

# 当“声音”成了一门大生意

是的，“声音经济”正在受到前所未有的关注——随着语音交互技术的不断成熟，智能音箱等打着人工智能旗号的硬件开始冒出来，各种音频内容更是爆发式产生，嗅觉灵敏的各路资本加速涌入。我们突然发现，“声音”已经成了一门大生意。

■本报记者 唐玮婕

## 语音交互的梦想 一场仍在持续中的研发

我们在很多科幻电影中都会看到类似的场景：故事主人公驾驶飞行器穿梭在宇宙空间时，只要轻松地与机器聊个天，或者开开玩笑，就能完成所有的操控。这是人类的梦想——用声音实现人机交互，体验完美的人工智能，甚至是带有情感交流。

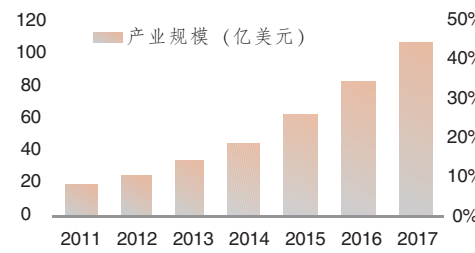
早在1967年，美国心理学家、传播学家艾伯特·梅拉比安等人经过大量实验，提出人类在沟通中全部的表达信息中，肢体语言信息占55%，声音信息占38%，语言信息占7%。而现在已经有专家预言，语音命令和听觉的结合，可能会成为任何自发行为的主要交互入口。

在现有技术支持的所有人机交互的方式中，声音大概是最为简单直接的一种，但它同时也是对配套技术支持要求最高的一种。要让机器听懂你在说什么，并不是件容易的事。事实上，各类研究机构、科技公司对语音技术的开发从未停止过。资料显示，从1970年代起，语音助理产品就已经陆续面世，但软硬件整体欠成熟。一直到本世纪初，计算机语音识别掀起了一轮小热潮，可惜的是，当时的最优成绩——80%识别准确度，并不足以说服大家收起键盘和鼠标。

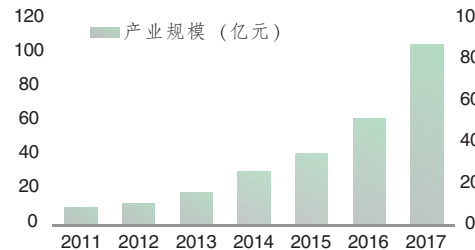
转变出现在最近10年，随着近几年人工智能技术的进步，深度机器学习的算法，终于让语音识别实现了质的飞跃。目前语音识别引擎已经可以达到95%的准确度，硬件端的麦克风阵列也在逐步解决远场语音交互难题，而大幅降低的成本，终于让语音交互、人机对话这类高大上的新技术走出实验室，进入普通人的生活——虽然没有出现类似触摸屏那样的“引爆效应”，但近几年各类智能音箱产品和语音交互软件相继推出并成为话题，印证了业界关于语音交互技术已实现“从0到1”突破、开始冲刺“引爆点”的判断。

语音交互能否成为“引爆点”？专家说了还无法作数，资本才是最后那个一锤定音的家伙。有意思的是，对语音交互的未来，市场研究机构齐刷刷展现乐观态度。中信证券的分析报告指出，人类人机交互的过程，大致经历了从人适应智能计算设备到设备适应人的过程，从早期黑白文字界面、到图形用户界面，以及目前的多点触控、语音会话等，人机交互方式的演进伴随着交互自然性、承载的信息量持续改善。智能语音产品——从硬件到软件的大量出现，推动人机交互方式更为顺畅、自然和富有效率，其发展遵循从人机交互方式不断进化的客观规律，具有历史必然性。而另一家国际市场调查机构最近给出的数据则显示，2026年全球智能语音市场预计将达到188.4亿美元，在2017年至2026年期间的年复合增长率为32.67%。

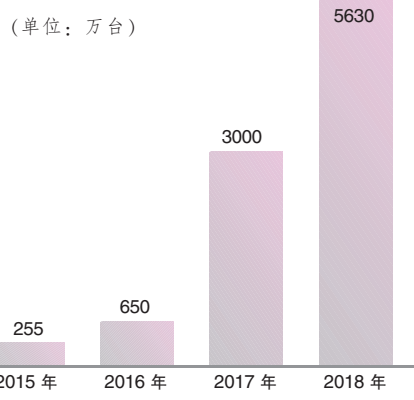
全球智能语音产业规模快速增长



中国智能语音产业规模增速显著高于全球市场



全球智能音箱产品出货量预测



## 智能音箱：被寄予厚望的入口

当然，我们距离梦想成真还是有点遥远。业界专业人士指出，目前的语音交互技术存在无法多轮对话、对话生硬等问题，需要进一步成熟，机器对自然语言理解也处于浅层语义分析阶段，基本局限为特定场景的弱人工智能。

