/@armandj.olivares/how-to-use-bert-for-lexical-simplification-6edbf5a4d15e)

With WordNet

Using **WordNet** and the **word frequency (zipfreq):**

- We compare the frequency of the actual word and the candidate (set of lemmas of the hypernym):
    - If the candidate is more frequent we replace the actual word
    - Else we keep the actual word

With WordNet

Original sentence	Simplified sentence
I have a **corgi** and a **bike**	I have a **dog** and a **bike**

Interface Presentation

Future work :

- Continue on syntactic simplification
- Continue on Lexical simplification
- Perfect the interface

Thank you

Questions?