BehancePR: A Punctuation Restoration Dataset for Livestreaming Video Transcript

Anonymous ACL submission

Abstract

Given the increasing number of livestreaming videos, automatic speech recognition and post-processing for livestreaming video transcripts are crucial for efficient data management as well as knowledge mining. A key step in this process is punctuation restoration which restores fundamental text structures such as phrase and sentence boundaries from the video transcripts. This work presents a new human-annotated corpus, called BehancePR, for punctuation restoration in livestreaming video transcripts. Our experiments on BehancePR demonstrate the challenges of punctuation restoration for this domain. Furthermore, we show that popular natural language processing toolkits like Stanford Stanza, Spacy, and Trankit underperform on detecting sentence boundary on non-punctuated transcripts of livestreaming videos. The dataset will be made available to the public to foster research in this area.

1 Introduction

Livestreaming is a powerful broadcasting medium that catches the attention of millions of users. Many video-sharing platforms have supported livestreaming for a wide range of topics such as Twitch for gaming, TikTok for short entertainment videos, Behance for visual creative work, and Youtube/Facebook Live accepting any topics. Among these videos, there are a substantially high number of videos that provide useful knowledge with exceptional visual demonstration. To this end, livestreaming videos are becoming a potential knowledge base waiting for being explored.

Mining videos on video/audio format directly is extremely hard and expensive because of their high data load and complexity in processing images and audio signals. Instead, mining video transcripts, transcribed by either human or machine, is much easier with the existing hardware and software. As such, livestreaming videos should be transcribed at high quality to facilitate future data mining research. As video transcription can be done using existing automatic speech recognition (ASR) systems, a reasonable step to improve the quality of transcribed texts for livestreaming videos involves post-processing produces to remove noises and restore correct language structures and texts from ASR-generated texts.

In this paper, we are particularly interested in punctuation restoration (PR) for livestreaming video transcripts. Punctuation restoration is the task to restore fundamental text structures such as sentences and phrases by inserting punctuation marks into non-punctuated text, e.g. text generated by an automatic speech recognition system for livestreaming videos in our paper. Punctuation restoration is an important post-processing step to improve the readability of ASR texts. Moreover, in natural language processing (NLP), PR is even more important as it enables the use of advanced techniques to process texts at the sentence level to achieve optimal performance for various tasks, e.g., part-of-speech tagging and dependency parsing. Prior studies have shown that with proper sentence split and punctuation, a downstream application can tolerate the word error rate of 25% (Alam et al., 2015), which is extremely high compared to the current state-of-the-art ASR. Figure 1 demonstrates how punctuation restoration improves the readability of ASR-generated texts.

In the literature, punctuation restoration is considered as a subtask of ASR, in which PR annotation is done as part of the ASR datasets such as the AMI (McCowan et al., 2005) and TED corpus (Federico et al., 2012). However, the speeches recorded in these audios are multi-speaker meetings, as in AMI corpus, or single-speaker talks, as in TED corpus. Our work is different from this work in that we consider livestreaming videos that feature many distinctive characteristics that are essential to study. In particular, the number of speakers in

use the marquee tool to draw a selection around the empty space on one side then hold shift and add the other areas to the selection too go to edit and fill then change the drop down menu to content aware photoshop should automatically generate a completely new background but it might make a couple of small mistakes
Use the marquee tool to draw a selection around the empty space on one side. Then hold shift and add the other areas to the selection too. Go to edit and fill, then change the drop down menu to content aware. Photoshop should automatically generate a completely new background. But it might make a couple of small mistakes.

Figure 1: Punctuation restoration improve the readability of an ASR-generated text.

livestreaming videos varies greatly ranging from 1 to a few main speakers together with up to thousands of audiences. The audiences might participate in question-answering and commenting during the whole duration of the video, hence, changing the topic of the video. Furthermore, the speech in livestreaming is much more spontaneous than those in meetings in AMI corpus and the well-scripted TED talk.

