

# 实验一：Praat 使用及语音信号处理算法基础

# 目 录

## Content

- 01** Praat 基本操作讲解
- 02** 任务一：Praat 基本操作实战
- 03** 任务二：探求声学参数与情感的关系
- 04** 任务三：主观评测实验
- 05** 实验结果评比与讲解

## Praat 的安装

### □ Windows

- ✓ 请下载并安装 <http://10.103.10.112:8000/Praat.exe>

### □ Mac

- ✓ 请下载并安装 [http://10.103.10.112:8000/praat6122\\_mac64.dmg](http://10.103.10.112:8000/praat6122_mac64.dmg)

### □ Ubuntu

- ✓ 请从终端依次执行下述指令

`sudo apt-get update`

`sudo apt-get install praat`

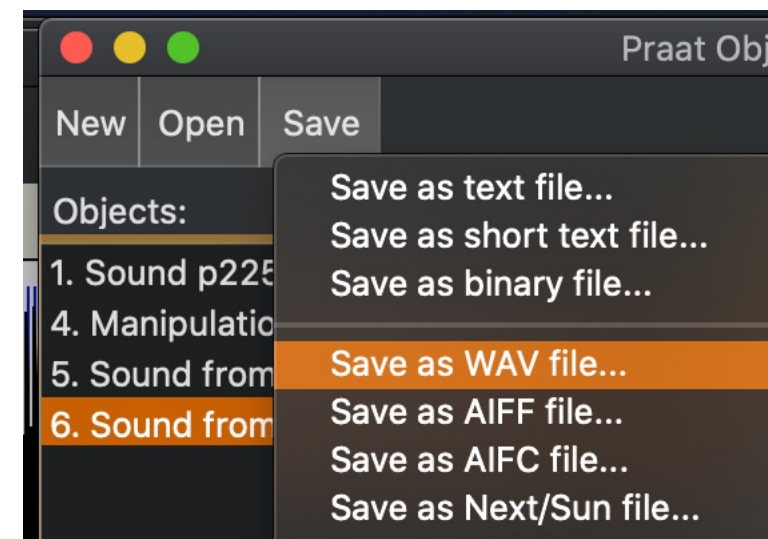
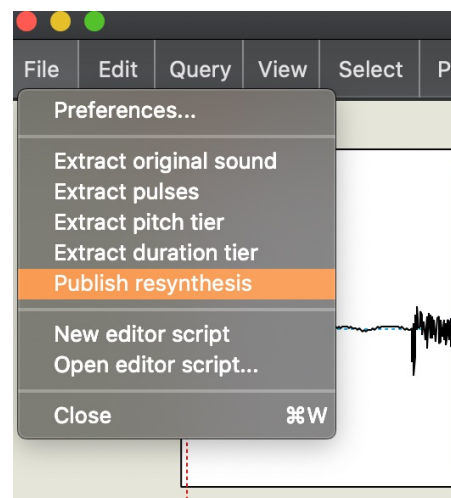
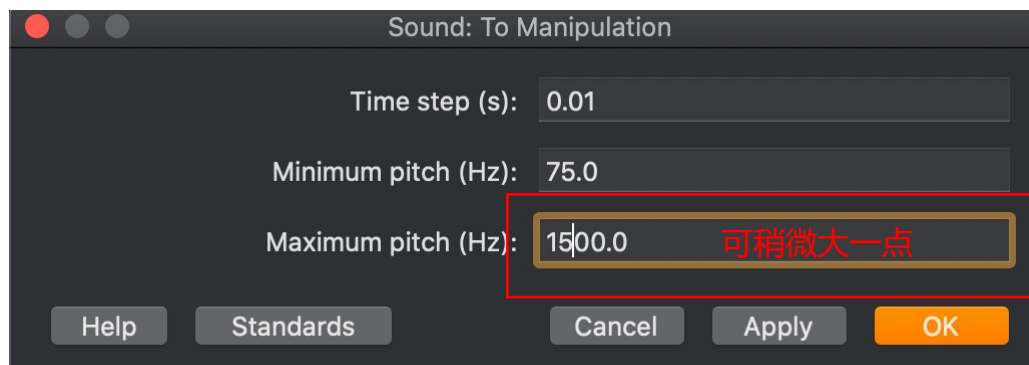
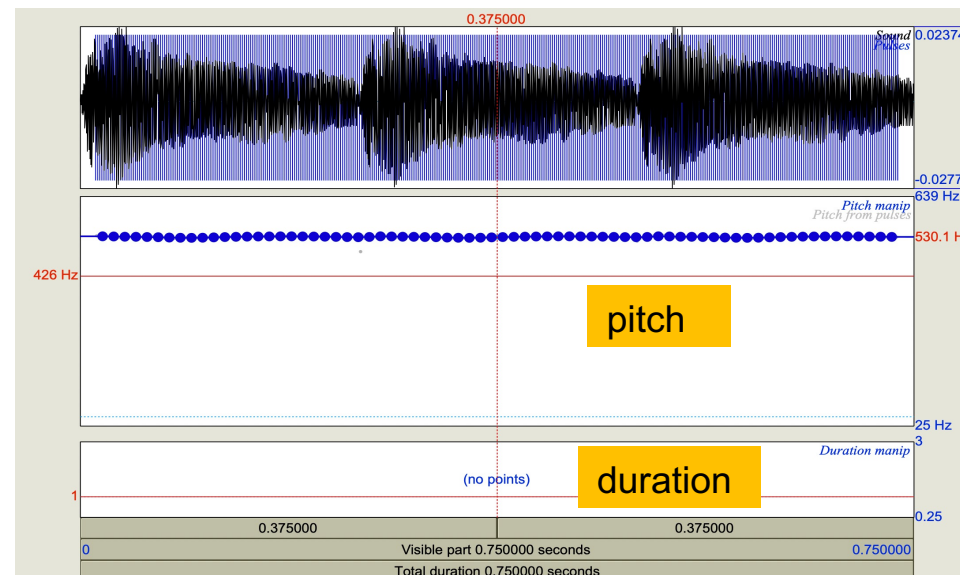
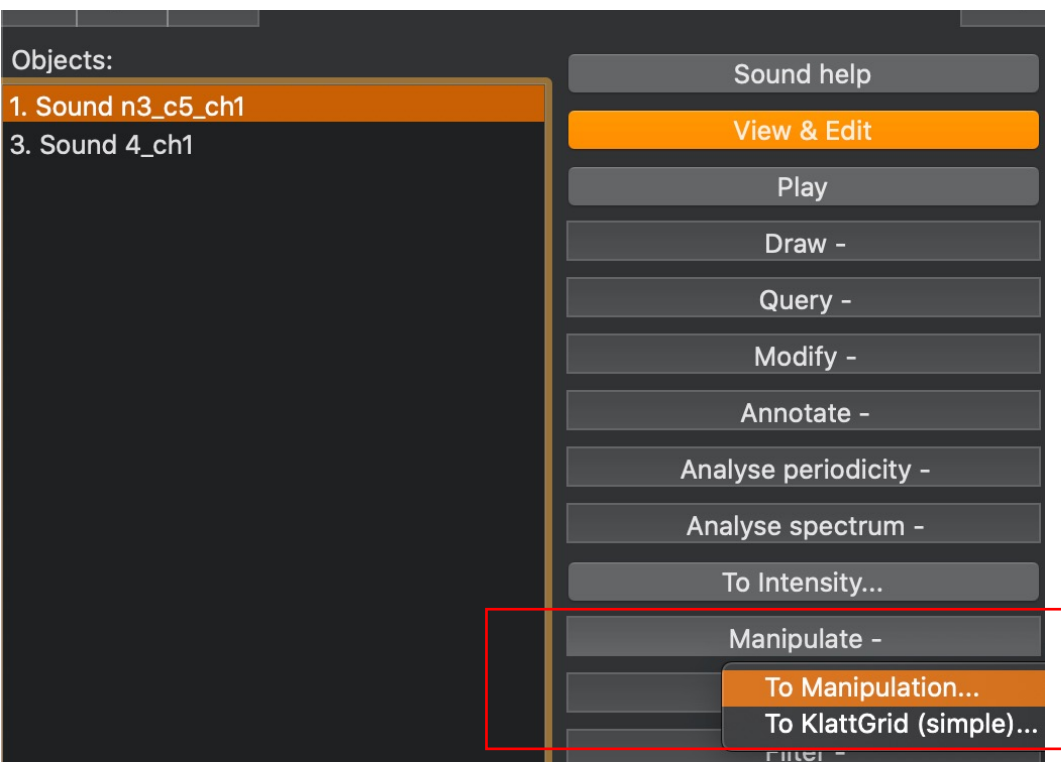
➤ 密码是 123456

