



安信电子：高度重视音箱行业的逻辑！我们联合计算机、传媒团队独家发布“声音的世界、巨头的战争”深度报告，再简明扼要说明逻辑：1，无论是耳机还是音箱，生产力和生产关系严重不匹配，存在被颠覆的刚需；2，声学的时间价值，每人每天平均 2 小时听音乐，时间超过打游戏等娱乐和汽车，时间是衡量价值的核心；3，all in one 的价值，以声学智能化和独立计算平台为核心，整合各类功能，成为家庭独立计算平台，级别广阔互联网价值；4，亚马逊的 Echo 已经证明音箱的成功，无论是销量还是智能家居和物联网入口，我们产业链充足信息可以相信，苹果和谷歌都很快进入这一市场；5，商业模式开始成熟，从内容的收费到平台的建立，有商业模式才有持续性；6，A 股产业链走进全球前沿。核心标的：国光电器（A 客户音箱供应商，弹性大，下周调研），歌尔声学（声学巨头，声学的互联网品牌战略明确），全志科技（整体解决方案提供者，很重要），和尔泰（智能家居解决方案，弹性大），北京君正（咕咚音箱芯片提供商，弹性大），漫步者（音箱品牌，开始智能化和互联网），东软载波（智能家居平台）。

■声学是下一个互联网的入口。Amazon Echo 销售大超预期，从产业链看，互联网厂商都将进入智能音响行业，之前的音乐版权之争有望通过硬件落地。全球进入数字音乐时代，智能音响满足声音的输出要求。在输入端智能语音技术发展成熟，智能音响又能满足人工智能交互的最佳输入方式。信息的输入与输出得到完

美诠释，声学将是互联网的新入口，基于声音的商业模式将得到延展。■智能语音技术包括了识别和交互环节，融合了多种人工智能尖端技术，将成为学习成本最低控制方式。人机语音交互的应用关键在于准确率与响应时间，语音识别技术日益成熟，即将突破量变到质变的临界点。语言识别技术和硬件发展驱动语音交互渗透率迅速提升。■智慧家居将把家庭的安全、娱乐、饮食、健康等结合起来，用户面对的将不再是各种家电和移动设备的零散组合，而是一个有机的整体。语音交互控制产品中音响类产品成主流方案。主要由扬声器、麦克风以及处理芯片、无线传输模块等构成。这类产品首先具有基础音频输出功能，可以通过语音交互控制音乐播放。■投资建议：我们预计智能音响行业有望成为新的互联网和物联网入口，互联网厂商加速切入，从 CES 等展会观察到约三分之一的厂商在布局声学产品。音乐占据人们的时间，平均家庭收入 33,800 美元的家庭平均每周花 19 小时听音乐，音乐支出占到娱乐花费的 18%。音乐的商业模式在数字时代正在变化，内容收费需要与硬件结合。智能音响更多的价值也体现在家庭物联网的整合控制和人工智能的实际载体。■电子行业相关标的：1、音响制造厂商：推荐国光电器，关注漫步者；2、消费电子芯片厂商：推荐全志科技，关注北京君正；3、声学元件厂商：推荐歌尔声学、信维通信。■计算机行业相关标的：1、软件平台处于产业核心地位：重点推荐科大讯飞、和而泰；2、家用服务机器人是产品最终发展形态：重点推荐参股 Jibo 智能家用机器人的东方网力；3、平台内容提供商受益新的商业模式：建议关注与图灵机器人合作的奥飞动漫、内容及版权提供商光一科技、安妮股份等。■传媒行业相关标的：1、音乐艺人：推荐苏宁环球，关注共达电声；2、音乐演出&音乐综艺：推荐华录百纳、中南文化；3、播映平台：推荐暴风科技，关注乐视网

## 1.重申声学是下一个互联网入口

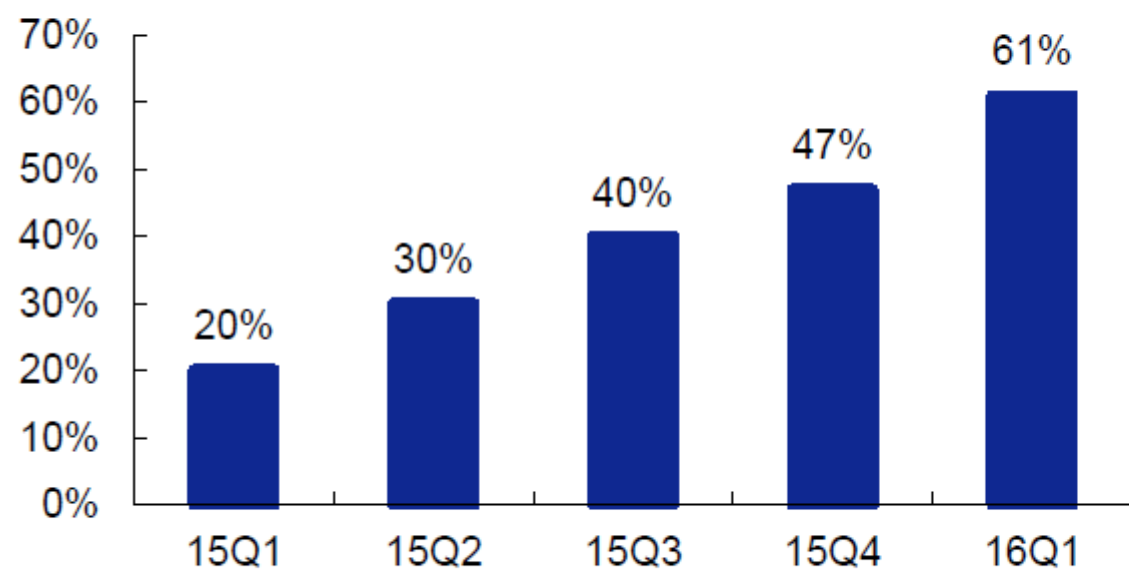
我们上周从亚马逊 echo 提示智能音箱的大机会并且此前一直看好音箱机会，我们在两年前的苹果产业链和歌尔声学的研究报告中就指出，声学将成为下一个互联网入口！Amazon Echo 再次吸引人们关注，我们在两年前判断的产业趋势进入临界点，声学完成从量变到质变的积累，智能语音技术、流媒体、大数据、人工智能切入智能音箱，互联网厂商与硬件厂商精密合作，家庭的互联网入口未来将基于声音的输入和输出！语音正成为计算接口，也是计算机输入最有效率的一种形式，自然语言处理软件 算法及降噪保真等硬件的基础完善，声学的世界将响起巨头们征战的号角！