An issue with the research of punctuation restoration for livestreaming videos is the lack of a human-annotated dataset for model development and evaluation. This is even more critical when livestreaming has become one of the most powerful communication mediums for not only entertainment but also education purpose. Toward this end, we introduce a new dataset for Behance Livestreaming Video Punctuation Restoration, called **BehancePR**. The dataset was annotated by skilled transcription annotators for 4 types of punctuation. Our experiments reveal the challenges of the BehancePR dataset that the performance of the existing state-of-the-art models for punctuation restoration on BehancePR lags far behind one on the TED dataset. Our further experiment of cross-domain punctuation shows that models that are trained on a PR dataset of a different speech scenario perform much worse than those trained on the BehancePR dataset even with a much larger training set.

2 Data Annotation

Preparation: The livestreaming videos that we annotate in this work are derived from `Behance.net`. Behance is an online platform to showcase and discover creative work such as digital drawing, graphic design, and photo/video editing. In those videos, one or a few creators stream their work on graphic design tools in English. The topics in the videos relate to design theories, graphical ideas, and tutorials to use those graphic design tools. The videos are split into shorter clips of 5 minutes per clip. Then, the videos are transcribed by the Microsoft ASR system. The video and the aligned automatic-generated transcript are presented to the annotators. To prepare for the annotation, we also created a taxonomy for livestreaming video punctuation restoration. Similar to prior studies in punctuation restoration, we inherit the taxonomy of three most popular markers: i.e. **period**, **comma**, and **question mark** (Federico et al., 2012). However, as livestreaming videos of creative works involve a lot of emotional expressions such as excitement, we include **exclamation mark** to better present strong feelings and emphasis.

Annotation: We recruit 8 annotators from `Upwork.com` crowdsourcing platform. As Upwork allows the freelancers to submit their resumes, we can choose the most experienced annotators with prior practices with audio transcribing. A detailed annotation guideline with many examples is provided to the annotators. We developed a customized web-based annotation tool that allows the annotator to work most efficiently with the material. To make sure the annotators understand the task and be familiar with the annotation tool, they are further trained on 2 hours of audio on the Behance videos, equivalent to approximately 6 hours of training and practice. We group 6 annotators into 3 pairs of annotators that work together in the same set of documents. The inter-annotation Cohen-Kappa agreement score is 0.59 , which indicates a moderate to substantial agreement. Toward the end, the annotators are allowed to discuss to make the final version of the dataset. Table 1 shows the detailed statistics and label distribution of the BehancePR dataset.

3 Dataset Challenges

Compare to existing punctuation restoration datasets, e.g., TED (Federico et al., 2012), AMI (McCowan et al., 2005), our dataset BehancePR features several unique challenges for punctuation restoration.

First, as its documents are obtained from live streaming video on the Behance platform, it features unique characteristics of spontaneous speech.

	Train	Dev	Test
Statistics			
#Documents	2,174	60	80
#Sentences	115,661	2,969	3,986
#Tokens	1,216,439	34,265	44,224
Label distribution			
#Periods	101,228	2,583	3,229
#Commas	126,739	3,291	4,388
#Questions	7,337	175	437
#Exclamations	7,096	211	320

Table 1: Statistics and label distribution of the BehancePR dataset.

They are much different from TED talks, in which the talks were heavily scripted, and AMI meetings, where the talks were also well prepared. As such, the live streaming video transcript text has a much lower cohesion. As it might be presented in a sudden change of topic and incomplete syntax. Besides, they come with a substantially high-frequency verbal pause, repetition of words and phrases, which are the results of hesitation and stutter of the speakers.

Second, as the documents in this dataset are generated from an automatic speech recognition system, it is expected that there is a certain number of word errors in the text even though the automatic speech recognition technology has improved significantly recently. In the worst case, word errors might affect punctuation prediction. Examples of these challenges are presented in the Appendix A.

Easy	***Owe my goodness****!* *I 'am* ***excited*** *for that!* *That's an* ***interesting*** *pick!* *Imagine it has this dynamics and* ***boom****!*
Hard	*It's in here!* *I found it!* *Let's jump into it!* *He moves!*

Table 2: Sentences with and without emotional words.