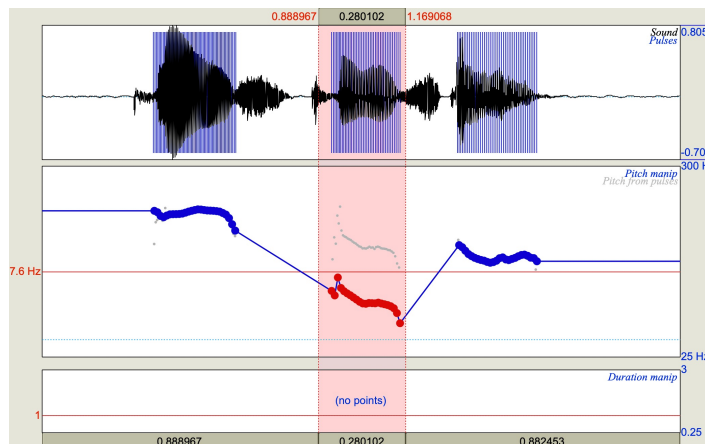
➤ 请慎用 sudo 权限，如搞崩了系统本节课内无法恢复的

- ✓ 然后即可从左上角菜单中搜索到 Praat

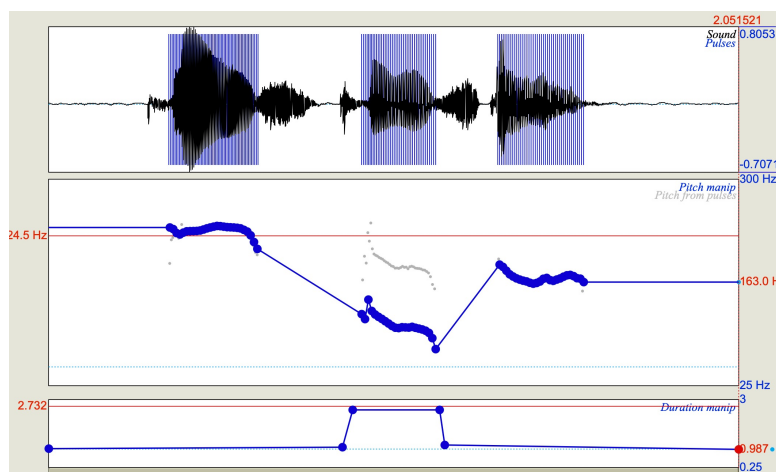
# Praat 基本操作讲解

## 基频、时长的查看与调整





Start:	Last:	Operations:
0.888967	0.280102	pitch:-80
听起来像男声		



Start:	Last:	Operations:
0	0.888967	duration:1.105
0.888967	0.280102	pitch:-80,duration:2.584
1.162554	0.873767	duration:1.165



- 在本实验课中，要求按照下述格式记录你在 Praat 中的操作以便复现的结果

起始时刻，持续时长，操作描述

- 例子：

1.82，1.42，基频上升 20 Hz

0，5.03，时长增加一倍

- 附注：

- ✓ 保证每行一个操作
- ✓ 保证每个操作描述人能看得懂
- ✓ 没有中英文、全半角等其他要求
- ✓ 不要存在空行或其它字符

□ 请根据 FTP 中提供的 input.txt 中描述的操作，调整 input.wav，并输出 output.wav 至 FTP 同目录

✓ 注意：每个人的题目是不一样的

□ FTP 登录方式

✓ 登录地址：<ftp://10.103.10.112>:(端口号)      端口：9999

✓ 用户名：学号      密码：在 info 预留的手机号

✓ Windows: 请从浏览器的地址栏输入 <ftp://10.103.10.112:9999>，登录之后选择：

➤ 【查看(V)】-> 【在文件资源管理器中打开FTP站点(O)】

✓ Ubuntu: 请从文件管理器的地址栏输入 <ftp://10.103.10.112:9999> 即可

✓ Mac: 请从 [http://10.103.10.112:8000/FileZilla\\_3.50.0\\_macosx-x86.app.tar.bz2](http://10.103.10.112:8000/FileZilla_3.50.0_macosx-x86.app.tar.bz2)