### 1.1.从 Echo 说起，互联网厂商的下一个战场是声学

《亚马逊 Echo 研发幕后全揭露》再次让市场想起 Amazon 的 Echo，语音交互革命的原型已经迭代，互联网厂商今年预计将大举进军智能音响领域，行业有望突破临界点。新产品的上市总是会经过热闹后归于平静，但是当我们再次关注时发现，产品的销量已经超过了很多人的预期。根据 2016 年 4 月 CIRP 发布的一份调查报告指出，亚马逊 Echo 的销量预计已经超过 300 万台，其中 2016 年第一季度的销量达到约 100 万台。约有 3.5 万人在亚马逊网站给这款扩音器打过分，评分有 4.5 星之高，满分为 5 星。2015 年以来，Echo 在亚马

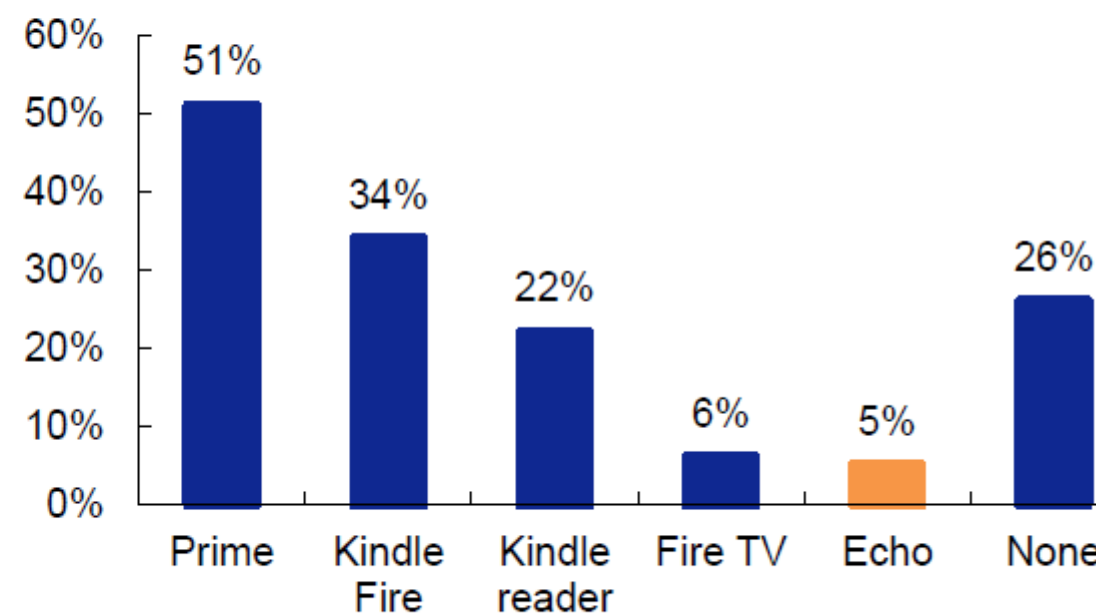
逊用户中的产品认知率和拥有率不断提升。其中 2016 年第一季度 Echo 在美国亚马逊用户中的认知率达到 61%，在客户的亚马逊设备拥有率提升至 5%。相较智能手机在 2016 年的出货量下降，语音交互产品的销量火爆与之形成鲜明对比。