Third, we introduce the exclamation mark for emotional sentences. This exclamation mark is a brand new challenge compared to other datasets, e.g., the TED and AMI talks in which emotion is rare. To detect these sentences, a model needs not only the textual features but also acoustic features such as frequency, the strength of the excitement. As such, without access to the audio, a model might find this a challenging task. Especially, this becomes even harder in a substantial amount of cases that emotional words do not appear in the sentence. Table 2 shows some examples of easy and hard sentences with and without the appearance of emotional words, respectively.

4 Experiments

Supervised learning: To reveal the complexity of the punctuation restoration for the livestreaming video, we evaluate the performance of the state-of-the-art model for punctuation restoration. Similar to prior work in this task, we model the task as a sequence labeling task on the token level. We investigate two major model architectures: neural-based model with Bi-Directional Long Short-Term Memory (BiLSTM) (Alam et al., 2020), and graphical-based model with Conditional Random Field (CRF) (Makhija et al., 2019). We also investigate the recent advance in data augmentation technique for punctuation restoration (Alam et al., 2020). These leave us four combinations of model and technique as presented in table 3. All of these models leverage a pre-trained language model RoBERTa (Liu et al., 2019) to obtain representation vectors.

Table 3(a) presents the performance of four models on the development and the testing sets of the BehancePR dataset. First, the graphical-based models with CRF show a slight improvement over the neural-based models in case no data augmentation is applied. In contrast, the performance on the test set decreases. This result is consistent with the prior study in punctuation restoration (Alam et al., 2020) that performance gain is not significant when CRF is used, which is opposite to the Named Entity Recognition task that CRF works very well. We attribute this to the difference between the PR and the NER tasks. In punctuation restoration, an entity, which is a sentence, is much longer than an entity in the NER task. As such, it is harder to model the dependency between tokens in such a long sequence. Hence, the CRF is incapable of producing significant improvement. Second, the data augmentation technique slightly improves the performance of the BiLSTM model but slightly decreases the performance of the CRF-based models. Importantly, we find that the performance of the PR models on BehancePR is far below on the TED talk dataset (F-score=84%). This suggests this dataset is much more challenging and calls for further study on this domain.

Domain adaptation: As punctuation restoration data is abundant, we further explore the cross-

Model	Dev			Test		
	P	R	F	P	R	F
(a) Behance → Behance						
BiLSTM	63.6	63.1	63.4	62.0	61.4	61.7
+aug	**64.8**	62.2	63.5	**63.8**	60.7	62.2
+CRF	62.8	**65.2**	**64.0**	62.2	**63.5**	**62.9**
+CRF+aug	62.8	64.5	63.7	61.1	62.8	62.0
(b) TED → Behance						
BiLSTM	53.5	59.1	56.2	54.6	59.6	57.0
+aug	55.8	58.0	**56.9**	55.7	58.5	57.1
+CRF	52.7	60.4	56.3	53.2	60.3	56.5
+CRF+aug	56.5	57.3	**56.9**	57.0	57.8	**57.4**

Table 3: Performance of the examined models on the Behance datatset.

Model	P	R	F
Stanza	70.4	1.4	2.8
Trankit	72.1	7.8	14.0
SpaCy	52.1	21.9	30.9
BiLSTM	70.3	75.6	72.8
+aug	71.5	72.8	72.1
+CRF	73.3	72.0	72.6
+CRF+aug	71.6	73.0	72.3

Table 4: Performance of sentence split.

domain evaluation setting where the models are trained on a different source domain and evaluated on the BehancePR dataset. In particular, we choose the TED corpus as the source domain because TED talks are monologues, which is the closest to the Behance videos. Table 3(b) presents the out-of-domain performance of the models. It is clear from the table that the performances of all punctuation restoration models degrade significantly when they are trained on TED talk and evaluated on Behance. This demonstrates the considerable difference in domains even they are almost monologual.

