

# The Many Voices of the Detached: Revisiting the Disputed Writings of *Lu Xun* and *Zhou Zuoren*

Haining Wang, Information & Library Science, IUB

## Research Questions

We attribute authorship of a collection of twenty essays that are disputed between two famous Chinese writers, *Lu Xun* (鲁迅) and *Zhou Zuoren* (周作人). The essays were written in classical Chinese and published under the pseudonyms *Du Ying* (独应) and *Du* (独), meaning “the Detached.”

- **RQ1: Who is the identity behind the pseudonyms?**
- **RQ2: Did the authors collaborate with each other?**

## Backgrounds

*Lu Xun* (or *Zhou Shuren*, 周树人, 1881–1936) and *Zhou Zuoren* (1885–1967) were brothers and leading figures of modern Chinese literature who greatly influenced the culture of East Asia. They studied abroad in Japan during the 1910s, where they began their writing careers.

Scholars have different opinions on the authorship of twenty essays published by the brothers between 1907 and 1912. The essays are signed with two pseudonyms, *Du Ying* and *Du*. Knowing the authorship is of great value for better understanding their intellectual trajectory.



Figure: *Zhou Zuoren* (left) and *Lu Xun* (right). Source: sinaimg.cn.

## Authorship Attribution

*Authorship attribution* infers the likely author of a disputed work based on their writing style. Two key assumptions of this method are:

- The candidate set of potential authors is known, e.g., two possible authors in our case.
- The disputed text is single-authored, i.e., there is no collaboration or significant editorial moderation present.

However, our case does not align well with the second assumption, and we hypothesize that the brothers collaborated on some long prose. To address this, we chunked down the long prose and treated them as if they were single-authored pieces. A well-performed classifier will yield different predictions with high confidence if the authors collaborated substantially.

## Highlights

- Both *Lu Xun* and *Zhou Zuoren* published using the pseudonyms *Du Ying* and *Du*.
- The brothers collaborated closely on *The Strings of Melancholy*.

## Corpus

We first collected a set of twenty essays published under the pseudonyms *Du Ying* and *Du*, which are considered a *test set* due to the uncertain authorship and therefore of interest. We then collected and proofread a set of essays known to be written by *Lu Xun* and *Zhou Zuoren*. These texts are written using classical non-fictional prose and were published roughly at the same time as the disputed works.

We split the non-disputed prose into *training* and *validation sets* in a way that the topics in two sets are different. This topic distinction was made to avoid classifiers from overfitting to thematic signals instead of stylistic cues.

## Feature Set

We transcribed two dictionaries of function words, both in classical and modern Chinese, and removed less distinguishing function characters and words using recursive feature elimination. The resulting list retains only 33 features, which includes 29 unigrams and 4 bigrams. The features are listed in Table 1.

**Table: Features and their coefficients.** All coefficients are significant ( $p \ll 1e-4$ ). Positive (reddish) features favor *Lu Xun* (encoded as 1 during the training).

Feat.	Coef.	Feat.	Coef.	Feat.	Coef.	Feat.	Coef.	Feat.	Coef.
Const.	0.53	随	-0.20	多	-0.24	是	0.23	但	-0.20
不	0.21	何	0.25	于	0.21	而	0.29	光	0.19
徒	0.24	原	-0.25	皆	-0.19	乃	0.22	进而	0.21
诚	0.34	别	-0.27	本	-0.34	则	0.19	自然	-0.19
焉	0.22	及	-0.32	故	-0.20	切	0.19	足以	0.23
夫	0.18	矣	0.26	惟	0.29	必	0.24	于是	0.26
之	0.38	全	0.20	各	-0.26	为	-0.22		

## Model

We chose a logistic regression model with l2 penalty ( $C = 1.0$ ) for simplicity and interpretability, as shown in Equation 1:

$$\hat{p}(X_i) = \text{sigmoid}(X_i w + w_0) \quad (1)$$

The parsimonious model achieves 93.3% accuracy on the validation set. Validation accuracy should give us some rough sense of how well the classifier performs. Note that we normalized the raw features in each sample by document length and standardized each feature to have zero mean and unit variance. During explanation, we dealt with the relative frequency of the features instead of the raw count.

## Interpretation

The following excerpt is the opening of *The Strings of Melancholy* (哀弦篇). We have marked function characters based on their signs and weights, with reddish characters indicating *Lu Xun*'s likely authorship. The text reads:

华土物色之黯淡也久矣。民德离散，质悴神亏，旧泽弗存，新声绝朕。处今日之世，虽步康庄之涂以临观市集，士女熙熙，盈吾左右，顾目击扰攘，而萧条之感乃不觉婴心而来，令人森然如过落日废墟或无神之寒庙者，其凄清也如是，盖所谓死寂者是也。萧条唯何？无觉悟是。曷无觉悟？无悲哀故。人唯不知自悲，而后零落所底，将更令他人悲之。盖哀弦断响而人心永寂，有如此也...

Much more reddish function characters reveal the author of this paragraph is likely *Lu Xun*, with a confidence of 0.98, to be exact.

## Findings

Our research provides evidence that the brothers shared two pseudonyms, *Du Ying* and *Du*. *Lu Xun* used these pseudonyms in *The Chinese Patriotism* (中国人之爱国) and four other essays, and *Zhou Zuoren* used them in *The Ordinary Folks' Responsibility* (庸众之责任) and six other works. We have demonstrated, for the first time, that the brothers collaborated closely on *The Strings of Melancholy*. Due to relatively low prediction confidence, we have withheld decisions on the remaining works. This study provides insight into the early stages of the authors' thinking trajectory.

## Corpus & Code

- [codeberg.org/haining/the\\_many\\_voices](https://codeberg.org/haining/the_many_voices)
- [hw56@indiana.edu](mailto:hw56@indiana.edu)