图 1: 美国亚马逊客户对 Echo 产品认知度



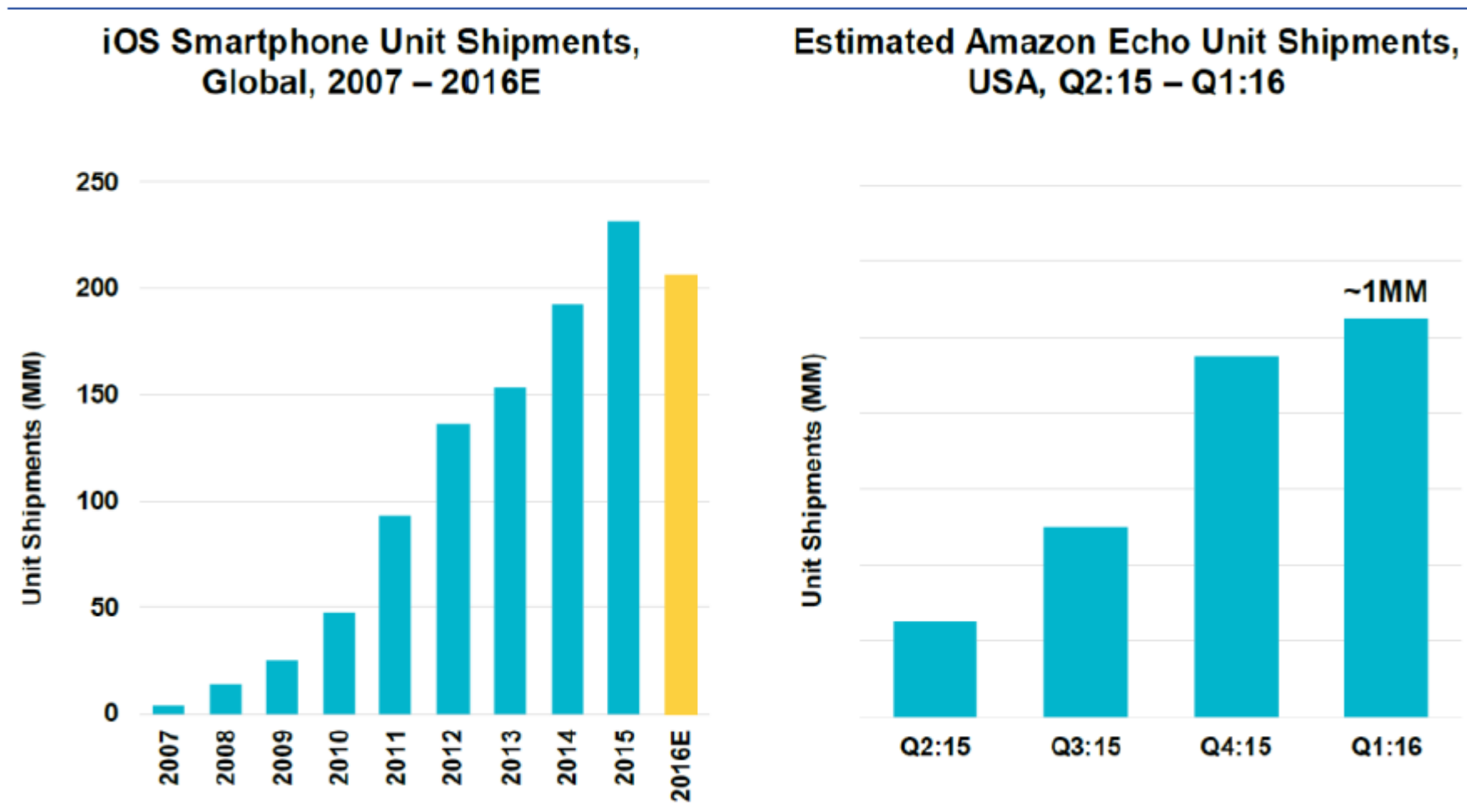
资料来源: CIRP、安信证券研究中心

图 2: 美国亚马逊客户的亚马逊设备拥有率



资料来源: CIRP、安信证券研究中心

图 3: Echo 销量大幅提升



资料来源: CIRP、安信证券研究中心

Amazon Echo 是智能音响的先驱，国外流媒体的成熟催生了声学硬件的创新。从 亚马逊官方的介绍来看，这款标价 199 美元的智能设备是一台可联网声控蓝牙音响，能够通过连接 Alexa 语音服务平台提供音乐播放、新闻、天气等语音推送 服务，交互模式是完全语音应答式的。Echo 拥有 7 个麦克风和音响集成，所以即使在室内有语音播放的情况下仍能够实现语音指令输入，同时它也可以非指向性 地实现 360 度全景声音输出服务。当你需要使用它时，通过呼喊“Alexa”或“Amazon”等关键词就可以立即触发设备。

Amazon Echo 外观上是一个没有屏幕的圆柱体，高约 9.25 英寸（约 23.5 厘米），直径 3.27 英寸（约 8.35 厘米），顶部外圈是代表设备唤醒状态的响应 灯，内部还有 7 个环状排列的麦克风，用于准确捕捉来自各个方向的声音信息。在设备内部，为了提供非指向性的 360 度全景声音，2 个扬声器被叠放在机身的下 部，上面是 2.5 英寸的低音单元，下面是 2.0 英寸高音单元，上方还有一个反射元件用于提高低音扩音输出表现防止失真。从外观上来看，Echo 的大小相当 于一个水杯，具有可移动性，而根据用户反馈的效果，Echo 的 7 个麦克风组成的听力识别系统非常强大，在隔着房间和杂音干扰下都能准确识别语音指令。



图 4: Echo 产品示意图



资料来源: Amazon、安信证券研究中心

Alexa 语音服务平台瞄准下一个计算界面平台，目标设备锁定“家居+汽车+手机移动端”三大领域。不同于 Siri 这类语音助手，Echo 提供的解决方案是从功能性出发的，而基于云端的 Alexa 语音服务平台更像一个语音库级别的 App Store，其中的服务包括亚马逊为 Echo 设计的原生语音功能，第三方开发者通过 Alexa App 为 Echo 设备开发的定制服务，还有一些基于 IFTTT 提供的语音服务触发器。

在 Echo 设计的原生语音功能方面：目前除了音乐、新闻、天气、提醒事项等语音服务功能外，还可以通过 Echo 连接亚马逊用户进行语音购物，语音购物使你能够在家中随时随地进行语音购物，不需要借助你的手机或其他智能设备，只需要通过语言指令或者对需求商品的描述就可以进行更快更方便的智能购物。

图 5: 通过 Echo 在亚马逊进行语音智能购物



资料来源: Amazon、安信证券研究中心

平台的价值在于生态圈，能够集成第三方的 OEM 语音服务，成为家居、汽车和手机移动端三个领域语音服务的集合中心。在智能家居方面，包括以 Philips Hue 为代表的智能照明，可通过语音随时调节亮度、开关灯等，以 Ecobee 为代表的智能控温器能根据外部监测数据进行中央空调温度调控，还有家居安全防护整套解决方案，包括电子锁控、摄像头监控、灾害监控等，可实现远程监控和操控功能；在汽车电子方面，Alexa 也积极布局与目前的车载智能平台进行整合，比如目前已有的和 Ford SYNC 平台合并，将汽车和家居联系起来，比如可以实现在家中通过 Echo 的语音指令远程启动汽车，也可以在汽车行驶过程中通过车载平台开启或关闭家居设备。而在手机移动端，如果安装了 Lexi app，同样可以通过手机来和 Alexa 语音服务平台连接，而不需要额外的 Echo 设备。通过把这三大领域连接到 Alexa 语音服务平台，真正通过云端的语音交互服务作为入口实现了物联网的设想，即只通过网络不依赖特定硬件的万物连接。

同时在 2015 年 11 月 Alexa 语音服务平台开始支持 IFTTT 的触发式命令，可以通过 IFTTT 来定制许多个性化语音服务，而不是仅仅依赖于亚马逊的原生语音服务和一些第三方集成的 OEM 解决方案。比如用户可以通过语音请求 Alexa 发出“trigger party time”指令可以触发 Phillips Hue 智能灯泡变成在 IFTTT 中定制过的不同颜色。