Sentence split: We evaluate the performance of the models on the sentence splitting task. In this task, the model is trained and evaluated to predict where the sentence ends. We trained the models presented in the supervised learning setting on the same BehancePR dataset. We further examine the performance of existing NLP toolkit for this task including Stanza (Qi et al., 2020), SpaCy (Honnibal and Montani, 2017), and Trankit (Nguyen et al., 2021). The performance of the models and toolkits are presented in Table 4.

The published toolkits perform very poorly on this kind of text, with the highest F-1 score of 30.9% by SpaCy due to their low recall performances. Stanza's sentence splitter is the worst with an extremely low recall of 1.4%, Trankit and Spacy are slightly better with higher recalls of 7.8% and 21.9%, respectively. The reason is that these models were trained on punctuated text (Nivre et al., 2016), hence, the model is heavily dependent on the presence of punctuation to trigger the sentence boundary prediction. In another word, they barely learn the context to predict the sentence boundary. Secondly, the models trained on the BehancePR dataset outperform the published toolkits with large margins at consistent F-1 scores above 72%. Without the punctuation in the text, these models learn the contextual information to predict the sentence ending, hence, successfully performing the task in this extreme case. This suggests the importance of the training data for even fundamental tasks like sentence splitting.

5 Related work

Early studies on the punctuation restoration task explored a wide range of features such as lexical, acoustic, prosodic, and their combination (Gravano et al., 2009; Levy et al., 2012; Xu et al., 2014; Che et al., 2016a; Szaszák and Tündik, 2019). Graphical models such as conditional random field have been widely used for this task (Lu and Ng, 2010; Zhang et al., 2013) before the emerging of neural networks. Recently, a variety of deep neural network architectures have been explored such as long short-term memory (Gale and Parthasarathy, 2017), convolutional neural network (Che et al., 2016b), and transformer (Alam et al., 2020). Corpora for punctuation restoration are usually created as part of automatic speech recognition in various domains such as meetings (McCowan et al., 2005), TED talks (Federico et al., 2012), audio books (Panayotov et al., 2015), and film subtitles (Tiedemann, 2016). Among these, the TED corpus is widely used as the benchmark corpus for punctuation restoration. However, livestreaming videos have not been explored for this task.

6 Conclusion

We presented BehancePR, the first dedicated corpus for punctuation restoration. BehancePR was manually annotated for 4 markers. The comprehensive experiments with state-of-the-art models show the challenges of the punctuation restoration for livestreaming video, as well as the poor performance of existing NLP toolkits for non-punctuated text.

7 Ethical Considerations

In this work we present a dataset on the transcripts of a publicly accessible video-streaming platform, i.e., "*Behance*"[1]. Complying with the discussion presented by Benton et al. (2017), research with human subjects information is exempted from the required full Institutional Review Board (IRB) review if the data is already available from public sources or if the identity of the subjects cannot be recovered. However, to protect the identity of the streamer and any other person whose information are shared in the video transcript, we impose extra consideration on the presented dataset. First, in this dataset, we exclude the username or any other identity-related information of the streamer in the transcripts to prevent disclosing their identity. Moreover, the proposed dataset only provides textual data (i.e., paragraphs), hence the other content of the videos (e.g., images, audios) are not revealed to protect human identity. Finally, to reduce the risk of disclosing the information of the people in the transcripts, in the final version of the dataset, we exclude the transcripts that explicitly or implicitly refer to the identify of the target people.

References

Firoj Alam, Bernardo Magnini, and Roberto Zanoli. 2015. Comparing named entity recognition on transcriptions and written texts. In *Harmonization and Development of Resources and Tools for Italian Natural Language Processing within the PARLI Project*, pages 71–89. Springer.

Tanvirul Alam, Akib Khan, and Firoj Alam. 2020. Punctuation restoration using transformer models for high- and low-resource languages. In *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pages 132–142, Online. Association for Computational Linguistics.

Adrian Benton, Glen Coppersmith, and Mark Dredze. 2017. Ethical research protocols for social media health research. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pages 94–102, Valencia, Spain. Association for Computational Linguistics.

Xiaoyin Che, Sheng Luo, Haojin Yang, and Christoph Meinel. 2016a. Sentence boundary detection based on parallel lexical and acoustic models. In *Interspeech*, pages 2528–2532.