但即便被视为“弱人工智能”，它还是催生出一个全新“物种”——智能音箱。正是借助智能音箱的普及，普通消费者才有机会与智能语音交互技术日常接触。

亚马逊的Echo音箱于2014年11月正式推出，产品测试阶段仅提供给少数客户购买，2015年正式上市后扩大至全美客户。就在去年年初，Echo音箱在全球获得了超过500万销量，这透露出一个令人兴奋的信号——智能语音交互产品第一次获得了规模化销量，并得到消费者们的认可。这个硬件迅速被寄予厚望，一些专家甚至预言，在不久的将来，智能音箱有机会成为人与机器交互的一大主流入口。

国内的厂家自然不会缺席这场剑指未来入口的竞争，去年第一轮“千箱大战”掀起后，阿里巴巴推出的天猫精灵和小米的小爱音箱在今年第一季度的销量分别达到100万台和60万台，分别位列全球市场的第三和第四位。目前，仍有一大批野心勃勃的公司决定加入战局，包括华为、联想、Oppo等手机厂商以及杭州的Rokid、出门问问等创业企业。

不过，现实的情况非常残酷，智能音箱在中国的市场始终不温不火，看似亮眼的销售规模，背后依靠的是巨头持续升级的补贴。例如，阿里巴巴和小米就双双采取了激进的定价策略，两款智能音箱价格都位于百元区间，可以说是“赔本也要赚吆喝”。业内专家指出，从中国市场的情况看，首先国内客厅文化并不普及，音箱暂时还止步于小众产品；与此同时，市场上销售的绝大部分产品，其语音交互技术还停留在“弱人工智能”的水平，用户在体验时，普遍反映会遇到远场识别差、误唤醒率高、语义理解能力差等问题。

从智能音箱近几年中的发展历程来看，这是一个“入门易深入难”的市场。所谓“入门易”指的是在市场启动阶段，消费者对产品的接受程度相对较高、增长势头迅猛。来自尼尔森的最新报告显示，亚马逊的Echo、谷歌Home和苹果的HomePod近年来销量增幅始终保持在30%以上，而且看起来这种趋势还在继续。到今年第二季度，已有24%的美国家庭拥有智能音箱，其中有40%的家庭拥有不止一台智能音箱。在已经拥有智能音箱的家庭中，45%的家庭还计划再购买智能音箱；62%的家庭是在过去6个月中购买智能音箱的……放眼全球，智能音箱保持着快速增长势头。根据市场调研公司Canalys的预测，智能音箱在全球范围内快速普及，预计2018年保有量将接近1亿台，用户量有望在2022年超过3亿。

在经历了启动阶段的突飞猛进之后，智能音箱接下来面临的问题，是如何让用户更多使用，所谓“深入难”，指的是以目前的技术而言，要让用户像离不开手机那样，对智能音箱产生使用粘性，难度非常大。

在强化用户粘性这点上，亚马逊依旧领先一步。国泰君安发布的报告《智能声学创新开启语音交互新时代》明确指出，在智能语音交

在智能语音交互领域，各路巨头布局周密，硬件+平台一个都不少



本版制图 杨天慧

## 翻译机的重生：主打智能语音

与智能音箱同时起飞的另一语音交互“爆款”产品，是智能翻译软件。

今年1月，搜狗在全球合作伙伴大会上发布了旅行翻译宝，支持文本翻译、对话翻译、语音翻译、实景翻译等。和讯飞翻译机2.0一样，也可以实现离线翻译。搜狗CEO王小川坦言，翻译是在人工智能和语言里面，比语音、图像复杂，但相对比对话问答又更简单，是更容易实用化的一个技术。

翻译机并不是什么高科技产品，学生一族过去有段时间几乎人手一部文字翻译机；而翻译技术也是早已有之，无论是谷歌、微软，或者

百度、网易有道等提供的在线翻译都相当成熟。而全新来袭的智能语音翻译机，主打的则是炙手可热的人工智能技术，完成自然语言之间的句子翻译，而依托的场景主要就是境外旅游。

这波热潮来得相当迅猛，就在今年上半年，最早推出翻译机单品的科大讯飞亮出升级后的2.0版产品；众多知名互联网企业集体宣布投入竞争——网易有道推出了名为“翻译王”的二代翻译机，百度有途的翻译机，搜狗有旅行翻译宝，小米、猎豹等也不甘落后；后来又有不少手机厂商、创业公司也加入了进来。

在国内，涌入人工智能翻译机这一领域的

企业，目前已经超过50家。而这50多家企业，基本上可以分为两大类：第一类是掌握核心技术的企业，无论是机器翻译、离线翻译都是自己来打造，比如科大讯飞、百度、搜狗等。另一阵营则来自消费电子领域的硬件公司，这些企业借助百度、微软、谷歌等的机器翻译技术，生产出自由品牌的翻译机产品。

在业内人士看来，智能语音目前最实在落地的场景就是翻译。智能翻译机的准确性最关键的是语音识别、语义理解、自然语言处理等人工智能技术。近几年来，各大科技巨头自动语音识别能力几乎均提升至95%以上。