图 6: Echo 拥有大量第三方合作伙伴



资料来源: Amazon、安信证券研究中心

图 7: Alexa 语音服务平台的目标语音设备

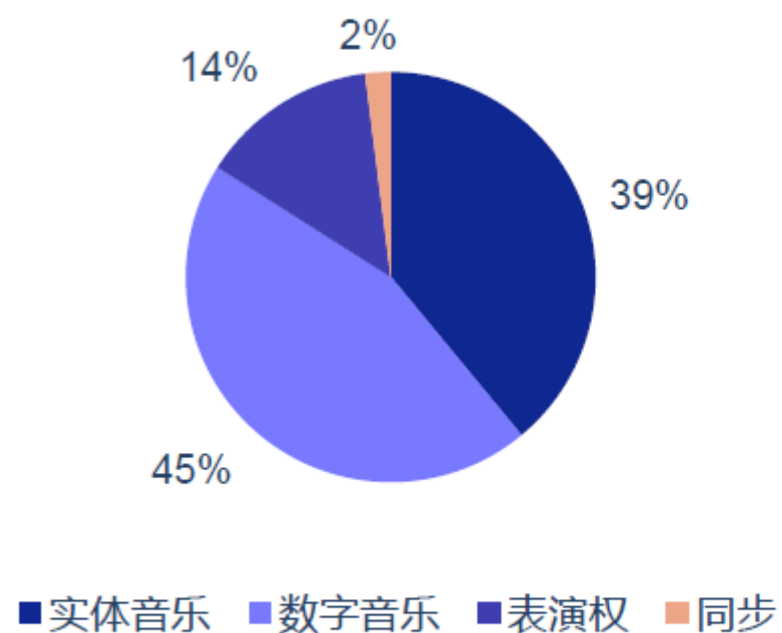


资料来源:《2016 Internet trend report》、安信证券研究中心

## 1.2. 互联网公司的音乐储备需要硬件落地

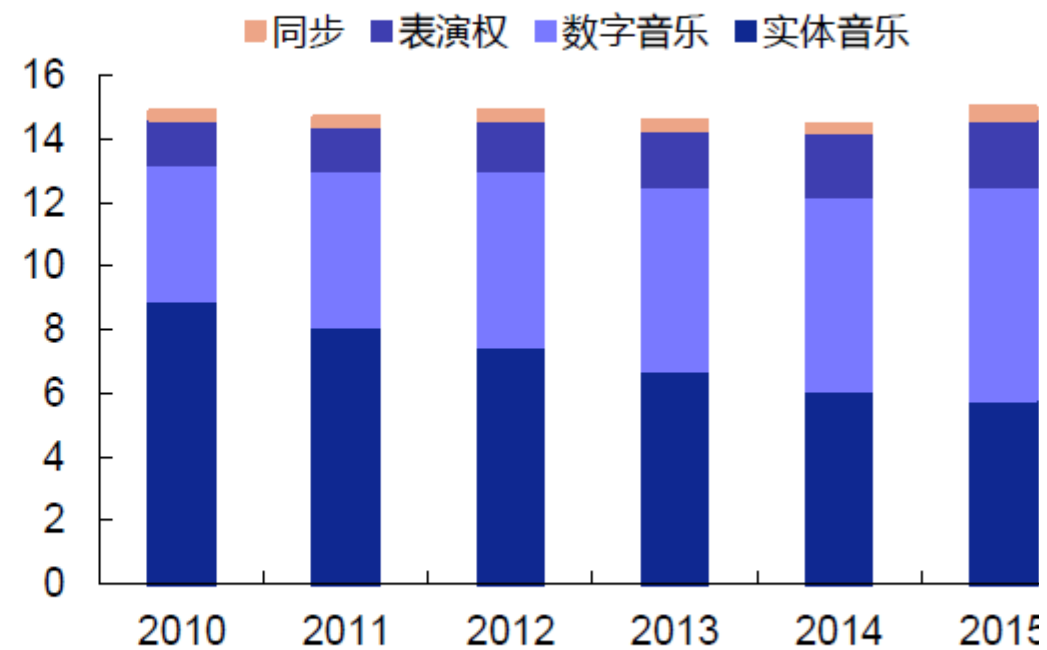
互联网公司频繁参与音乐版权争夺，全球进入数字音乐时代，但是依然需要硬件落地，智能音响满足声音的输出要求。国内外互联网公司纷纷加入音乐版权的争夺。一个有趣的现象可以看到，现在出现很多歌曲只有单一的平台可以收听，比如 QQ 音乐成为《中国好歌曲第三季》全网独家音频首播平台，虾米上的正版限制越发严格。现在已经很难在一个平台上听到所有的歌曲，音乐版权的争夺基本进入平稳期，精英消费群体的付费习惯也已养成，预计下一步的变现方式将是智能硬件。得益流媒体迅速增长，全球数字音乐规模首超实体音乐。IFPI 的 2016 年全球音乐报告指出，得益于流媒体收入 45.2% 的增长率，数字音乐创造了 10.2% 的收入增幅，总量达 67 亿美元，销售占比 45% 首超实体音乐。音乐行业的整体收入增长了 3.2%，达到了 150 亿美元的规模，这也是近 20 年来音乐行业首次实现逐年增长。

图 8: 2015 数字音乐市场规模首超实体音乐



资料来源: IFPI、安信证券研究中心

图 9: 全球音乐产业收入情况 (十亿美元)



资料来源: IFPI、安信证券研究中心



**表 1：流媒体与传统下载模式比较**

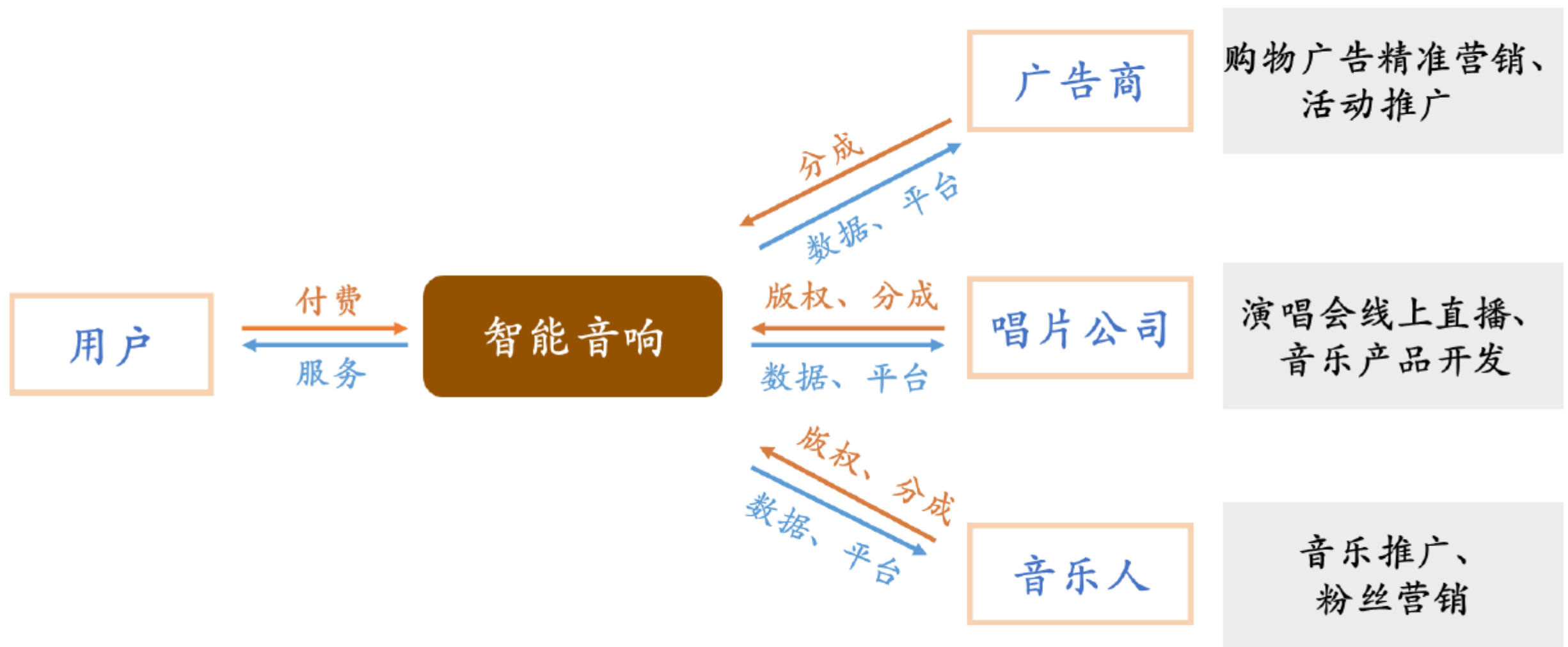
比较方面	传统音乐下载	流媒体
软硬件结合	传统音乐下载模式不可能也没有必要做到聚合，它本身就是一个个单一的文件，没有平台，只有不同的压缩格式。和软硬件之间的结合也没有那么紧密，无非是将数字文件放进哪个硬件之中通过什么软件播放而已。	流媒体在这一点上有无限可能，基于 WiFi 环境的智能音响搭载官方 APP 能够实现音乐推送和在线播放。
传输	需要通过硬件介质进行传输和迁移，如 TF 卡、存储卡。	流媒体可以实现在线传输和迁移，不需要硬件介质。
音乐来源	需要通过多个渠道去寻找资源，再下载，导入 itunes 或者 FB2K，无法及时更新歌曲	流媒体 APP 厂商采购版权，然后将版权方的东西放置在服务器里，就等于给了用户一个无限大的歌库，并且能够随时更新