Xiaoyin Che, Cheng Wang, Haojin Yang, and Christoph Meinel. 2016b. Punctuation prediction for unsegmented transcript based on word vector. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 654–658.

Marcello Federico, Sebastian Stüker, Luisa Bentivogli, Michael Paul, Mauro Cettolo, Teresa Herrmann, Jan Niehues, and Giovanni Moretti. 2012. The iwslt 2011 evaluation campaign on automatic talk translation. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey. European Language Resources Association (ELRA).

William Gale and Sarangarajan Parthasarathy. 2017. Experiments in character-level neural network models for punctuation. In *INTERSPEECH*, pages 2794–2798.

Agustin Gravano, Martin Jansche, and Michiel Bacchiani. 2009. Restoring punctuation and capitalization in transcribed speech. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 4741–4744. IEEE.

Matthew Honnibal and Ines Montani. 2017. spacy 2: Natural language understanding with bloom embeddings. *convolutional neural networks and incremental parsing*, 7(1).

Tal Levy, Vered Silber-Varod, and Ami Moyal. 2012. The effect of pitch, intensity and pause duration in punctuation detection. In *2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel*, pages 1–4. IEEE.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Wei Lu and Hwee Tou Ng. 2010. Better punctuation prediction with dynamic conditional random fields. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 177–186.

Karan Makhija, Thi-Nga Ho, and Eng-Siong Chng. 2019. Transfer learning for punctuation prediction. In *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pages 268–273.

Iain McCowan, Jean Carletta, Wessel Kraaij, Simone Ashby, S Bourban, M Flynn, M Guillemot, Thomas Hain, J Kadlec, Vasilis Karaiskos, et al. 2005. The ami meeting corpus. In *Proceedings of the 5th international conference on methods and techniques in behavioral research*, volume 88, page 100. Citeseer.

Minh Van Nguyen, Viet Dac Lai, Amir Pouran Ben Veyseh, and Thien Huu Nguyen. 2021. Trankit: A light-weight transformer-based toolkit for multilingual natural language processing. In *Proceedings of the 16th*

[1] www.behance.net

Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations, pages 80–90, Online. Association for Computational Linguistics.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty, and Daniel Zeman. 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, Portorož, Slovenia. European Language Resources Association (ELRA).

Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. 2015. Librispeech: an asr corpus based on public domain audio books. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pages 5206–5210. IEEE.

Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: A python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 101–108, Online. Association for Computational Linguistics.

György Szaszák and Máté Akos Tündik. 2019. Leveraging a character, word and prosody triplet for an asr error robust and agglutination friendly punctuation approach. In *INTERSPEECH*, pages 2988–2992.

Jörg Tiedemann. 2016. Finding alternative translations in a large corpus of movie subtitle. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3518–3522.

Chenglin Xu, Lei Xie, Guangpu Huang, Xiong Xiao, Eng Siong Chng, and Haizhou Li. 2014. A deep neural network approach for sentence boundary detection in broadcast news. In *Fifteenth annual conference of the international speech communication association*. Citeseer.

Dongdong Zhang, Shuangzhi Wu, Nan Yang, and Mu Li. 2013. Punctuation prediction with transition-based parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 752–760.

A Noisy texts in BehancePR

Table 5 shows the examples of different types of noisy texts including verbal pauses, duplicate words and phrases, incomplete syntax, instructional steps, and word errors as presented in Section 3.

In the word error example, the streamer has just hurt herself, the ASR system was incorrect at detecting the words "*Oww*". Instead, it generated "*Oh*" and "*How*". This error is highly adverse as it might turn a declarative sentence into a WH question starting with the word "*How*".

