



用眼睛听世界

我们能够听到什么在很大程度上依赖于我们看到什么和摸到什么。

撰文 劳伦斯·D·罗森布拉姆 (Lawrence D. Rosenblum) 翻译 毛利华

20世纪70年代末，美国联邦调查局雇佣了苏·托马斯 (Sue Thomas) 和其他8个失聪人士进行指纹模式分析工作。联邦调查局认为，这项工作需要极其谨慎、细心，而失聪的人应该更容易保持专注。但是，从第一天开始，托马斯就发现这个工作单调得让人无法忍受，并跟上司抱怨了很多次。当上司安排她和其他联邦探员见面的时候，她甚至做好了被辞退的准备。

但托马斯并没有被辞退，从某种意义上来说，她还获得了晋升。联邦探员给她看了一段两个犯罪嫌疑人交谈的无声录像，并让她破译谈话内容。

私底下和托马斯交流时，这些探员就注意到她能熟练地通过唇读来理解他们的意思。正如同事们所料，托马斯轻松地破译了嫌疑人的对话，发现这些人与一个非法赌博团伙有染。从此，托马斯便开始了她作为联邦调查局第一位唇语专家的职业生涯。

托马斯需要依赖唇读与人交流，这磨练了她的技术。但事实上，正常人对唇读依赖的程度，远超我们的想象。如果我们看不到说话人的嘴唇动作，特别是在嘈杂的环境中，或讲话者是

陌生人的情况下，我们理解言语的能力便会削弱。学会同时用眼睛和耳朵去感知言语，是言语发展过程中的一个重要部分。因此，失明的婴儿由于不能看到说话人的嘴，通常要花费更多的时间去学习言语的某些方面。我们会把从别人口型中看到的词语与听到的词语，自动地结合在一起。近年来，对多感觉通道言语知觉的研究，使我们的大脑如何组织多种不同感觉信息的理解，产生了翻天覆地的变化。

神经科学家和心理学家已经基本放弃了那种“大脑像瑞士军刀一样，不同脑区负责不同感觉信息加工”的早期观念。如今的科学家认为，多感觉通道间的交流是大脑进化的趋势，大脑中不同的感觉区域在物理结构上是交织在一起的。

我们的感觉常常喜欢“多管闲事”，总是希望在其他感觉加工中插上一脚。譬如，视觉皮层主要负责视觉加工，但它也能很好地加工来自其他感觉通道的信息。一个具有正常视觉的人被蒙上眼睛90分钟之后，视觉皮层会对触摸物体变得异常敏感。同样，对盲人大脑的扫描也表明，他们的视觉皮层发生了重构，开始参与听觉的加工。当我们吃薯条时，薯条在嘴里嚼碎时发出的清脆响声也会部分影响我们对薯条好吃程度的感觉——研究者能够通过调整吃薯条时人们听到的声音，改变接下来他们对味道的评价。我们站立不动时，看向哪儿和看到什么，也会影响我们整个身体的姿势。总之，过去15年的研究已经证明，没有哪个感觉是独立工作的。对多重感觉的认识，就像一场革命，将为科学家提供新的方法，改进那些帮助失明、失聪人士的仪器，比如人工耳蜗。

触摸到的“言语”

麦格克效应 (McGurk effect) 是最早被发现、最稳定的多重感觉加工的例子。1976年，哈利·麦格克 (Harry McGurk) 和约翰·麦克唐纳德 (John MacDonald) 首次报告了这个效应：如果让你看一段录像，录像中的人不出声地重复做出音节“ga”的口型，同时让你听同一人发出的音节“ba”的录音，这时你就会听到“da”的发音。看见“ga”的口型，改变了你对听到的“ba”音节的知觉，是因为大脑将你听到和看到的信息，进行了整合。麦格克效应在所有语言中都会出现，而且即使你已经研究了它25年，在你身上这种效应依然会发生——我本人就能为此作证。

你所听到的言语也会受到你所“触摸”到的言语的影响。1991年，当时在美国达特茅斯学院的卡罗尔·富福勒 (Carol Fowler) 和同事，让没有相关知识的自愿者尝试一种被称为泰德马 (Tadoma) 的技术——即通过把手指放在讲话者的嘴唇、下巴和脖颈处来理解他人说的话。在人工耳蜗发明之前，很多又聋又盲的人 (包括海伦·凯勒) 就是依靠泰德马技术与他人交流的。这些自愿者用手指“触摸”到的言语改变了他们对身边扬声器发出的声音的解读。