资料来源：APP Annie、安信证券研究中心

流媒体音乐时代助力智能音响市场发展。我们认为在流媒体音乐时代，具备 WiFi 联网、流媒体音乐播放接口和语音交互功能的智能音响产品能够受益音乐行业发展潮流，市场空间广阔！从 CES 2015\2016 参展情况来看，超过三分之一的消费电子品牌都展出了智能音响产品，除去国内新兴的 Sonos、叮咚智能音箱、电蟒智能云音响外，三星、LG、松下、索尼也纷纷推出智能音响产品。我们从以下三个角度看好智能音响市场：

- **符合流媒体音乐发展潮流：**大多数智能音响具有流媒体播放接口。以 Sonos 为代表的智能音响采用开放式策略，与第三方音乐服务商合作，为用户提供个性化流媒体音乐播放服务。目前 Sonos 在国内的合作伙伴包括网易云音乐、QQ 音乐、虾米音乐等主流音乐服务商，用户可以将这些音乐服务商中的曲目同步至 Sonos 自身 APP 从而在智能音响上播放。也有以电蟒云音响为代表的智能音响厂商通过自建云曲库，为用户提供高品质流媒体音乐服务；
- **结合智能语音操作及其便捷：**智能音响搭载语音交互功能是大趋势，通过语音进行操作较传统音乐播放更加便捷。同时也能通过 APP 远程操控音响系统打破距离限制，无需复杂的布线安装。
- **盈利模式变化，未来目标是打造声学生态圈：**智能音响的核心优势不是信号传输无线化，而是声学生态圈的构建！智能音响硬件产品获得用户认可后，厂商可以通过自有平台掌握用户大数据，进而联合流媒体音乐服务商、音乐人和唱片公司提供精准营销、音乐推广等增值服务，**摆脱单纯靠硬件销售盈利的模式，构建“用户+智能音响+大数据+增值服务”的声学生态圈！**