Verbal pause
***So**, this is what we got for the site map.* *We talked about six of these being virus killers maps triall assets forum another* ***So**, those will be that.* ***So**, then will also have a footer.* ***Um,** so this will be the home page going to,* ***um**, start grab some assets...*
Duplicate words and phrases
Alright**, **alright**, **alright *So, but **there are all set*** *All of **these are all set***
Incomplete syntax
*Um so this will be the homepage **going to hum** start grab some assets I guess image that **google "survivor"***
Instructional steps
What if I click it open *2 xbox hub* ***Inspect*** ***Header*** *1920 by 1080* ***Copy*** *I got it*
Word error
***Oh** definitely just stubbed my toe.* *And in not very fun pain.* ***How** just making sure there will be no bleeding.*

Table 5: Examples of noisy texts in transcripts of livestreaming videos. Noisy words are highlighted.

B Reproducibility Checklist

Dataset: The statistics of the created dataset BehancePR (including training/development/test portions) and the annotation process are presented in Section 2. We will publicly release the dataset upon the acceptance of the paper. A URL to the publishing site will be included in the paper.

Source code with the specification of all dependencies, including external libraries: We will publicly release the code to run the models upon the acceptance of the paper.

Description of computing infrastructure used: All the experiments were run on a machine with 2 Intel Xeon E5-2620 v4 CPUs, 128GB of RAM, and 4 NVIDIA RTX 2080 Ti GPUs with 11GB RAM. We only train the models with a single GPU. The amount of GPU memory is from 8GB-10GB depending on the variant being used.

Average runtime for each approach: We train the models for 20 epochs; the whole training takes 2 hours. The best epoch is chosen based on the performance on the development set.

Number of parameters in the model: Our models have total 360M trainable parameters.

Explanation of evaluation metrics used, with links to code: Following prior work in punctuation restoration, we use the precision, recall, and F1 scores for performance metrics.

Hyperparameter bounds and configurations for best-performing models: We found that RoBERTa large version delivered the best performance among RoBERTA, AlBERT, and bert-large-uncased version of BERT. This confirms prior result by Alam et al. (2020). The texts are split into sequences of 256 word pieces, searched in the range of $\{64, 128, 256, 384\}$. The best batch size is 64 searched from $\{16, 32, 64\}$. We use a single BiLSTM layer to encode the text. The augmentation rate is set to 0.2 following previous research's result (Alam et al., 2020).

C Annotation Tool

We developed a customized web-based annotation tool for this work. The annotation tool focuses on improving the readability of the annotated text, as the result, improves the annotation quality. Toward this end, we use color coding for punctuation markers. More importantly, whenever a sentence ending marker is assigned, such as period, question mark, and exclamation, it automatically creates a new line to separate sentences. Figure 2 shows the interface and the annotated text using our tool.

D Annotation guidelines

This section summarizes the taxonomy, annotation guideline, and annotation examples. Examples are