1997年，英国牛津大学的吉玛·卡尔弗特 (Gemma Calvert) 用大脑成像的方法，通过让没有唇读经验的自愿者，在无声的环境下，观察一张缓慢清晰地读出数字1到9的面孔的唇动，找到了唇读时最活跃的脑区。卡尔弗特和同事发现，唇读激活了大脑中加工声音的听觉皮层，以及其他在听别人说话时所激活的区域。这是关于感觉联合的最早证明之一：其他感觉通道的信息会影响原本被认为仅处理单一感觉信息的区域。

最近的研究进一步提供了感觉联合的证据。科学家已经知道，脑中负责听觉的部分，也会对“看到”的言语作出某些

劳伦斯·D·罗森布拉姆是美国加利福尼亚大学河滨分校的心理学教授，也是《感知力》(See What I'm Saying: The Extraordinary Power of Our Five Senses) 一书的作者。



特征反应，而此前他们一直认为这些部分只参与对声音的初级加工。神经成像研究也表明，在麦格克效应中——录制的声音是“ba”，听到的却是“da”——大脑的反应就如同真的听到音节“da”一样。

这些发现说明，大脑对从耳朵、眼睛甚至皮肤搜集到的言语信息都给予同等重要的地位。当然，这并不意味着，不同的感觉通道提供了等量的信息：显然听觉能比视觉或触觉捕捉到更多语音细节。而是，大脑会努力加工、整合它所接收到的不同形式的言语信息，无论它们来自哪个感觉通道。

读脸识声

在另外一些情况下，不同的感觉会互相配合，加工相同类型的信息。例如，我们可以从一个朋友讲话的特定方式，认出他来，无论他的言语是被听到还是被“看到”。我和同事录制了人们讲话时的录像，然后去除讲话者所有可识别的面部特征——将面部处理成像萤火虫一样快速移动的光点，只露出下巴和嘴唇。当我们播放这些录像时，自愿者们能够唇读这些没有面部特征的光团，并认出他们的朋友。

言语中抽取出来的声音特征，也能够帮助我们识别出自己的朋友。美国哥伦比亚大学的罗伯特·雷米兹 (Robert Remex) 及其同事，将正常的讲话录音处理成正弦波。这种正弦波听起来就像《星球大战》里机器人R2-D2发出的奇怪声音。尽管失去了音高、音色等可以用来分辨人声的典型特征，这些正弦波却保留了关于讲话方式的信息，听者仍可根据这些信息辨认出他们的朋友。最惊奇的是，自愿者甚至可以将这些正弦波与仅有讲话人嘴唇信息的光点面孔匹配起来。

对听到的言语和“看到”的言语进行删减处理之后，言语中依然保留有相似说话风格的信息。这个事实表明，不同的知觉模式在大脑中是交织在一起的。神经成像研究给这种联系的存在提供了支持：听到熟人的声音会激发梭状回的神经活动，而梭状回是大脑中参与辨认面孔的区域。

这些发现可以启发我们做一个更神奇的预测：如果这些不同形式的知觉是混合在一起的，那么读懂一个人的唇语，就能更加轻松地理解他说出的话语。我们先让没有唇读经验的自愿者花一个小时去练习，尝试通过唇语理解一段无声录像中某人的讲话，随后再让自愿者听一组有随机背景噪音的句子。其中

精彩链接

神经科学家曾经认为，大脑就像瑞士军刀，不同的脑区负责截然不同的感知觉过程，譬如视觉、听觉、嗅觉、味觉和触觉。

近30年来，心理学和神经科学的研究表明，大脑其实是一个具有广泛多重感觉性的器官，它不停地整合着来自不同感觉的信息。

这场对于大脑认识的多重感觉革命，不仅改变了科学家对大脑功能的理解，还提供了帮助盲人和聋哑人的新方法，改进了语音识别软件的性能。

一半自愿者听到的句子是由刚刚录像中的人说出，另一半自愿者听到的句子则由另一个人说出，但自愿者事先并不知道。结果，当说出这些句子的人是自愿者在唇语练习中看到的那个人时，自愿者能够更好地从噪音中辨别出句子。