图 10：智能音响构建声学生态圈



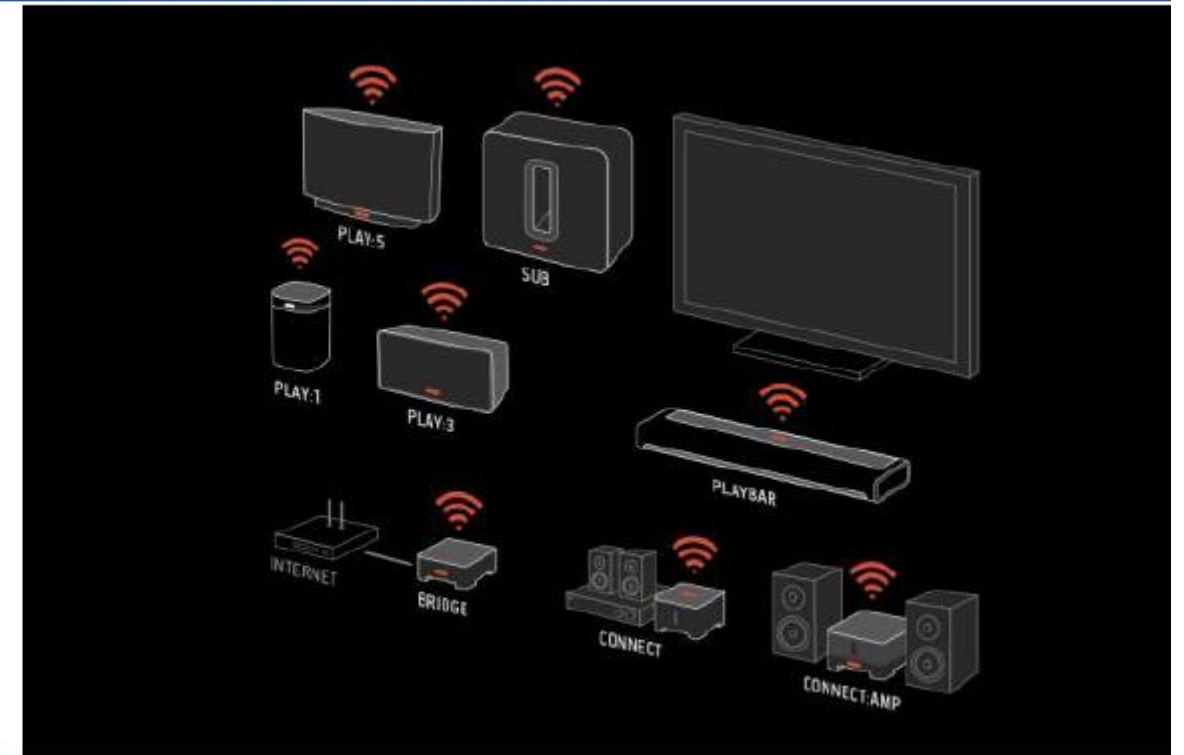
资料来源：Sonos，电蟒云音响，安信证券研究中心整理

图 11: Sonos 已与多家音乐服务商达成合作



资料来源: Sonos、安信证券研究中心

图 12: Sonos 家庭音响系统

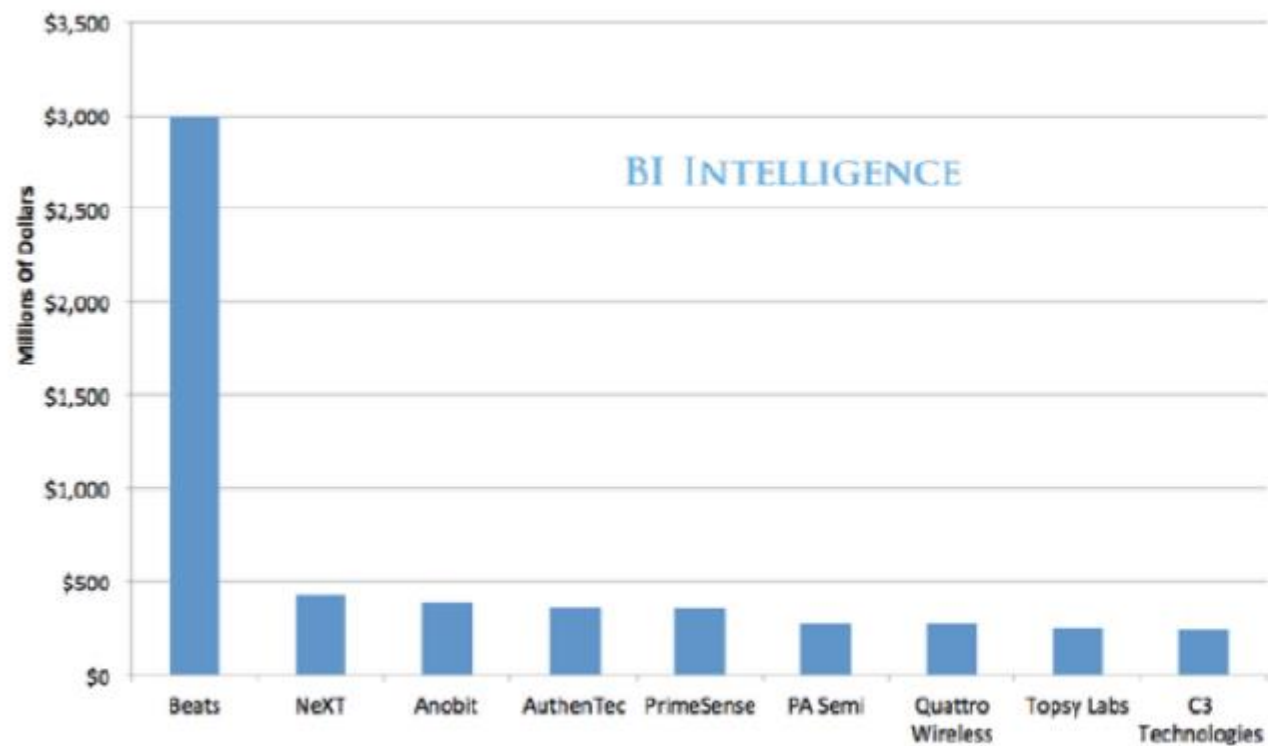


资料来源: Sonos、安信证券研究中心

苹果重金并购 Beats，意在整合流媒体大数据与智能声学产品。14 年 5 月，苹果以 30 亿美元收购耳机生产商 Beats Electronics 和流媒体音乐服务提供商 Beats Music，是苹果历史上价值最高的一笔收购。流媒体音乐服务 Beats Music 的 CEO Ian Rogers 接管苹果 iTunes Radio 服务，直接向苹果负责互联网软件和服务的高级副总裁 Eddy Cue 汇报。



图 13: 苹果历史上几次大规模收购



资料来源: BI Intelligence、安信证券研究中心

图 14: Beats 具备音频设备、音效系统和流媒体服

① Beats Audio 音效系统：由惠普和 Beats 共同开发，目前应用于 HTC 手机、克莱斯勒车载音响等多个移动和交通领域。

② Beats 音频设备业务：包括多款头戴式耳机和入耳式耳机以及便携式户外扬声设备。有调查显示，截至 2012 年，Beats 公司在高端头戴式耳机市场占有 51% 的份额。

③ Beats Music 流媒体服务：今年一月推出，基于算法向用户推荐音乐，收费 9.99 美元/月。业界估计，该服务订阅用户为 20 万左右。

资料来源: 凤凰网、安信证券研究中心

苹果收购 Beats 的 30 亿美元出价，远远高于苹果成立以来任何一次收购，显示苹果对互联网音乐志在必得之意。除 Beats 以外，苹果还收购了拥有 Musicmetric 音乐分析工具的英国创业公司 Semetric。我们推测苹果未来将以 Beats 品牌推广智能耳机及智能音响产品，结合 Apple Music 等音乐服务，打造类似过去 iPod+iTunes 的软硬件结合模式，与 Spotify 等互联网音乐服务商竞争。

Spotify 用户、估值迭创新高，引领流媒体音乐发展潮流。2015 年 5 月，总活跃用户增长至 7500 万，其中付费用户 2000 万。付费用户占比达到 26.7%。自该平台发布以来，Spotify 用户中的付费用户数量和免费用户数量占比一直稳定在 25% 和 75% 左右，付费用户每月支付 9.99 美元即可收听到 320 kbps 高音质无广告版音乐，如果不愿意付费，也可以使用音质 160 kbps 的免费流媒体服务。2015 年 5 月，Spotify 宣布将融资 5.26 亿美元，估值达到 85 亿美元，跻身全球身价最高的非上市科技企业行列，已超过大部分唱片商估值水平。目前 Spotify 尚未实现盈利，近七成营收用于向版权持有者支付版权费。