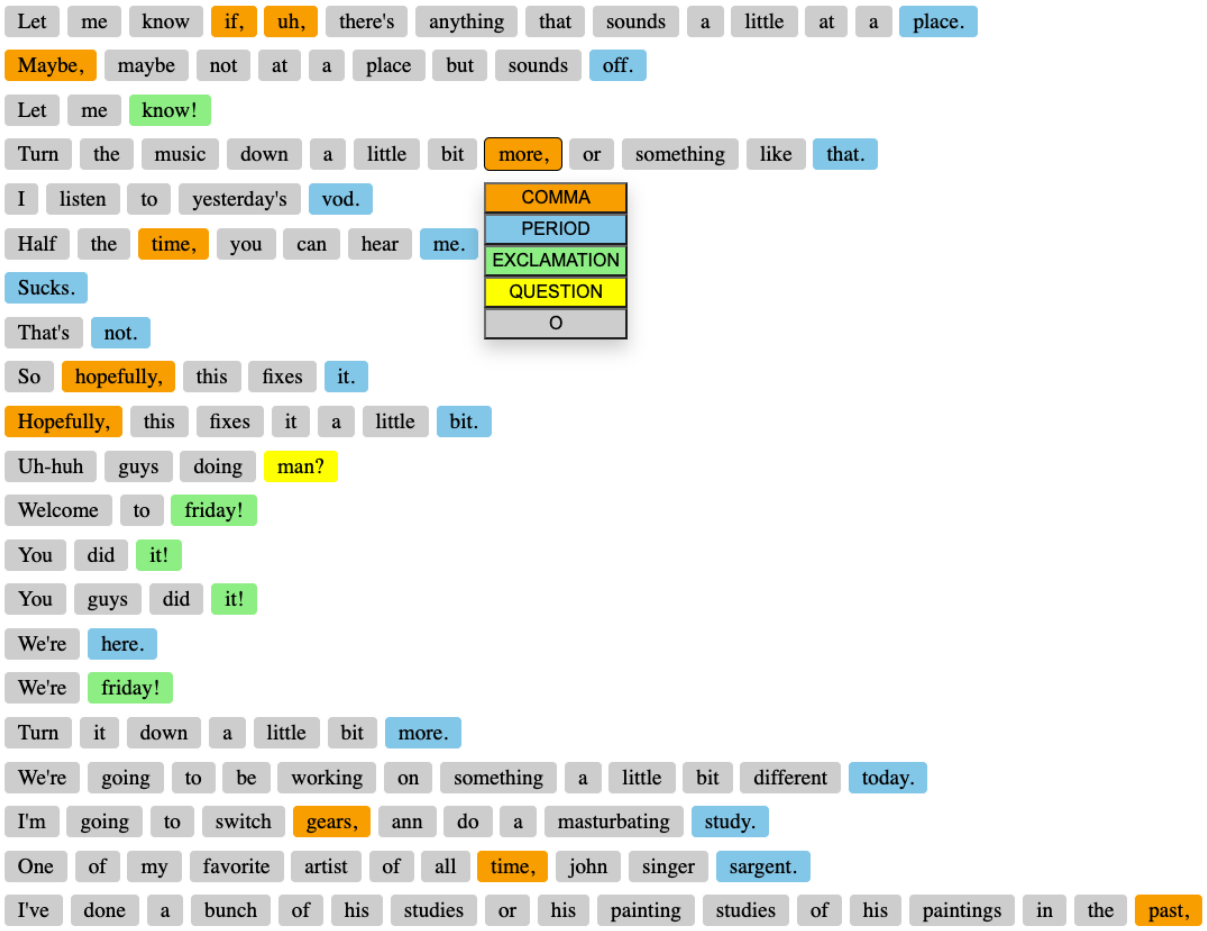

Figure 2: The colorful, easy-to-use interface of our annotation tool designed for the annotation of BehancePR. The color codes for comma, period, exclamation mark, and question mark are orange, sky blue, light green, and yellow, respectively.

shown in figure 2.

A **period** is used for:

- Marking the end of a **declarative sentence**.
- Separating independent clauses **without** a conjunction when a semi-colon is usually used (to distinguish with the case that a comma is used when a conjunction presents).

A **question mark** is used for:

- Marking the end of a **question**.

An **exclamation mark** is used for:

- Exclaiming something. They are commonly used after interjections (words or phrases that are used to exclaim, command, or protest like "wow" or "oh").
- Express the following emotions: excitement, surprise, astonishment, emphasizing a point, and other types of strong emotions.

A **comma** is used for:

- Separating independent clauses when they are joined by any of these seven coordinating conjunctions: and, but, for, or, nor, so, yet.
- Separating introductory clauses, phrases, or words from the main clause.
- Setting off clauses, phrases, and words that are not essential to the meaning of the sentence. Use one comma before to indicate the beginning of the pause and one at the end to indicate the end of the pause.
- Separating three or more words, phrases, or clauses written in a series.
- Separating two or more coordinate adjectives that describe the same noun. Be sure never to add an extra comma between the final adjective and the noun itself or to use commas with non-coordinated adjectives.

- Separating contrasted coordinate elements or to indicate a distinct pause or shift near the end of a sentence.
- Setting off phrases at the end of the sentence that refers back to the beginning or middle of the sentence. Such phrases are free modifiers that can be placed anywhere in the sentence without causing confusion.
- Setting off all geographical names, items in dates (except the month and day), addresses (except the street number and name), and titles in names.
- Shifting between the main discourse and a quotation.
- Preventing possible confusion or misreading.