



链滴

# Caps Writer 发布：按住大写锁定键，进行语音识别输入

作者：[HaujetZhao](#)

原文链接：<https://ld246.com/article/1594371212477>

来源网站：[链滴](#)

许可协议：[署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

# Caps Writer

## 简介

一款语音输入工具，下载运行后，把它最小化，按下大写锁定键超过 0.3 秒后，开始语音识别，松开键之后，自动输入识别文字。

## 地址

[Github](#)

## 描述

本工具 (Caps Writer) 是一个电脑端的语音输入工具，使用了阿里云的一句话识别 api

(有兴趣的可以改成百度、腾讯、讯飞的 api 试试)。

使用方法很简单：运行程序后，按下 **Caps Lock** (也就是大写锁定键) 超过 0.3 秒后，就会开始用阿云的 api 进行语音识别，松开按键后，会将识别结果自动输入。(不用担心会启用大写锁定，在松开时候，大写锁定状态会自动恢复)

因为使用了阿云的 api，所以需要用户自己到阿里云申请，再填到 **token.ini** 中才能正常使用。

对于聊天时候进行快捷输入、写代码时快速加入中文注释非常的方便。

## 开发动机

我就气抖冷，为什么直到现在，仍然没有开发者做过一个好用的语音输入工具？

有人建议用搜狗输入法、讯飞输入法的语音输入，但这几个方面是真让人受不了：

- 广告太多，拒绝安装
- 我主力五笔，不使用搜狗输入法、讯飞输入法，顶多临时用下微软拼音
- 就以搜狗输入法为例，它的语音输入快捷键只能是 **Ctrl + Shift + A/B/C.....**，有以下槽点：
  - 这个快捷键会和许多软件的快捷键冲突，且不好记
  - 打字时，按这样三个快捷键，手指很别扭，不爽
  - 它的逻辑是按下快捷键后，启用语音输入，你一停顿一下，要说下一名，语音输入却结束了，能让用户决定什么时候结束语音输入。

为了在电脑上语音输入，我之前是用的 Quicker 的手机端进行语音识别，输入到电脑上，需要两个设，非常麻烦。今天终于做好我心中最好用的电脑端语音输入工具了！

## 视频演示

[Caps Writer 使用教程](#)

<https://b3logfile.com/file/2020/07/VID20200710150850new-19dfadfe.mp4>

## 开箱即用

小白用户，只需要在 [Release](#) 界面下载打包好的 exe 文件，运行，会在同级目录生成一个 `token.ini` 文件，在 `token.ini` 中填入你阿里云拥有 **管理智能语音交互 (NLS)** 权限的 **RAM访问控制** 用户的 **Accesskey Id**、**Accesskey Secret** 和智能语音交互语音识别项目的 **appkey**，就可以正常使用了。详细步骤之后我录一个视频

## 安装使用

本工具是一个python脚本，上面小白下载的 Release 其实是用 pyinstaller 导出的 exe 文件，如果你在源码基础上使用，就需要安装以下模块：

- keyboard
- pyaudio
- configparser
- aliyunsdkcore
- alibabacloud-nls-python-sdk

其中：

- pyaudio 在 windows 上不是太好安装，可以先到 [这个链接](#) 下载 pyaudio 对应版本的 whl 文件再用 pip 安装
- alibabacloud-nls-python-sdk 不是通过 python 安装，而是通过 [阿里云官方文档的方法](#) 进行安装。

另外，需要在 `run.py` 中填入阿里云拥有 **管理智能语音交互 (NLS)** 权限的 **RAM访问控制** 用户的 **accessID**、**accessKey** 和智能语音交互语音识别项目的 **appkey**。

做完以上步骤后，只要运行 `run.py` 就可以用了！

本文件夹内有一个 [安装指南](#) 文件夹，在里面可以找到详细的安装指南，还包括了提前下载的 alibabacloud-nls-python-sdk 和 pyaudio 的 whl 文件。

## 后话

因为作者就是本着凑合能用就可以了的心态做这个工具的，所以图形界面什么的也没做，整个工具单就一个脚本，功能也就一个，按住大写锁定键开始语音识别，松开后输入结果。目前作者本人已经很满意。

欢迎有想法有能力的人将这个工具加以改进，比如加入讯飞、腾讯、百度的语音识别api，长按0.3秒开始识别时加一个提示等等等等。