【Siri】

2011 年,Phone 4S搭载的Siri，或许就是这样的一个时间点。现在我们看起来非常鸡肋的功能，比如语音查时间、订日历等，在当时看来却是惊为天人。很多人第一次发现，原来还可以用语音去控制程序，更可以和一个虚拟人物对话聊天。



但是到了现在，或许Siri的发展不像当初预见的那样。

Creative Strategies 2016 年的一项调查数据显示，70% 的苹果手机用户很少或偶尔使用 Siri，甚至还有 2% 左右的用户从未使用过 Siri。

在公共场合中使用 Siri 也会让用户感觉尴尬。Creative Strategies 的数据显示，98% 的 iPhone 受访用户用过 Siri，但是，仅有 3% 的人在公共场合或其他人面前使用这款语音助手。

【Cortana】

2014年，微软正式推出Cortana，并将其命名为“微软小娜”。它的功能与Siri相似，但是添加了一些更加有趣的功能，例如必应美图、节日特辑等。

因为Cortana的某些功能，它会将某些问题在必应上进行搜索，并且它的后台会存储用户数据并建立索引，虽然可以关闭，但是关闭并不会删除存储在微软服务器上的用户数据。用户的隐私受到了威胁。



【Alexa】

2014年，Amazon推出了基于Echo发展的Alexa。相比于以上两款语音接口，Alexa还可以通过语音控制一些智能家居，同时用户也可以为Alexa安装一些功能。

但是大部分设备中的Alexa需要用户给出激活命令，例如“Echo”，另一部分设备需要用户先按键开启Alexa语音，才能与其进行交互，这在日常使用中十分的不方便。并且与Alexa的交互和交流仅限于英语、德语和日语，极大程度的限制了Alexa的使用人群。



尽管上述平台都提供了基于语音的强大交互，但是它们有着共同的缺陷，限制了它们作为主要人机接口的可能性：

1. 会话隐私：当通过这些接口进行通信时，用户将语音广播到环境中，因此不能保证用户隐私。例如与另一个人打电话的时候、与Siri等进行通信的时候。
2. 窃听：当不需要的时候，语音接口其实并没有被关闭，语音界面总是监听对话内容，只有稍后通过特定的触发词才能被激活。例如“Ok Google”激活Google智能助理、
3. 非个人设备：这些设备不是个人设备，任何其他用户都可以有意或无意地向这些设备发送有效的语音输入。
4. 需要集中注意力：当前的语音交互设备作为设备的可用性较低，用户不能随意使用语音界面，这种情况通常与不能移动的专用语音设备（例如Amazon-Echo）一起使用。此外，为了优化语音识别，需要用户靠近设备。

以前有很多次对于无声的语音通信的尝试，这些系统可以分为两种主要方法：侵入式和非侵入式系统。

侵入式系统

1. Brumberg等人在语音运动皮层中使用直接的脑植入来实现无声的语音识别，从而在有限的词汇数据集上证明了合理的准确性。他们通过将传感器放置在内部语音的发音器官里，对发音器官周围的运动进行了测量和探索。
2. Hueber等人使用放置在舌头上的传感器来测量舌头的运动。
3. Hofe和Fagan等人使用永磁体（PMA）传感器来捕捉用于讲话发音的肌肉上特定点的运动。这种方法需要永久固定磁性微珠，这在现在的世界中并不能很好地扩展
4. Florescu等人提出使用超声波来表征声道以实现无声言语。只有在摄像机正对用户嘴巴的时候，这个系统才能取得良好的效果。

除了临床的情况之外，设备的侵入性、突兀性或者静止性妨碍了这些解决方案在实际环境中的可扩展性。

非侵入式系统

目前有许多种以非侵入的方式来检测和识别无声语言的方法。

1. Porbadnik等人使用EEG传感器进行无声语音识别，但是这种方法需要以低信噪比来加大语音检测的准确度，因此它在实际环境中的表现并不好
2. Wand等人对没有声学发声的视频进行了深度学习，但需要有外置摄像头来解码唇部运动的语言。
3. HHirahara等人使用听不见的低音麦克风将数字转换成信号
4. Wand和Schultz等人使用基于音素的声学模型演示了表面肌电图的无声语言，但是用户必须明确地说出单词并且必须有显著的脸部运动。
5. Jorgensen等人使用表面肌电图来检测准确波动低于33%的次声词，同时系统也无法高准确度的识别肺泡声音，这是实际用作语音接口的重要障碍。

非侵入式系统虽然不会给用户带来不愉快的体验，但是他的准确率大大降低，很多时候需要用户作出夸张的动作来保证检测的准确性。