纵观五花八门的翻译机，产品的定价相差较为悬殊。最低的不超过300元，例如猎豹移动发布的翻译机只要299元，小米生态链公司香蕉出行在5月更是推出了一款249元的产品。至于高端的价格则是千元级别，科大讯飞和网易有道的翻译机都接近3000元。按照科大讯飞

的说法，2.0产品的语音识别准确率能够达到98%，目前已能够实现34种语言的即时互译。

坦率地说，在产品体验上，现有的翻译机依然存在硬伤，包括在远场、有噪音的环境下，识别、翻译的准确度普遍不理想等。从语音合成发展历史来看，表现力、音质、复杂度和自然度是合成技术的技术难点所在。与此同时，传统翻译机产品功能过于单一，无论是摄像头、屏幕都不能满足用户在多场景下获得更好的体验，云端数据更新也比较慢。

“翻译机整体市场的规模预计在60万台左右，三到五年内能够达到三四千万台”。有业内人士估算，尽管各大厂商都看好基于智能语音交互技术的翻译机市场，对未来的前景相当乐观。然而，客观来看，整个市场尚未大规模发展起来，用户对产品有需求，却依然需要培养。从实验室的理论模型，到生活场景下的真实产品，智能翻译机还有很长的一段路要走。

### 链接

与声音相关的硬件如此红火，音频的在线平台、内容制作自然而然也站上了风口。纵观整个市场，在线网络音频行业已经挤满了各式各样的竞争者，喜马拉雅、蜻蜓FM、荔枝、懒人听书等都在攻城略地，资本也在加速投入。其中，蜻蜓去年获得了百度和微影资本领投的10亿元人民币融资；荔枝则在今年1月拿到由兰馨亚洲领投，媒体和互联网投资平台EMC跟投的5000万美元D轮融资；有声阅读平台“懒人听书”近日完成C轮2亿元融资，由时代出版、前海天和、厚惠成长共同领投……

在商业模式上，这些公司的定位略有差异。喜马拉雅FM一直自称是“声音的淘宝”，目前总用户规模突破4.7亿，平台内共有500万主播，其中包括20万名认证主播。活跃用户日均收听时长超过135分钟。蜻蜓FM则通过大规模自媒体入驻并制作发布音频节目，同时对接大量的广播资源，并未开放用户上传，也没有放弃直播模式。而荔枝在今年拿到融资后，已经宣布主营业务转为语音直播，希望走出一条不一样的道路，对外号称的月收入接近1亿元规模。

来自《2018中国网络音频全场景发展研究报告》的数据显示，2017年中国网络音频用户规模达2.6亿，同比增长18.2%，预计2018年用户规模将达到3亿，同比增长15.4%。与此同时，音频内容生产者也得到了资本的垂青，先后公布融资的好消息。例如，2014年5月成立

“声音经济”可以说是无处不在，其中属于网络音频的市场还算小众，渗透率并不高。有数据显示，当前网络音频的渗透率大概还不到10%，加上开车听音频的用户也不超过17%。网络音频的用户渗透率与视频、支付工具、电商还有较大的差距。

程一电台，在今年8月宣布获得A轮融资。数据显示，程一的全网播放量累积超过25亿，粉丝已经超过1200万。而听众累计达到2500万的微信公号“夜听”去年获得了来自真格基金的投资。

值得注意的是，喜马拉雅FM正在逐渐搭建音频创业生态圈，通过广告、知识付费以及直播等形式，让入驻的主播们以及平台都能获得收益，从而形成商业变现体系。就在今年，喜马拉雅FM推出了“万人十亿新声计划”，计划在未来一年投入三个十亿，从资金、流量及创业孵化三个层面全面扶持音频内容创业者。

喜马拉雅FM联席CEO余建军曾多次以“有声的紫襟”作为有声书创业者的典型代表，“我们平台上有一位主播叫紫襟，刚毕业不久，月收入已达到六十多万元，已经实现了财务自由。他就是比较专注，比较执着地做这么一件事，做到了极致。”目前，“有声的紫襟”的粉丝超过340万，节目累计播放量达到21.1亿。

有差距，也意味着有潜力。经过前期的蓄势，现在看好音频内容的已经成为主流，因为音频可以在很多无法使用视频、文字的场景中使用，满足了人们碎片化时间的娱乐和知识需求，未来还有很大的发展空间。“声音最大的魅力在于陪伴”，余建军曾公开表示，“声音经济”的迅猛发展将与“眼球经济”并驾齐驱。

## 音频逆袭 时间碎片的商机



### 智能音箱 能为你做什么

90%的用户用智能音箱来听音乐；68%的用户与他们的智能音箱聊天；81%的用户利用智能音箱上的语音助手来获得各种资讯，例如交通状况和天气情况；75%的用户会查询“事实性信息”；68%的用户利用智能音箱来收听新闻、设置闹钟以及时间。

图像处理 王梓含