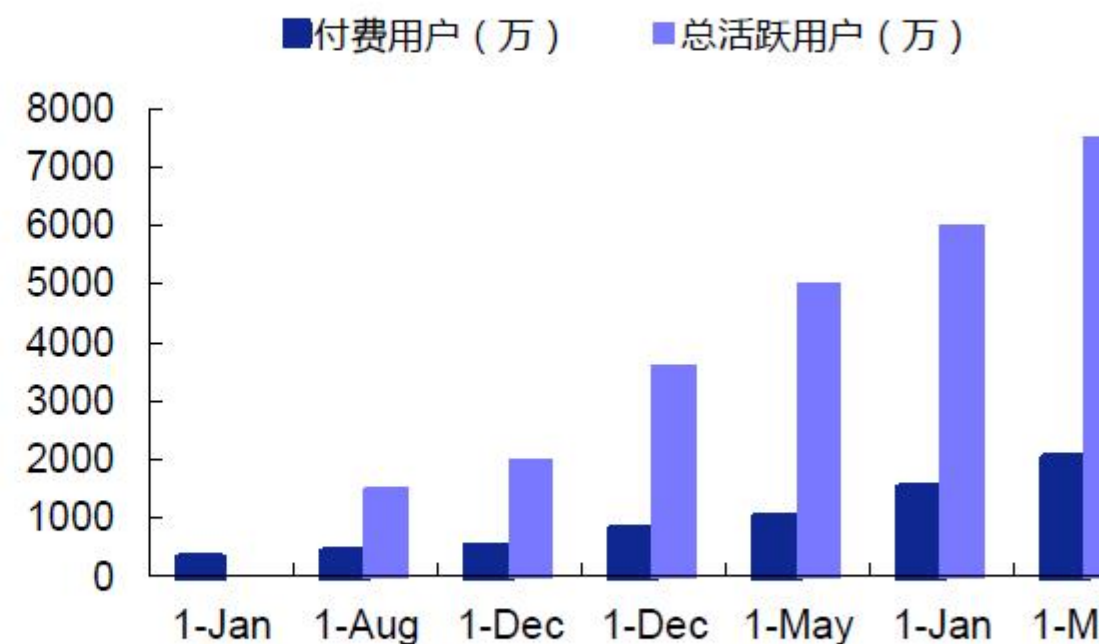


图 15: Spotify 成为互联网音乐市场最强大力量



资料来源: 互联网、安信证券研究中心

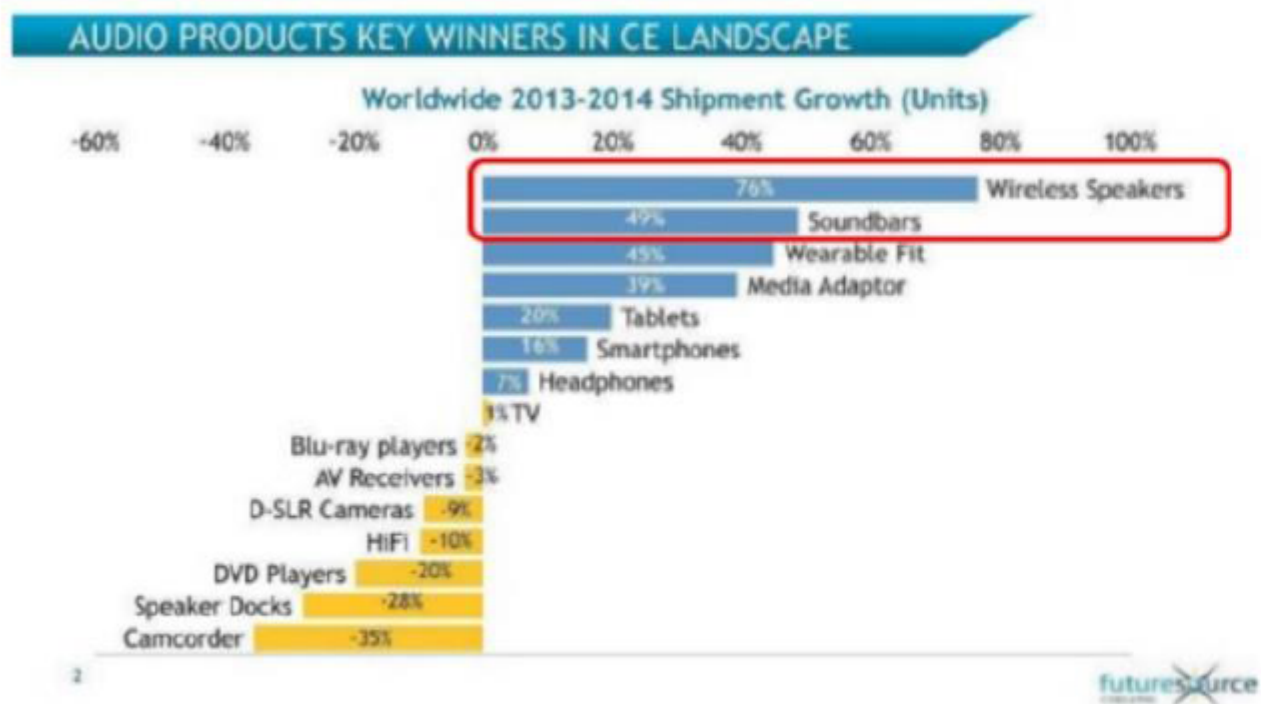
图 16: Spotify 的用户变化 (万人)



资料来源: Kobalt、安信证券研究中心

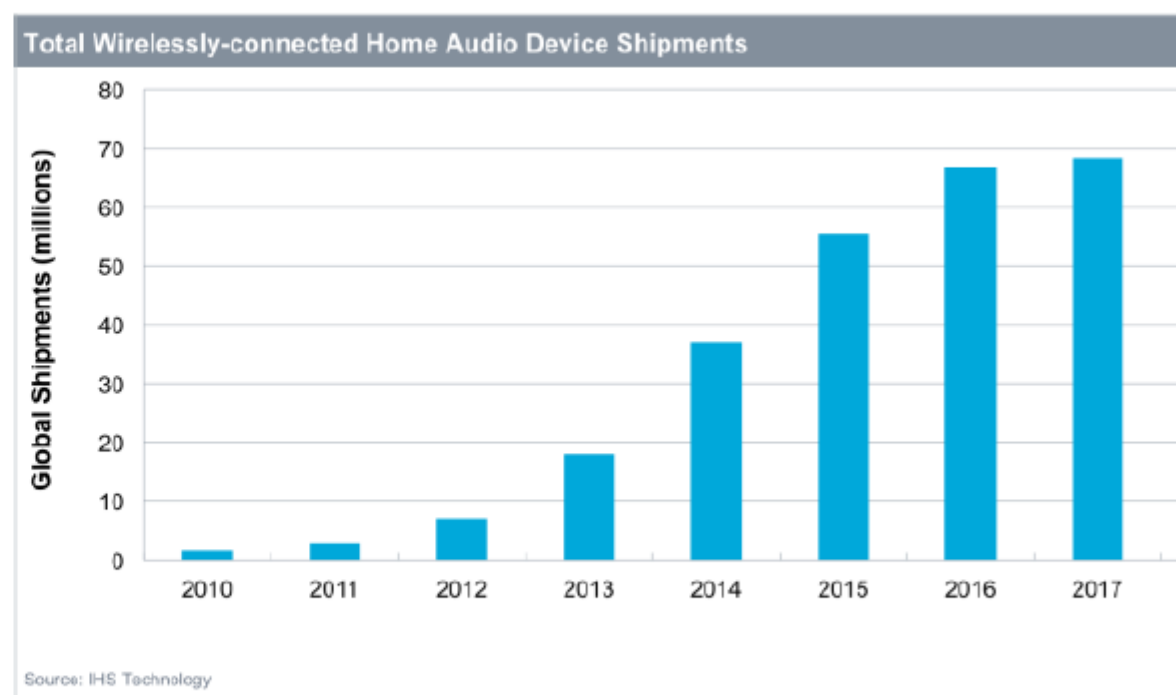
目前智能音响市场保持较快增速增长, 增速是手机三倍。根据 IDC 报告, 在 2014 年, 全球无线音响市场规模为 69 亿美元, 而到了 2022 年, 这一数字将达到 385 亿美元。在智能家居和智能音乐催化下, 市场尚具有较大提升空间!