下载并安装 FileZilla FTP 客户端，并输入地址、用户名、密码登录

□ 提交状态查询

✓ 如不确定提交成功，请访问 <http://10.103.10.112:8000/> 查看提交状态

# 任务一：Praat 基本操作实战

## Windows 访问 ftp 方法示意图

- ✓ Windows: 请从浏览器的地址栏输入 <ftp://10.103.10.112:9999> , 登录之后选择 :
  - 【查看(V)】 -> 【在文件资源管理器中打开FTP站点(O)】





## 任务二：探求声学参数与情感的关系

具体评分标准：

- 请根据 FTP 中提供的 neutral.wav ，通过 Praat 调整，使其表达的情感接近 happy.wav
  - ✓ 输出要求：happy-k.wav 与 happy-k.txt
    - 其中  $k$  表示通过 Praat 调整的操作数，即 happy-k.txt 应该恰好有  $k$  行
    - 你可以至多提交两组 happy-k.wav 与 happy-k.txt ，其中  $k$  是不同的；请自行选择要提交的两组。
  - ✓ 评分标准概述：
    - 合成效果应当越接近 happy.wav 越好
    - 操作步数应当越少越好
      - 你需要在操作数  $k$  与合成效果中做权衡 😊
      - 具体评分计算方式见下一页
- 同理，请提交 sad-k.wav 与 sad-k.txt
  - ✓ 注意，sad.wav 的内容与 neutral.wav 是不平行（不同）的，请大家自行发挥 😊
- 提交状态查询
  - ✓ 如不确定提交成功，请访问 <http://10.103.10.112:8000/> 查看提交状态，会显示已提交的操作数

### □ 具体评分标准：

- ✓ 对每一个操作数  $k$ ，规整所有操作数为  $k$  的音频的合成效果评分到一个单位分布
  - 若某个操作数  $k$  对应的音频数量太少，则会先对合成效果评分做惩罚后与操作数为  $k$  附近的音频一起计算
    - 简单来说就是  $k$  太过另类的话会吃亏，不要钻空子 😊
- ✓ 对每一个学生，取每个提交的音频在对应操作数的单位分布中的位置的最大值记为得分
  - 提交多个总是好的
- ✓ 对所有学生得分，规整为单位分布，并以此计算最终成绩

□ 请对 FTP 中提供的每个音频做出主观评测，并整理成符合格式的 result.txt 上传至 FTP 同目录

□ 示例：

✓ 若 FTP 中提供文件 abc.wav bcd.wav cde.wav def.wav efg.wav

✓ 则要提交的 result.txt 如下

abc 1

bcd 5

cde 3.5

def 4

efg 3.5

➤ 其中，文件名和评分之间的间隔为一个空格

➤ 请不要删除 FTP 目录中的 wav 文件来试图偷懒 ☺，仅需要再提交 result.txt

□ 提交状态查询

✓ 如不确定提交成功，请访问 <http://10.103.10.112:8000/> 查看提交状态

### □ 评测标准

- ✓ 评分项：所给音频是否表达出了目标音频中的情感
- ✓ 评分制：采用 5 分制评分，最低为 1 分，最高为 5 分
  - 5 分：完全表达出目标音频中的情感，堪比目标音频本身 “以假乱真”
  - 4 分：很好表达出目标音频中的情感，感觉强烈，但还可以改进 “有滋有味”
  - 3 分：基本表达出了目标音频中的情感，但感觉不强烈，也不像中立 “食之有味”
  - 2 分：表达了一点点目标音频中的情感，但感觉接近中立 “食之无味，弃之可惜”
  - 1 分：完全没有表达出目标音频中的情感，或表达了相反的情感 “弃之不可惜”
- ✓ 可以精确至一位小数
  - 如：3.5 分，4.8 分

- 请访问 FTP 来查看各自的评测具体结果
  - ✓ 如有疑问请于一周内与助教联系
  - ✓ 经过标准化后的得分将于之后给出
  
- 高分结果选讲
  - ✓ 请经过初步统计后，得分靠前的同学来分享自己的调整方法及对情感的理解
    - 注：选讲不代表最终成绩