多重感觉革命

通过多感觉通道言语知觉的研究，科学家对以前被忽略的、不同感觉通道间的交互也产生了兴趣。例如，大多数人都知道嗅觉会影响对味道的判断，但一些研究表明，视觉和听觉同样也能改变味道。一个有趣的例子是，科学家发现橙子味的饮料被染成红色后尝起来会像樱桃，反之亦然。2005年，意大利特伦托大学的马希米利亚诺·赞皮尼（Massimiliano Zampini）发现，在自愿者吃薯条的同时，给他们听自己咀嚼薯条的录音，如果改变他们听到的咀嚼声的音色，会在某种程度上影响他们对薯条新鲜与酥脆程度的判断。另外，看着一个有不断下降视觉效果纹理（如瀑布），会让人们感到他们用手触摸着的纹理面是上升的。其他证据也表明，跨感觉的信号输入（cross-sensory input）会在无意识层面改变我们的行为。美国明尼苏达大学的汤姆·施托夫雷根（Tom Stoffregen）和同事让自愿者垂直站立，然后将他们的注视点从身边的一个目标移向远处的一个目标，这种视觉焦点的简单转换会引发身体姿势细微却系统的变化。

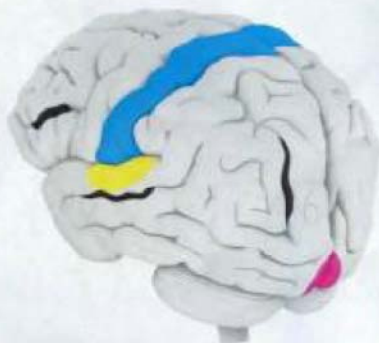
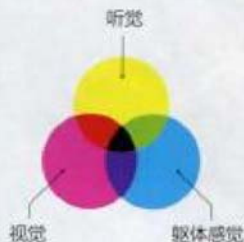
类似的发现相当普遍，以至于很多研究者都认为，大脑的感觉区域本身就具有多重感觉性。这种改进版的大脑模型，和那些表明大脑具有极大可塑性的证据是一致的——即使面临短期或轻微的感觉剥夺，大脑也能转换某个脑区的主要功能。过去4年，大脑成像研究已证实，蒙住一个人眼睛一个半小时，就会使他的视觉皮层对触觉产生反应。而且，视觉皮层的参与也确实增强了对触觉的敏感度。还有一个例子，近视通常会增强人们的听力和空间感知能力，即便他们戴着眼镜（大部分外周视野还是模糊的）。总的来说，跨感觉补偿比我们先前所认为的更加普遍。

这场多重感觉革命，已开始帮助那些丧失了某种主要感觉的人们。研究表明，如果一名失聪者的大脑，长期将被闲置的听觉皮层用于视觉或触觉等其他知觉，那么植入人工耳蜗的效果就会降低。因此医生通常会建议先天失聪的儿童尽早植入人工耳蜗。同样，类似的研究也告诉我们，应该让已经植入耳蜗的儿童观看人们讲话的录像，从而使他们学会把从别人口中看到的“言语”，跟他们听到的言语整合在一起。

研究面部和语音识别装置的工程师，也从多重感知的研究

对多重感觉的新认识

几十年前科学家就已经知道，某些脑区可以整合来自不同感觉的信息。例如，某个区域会整合视觉信息和躯体感觉信息，如触觉和温度觉。现在，新的研究已证明，多重感知是一个比研究者先前所知更为普遍的大脑神经架构特性，支持不同感觉间的相互交流，这符合大脑进化规律。



传统的图式



新近的图式

上图部分数据来自于灵长类大脑的研究，且仅标示了主要感觉区。

中有所获益。即便只是在中等水平的背景噪音下，语音识别系统也会表现得很差。但是，如果教会这些系统，同时分析录像中人们的口型，就可以大幅提高语音识别的准确性——这个方法甚至对安装在手机和笔记本上的摄像头也是有效的。

从某种角度来说，多重感知似乎违背了我们的日常经验。我们的直觉是，这些感觉是不同的种类，每一种感觉用来理解大千世界中一个不同的方面。我们用眼睛观察别人，用耳朵倾听别人；我们用手触摸苹果，用舌头品尝它。但是一旦感觉信息到达大脑，这种严格的分类便瓦解了。大脑并不像将一堆硬币分类一样，把视觉信息从眼睛传导到一个神经容器中，把听觉信息从耳朵传导到另一个神经容器中，而是将多种形式的感知混合起来，以尽可能多的方式，去感受和理解外面的世界。

本文译者 毛利华为北京大学心理学系副教授，心理学博士，主要研究方向为社会认知神经科学。

扩展阅读

Speech Perception as a Multimodal Phenomenon. Lawrence D. Rosenblum in *Current Directions in Psychological Science*, Vol. 17, No. 6, pages 405-409, December 2008.

The New Handbook of Multisensory Processing. Edited by Barry E. Stein. MIT Press, 2012.