

kǒuhú

口胡

v0.2.0

2024-12-27

MIT

Chinese lipsum text generator; 中文乱数假文 (Lorem Ipsum) 生成器

Shengqi Chen <harry-chen@outlook.com>

<https://github.com/Harry-Chen/kouhu>

kouhu (口胡 in Chinese) is typst package aimed at generating Chinese lipsum text (a.k.a. blind text). It provides a set of built-in text samples containing both Simplified and Traditional Chinese characters. You can choose from generated fake text, classic or modern Chinese literature, or specify your own text.

The idea of kouhu is inspired by **zhlipsum** LaTeX package and **roremu** Typst package. All builtin text samples are excerpted from zhlipsum (w/o non-UTF-8 paragraphs). Please refer to [its documentation](#) for their detailed description.

1. **simp**: 劳仑衣普桑，认至将指点效则机，最你更枝。想极整月正进好志次回总般，段然取向使张规军证回，世市总李率英茄持伴。用阶千样响领交出，器程办管据家元
2. **trad**: 勞侖衣普桑，認至將指點效則機，最你更枝。想極整月正進好志次回總般，段然取向使張規軍證回，世市總李率英茄持伴。用階千樣響領交出，器程辦管據家元
3. **nanshanjing**: 南山經之首曰騫山。其首曰招搖之山，臨于西海之上，多桂，多金玉。有草焉，其狀如韭而青花，其名曰祝餘，食之不飢。有木焉，其狀如穀而黑理，其花四照
4. **xiangyu**: 項籍者，下相人也，字羽。初起時，年二十四。其季父項梁，梁父即楚將項燕，為秦將王翦所戮者也。項氏世世為楚將，封於項，故姓項氏。
項籍者，下相人也
5. **zhufu**: 旧历的年底毕竟最像年底，村镇上不必说，就在天空中也显出将到新年的气象来。灰白色的沉重的晚云中间时时发出闪光，接着一声钝响，是送灶的爆竹；近处
6. **aspirin**: 阿司匹林（英语：Aspirin），也称乙酰水杨酸（英语：acetylsalicylic*acid），是水杨酸类药物，通常用作止痛剂、解热药和

Part I.

Usage

`#builtin-text-list()` → `dictionary`

Get all builtin texts. Returns in `dictionary[string, list[string]]`.

Example: Show the number of paragraphs in each builtin text.

```
#builtin-text-list().pairs().map(p => (p.at(0), p.at(1).len()))  
  
(  
  ("simp", 50),  
  ("trad", 50),  
  ("nanshanjing", 43),  
  ("xiangyu", 43),  
  ("zhufu", 110),  
  ("aspirin", 66),  
)
```

```
#kouhu(  
  <builtin-text>: "simp",  
  <custom-text>: none,  
  <offset>: 0,  
  <indices>: none,  
  <length>: 0,  
  <before-para>: none,  
  <after-para>: none,  
  <between-para>: "parbreak()  
) → doc
```

Output selected Chinese lorem ipsum paragraphs.

Example 1: Cut the first several characters of a paragraph from Lu Xun's *Zhufu* (《祝福》) .

```
#kouhu(builtin-text: "zhufu", indices: 18, length: 99)
```

我因为常见些但愿不如所料，以为未毕竟如所料的事，却每每恰如所料的起来，所以很恐怕这事也一律。果然，特别的情形开始了。傍晚，我竟听到有些人聚在内室里谈话，仿佛议论甚么事似的，但不一会，说话声也就止了，

Example 2: Cut multiple paragraphs from *Zhufu*.

```
#kouhu(builtin-text: "zhuifu", offset: 5, indices: (2, 18), length: 31,
between-para: "—")
```

我万料不到她却说出这样的话来，诧异的站着。——“祥林嫂？怎么了？”

Example 3: Repeat some text until the specified length.

```
#kouhu(custom-text: ("奥利",), between-para: none, length: 31)
```

奥利

—Argument—

⟨builtin-text⟩: "simp"

str

Name of the builtin text, see [#builtin-text-list\(\)](#) for a full list and length.

—Argument—

⟨custom-text⟩: none

array

Custom text to use. If not none, builtin-text will be ignored.

—Argument—

⟨offset⟩: 0

int

Offset of the paragraph to start from.

—Argument—

⟨indices⟩: none

int | array

Indices (**NOT RANGE**) of paragraphs to use (offset will be added). A single integer means a single paragraph, an array means multiple paragraphs, and none means all paragraphs. Any out-of-bound index will be ignored.

—Argument—

⟨length⟩: 0

int

Length of grapheme (characters) to print. kouhu will repeat over selected paragraphs until length is reached, and the final paragraph will likely be truncated. 0 for unlimited, i.e. print all selected paragraphs for only once.

—Argument—

⟨before-para⟩: none

content

Content inserted before each paragraph.

—Argument—

⟨after-para⟩: none

content

Content inserted after each paragraph.

Argument

⟨between-para⟩: "parbreak()"

content

Content inserted between two paragraphs (has no effect if only one paragraph is selected).

Part II.

GitHub README

II.1. kouhu (口胡)

`kouhu` (口胡) is a Chinese lipsum text generator for [Typst](#). It provides a set of built-in text samples containing both Simplified and Traditional Chinese characters. You can choose from generated fake text, classic or modern Chinese literature, or specify your own text.

`kouhu` is inspired by [zhlipsum](#) LaTeX package and [roremu](#) Typst package.

All [sample text](#) is excerpted from `zhlipsum` LaTeX package (see Appendix for details).

II.1.1. Usage

```
#import "@preview/kouhu:0.1.0": kouhu

#kouhu(indices: range(1, 3)) // select the first 3 paragraphs from default text

#kouhu(builtin-text: "zhufu", offset: 5, length: 100) // select 100 characters
from the 5th paragraph of "zhufu" text

#kouhu(custom-text: ("Foo", "Bar")) // provide your own text
```

See [manual](#) for more details.

II.1.2. What does `kouhu` mean?

GitHub Copilot says:

`kouhu` (口胡) is a Chinese term for reading aloud without understanding the meaning. It is often used in the context of learning Chinese language or reciting Chinese literature.

II.1.3. Changelog

II.1.3.1. 0.2.0

- Rename misspelled parameter `indicies` to `indices`, and supports `int` as input (see: [#2](#)).
- Include README on GitHub repository in the documentation.

II.1.3.2. 0.1.1

- Fix some wrong paths in `README.md`.
- Fix generation of `indicies` when not specified by user.
- Add repetition of text until `length` is reached.

II.1.3.3. 0.1.0

- Initial release.

II.1.4. Appendix

II.1.4.1. Generating zhlipsum.json

First download the zhlipsum-text.dtx from [CTAN](#) or from local TeX Live (kpsewhich zhlipsum-text.dtx). Then run:

```
python3 utils/generate_zhlipsum.py /path/to/zhlipsum-text.dtx src/  
zhlipsum.json
```