图 17: 家庭多媒体设备中无线音箱增速最快



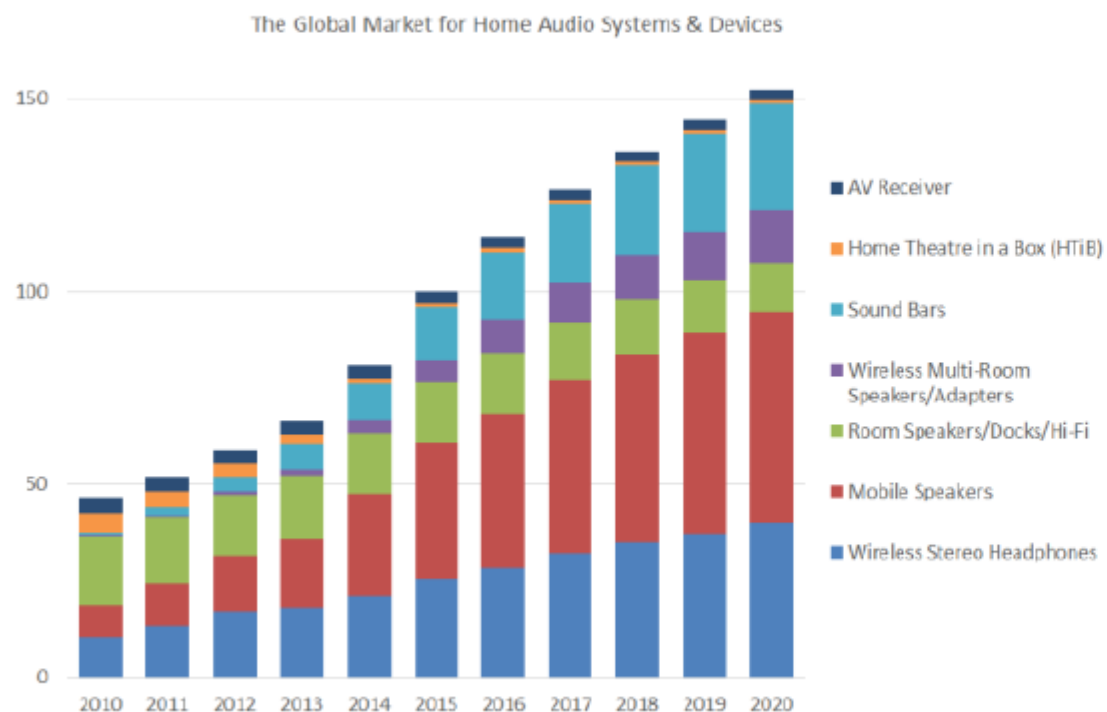
资料来源: Future Sorce、安信证券研究中心

图 18: 全球无线家庭音响出货量



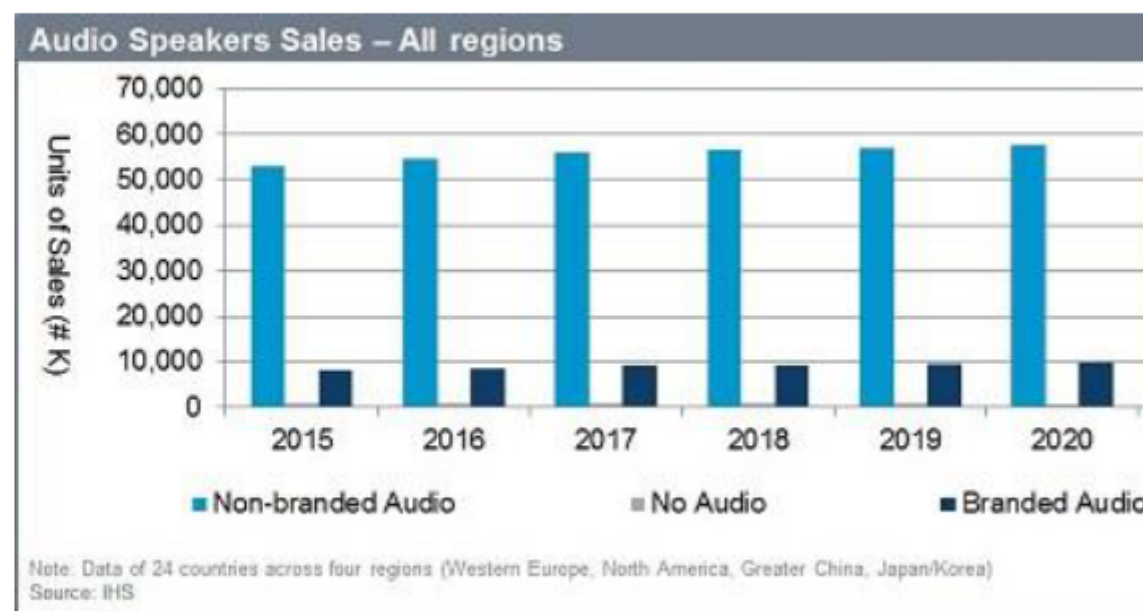
资料来源: HIS、安信证券研究中心

图 19: 家用音响设备细分零售市场价值 (亿美元)



资料来源: SAR INSIGHT & CONSULTING、安信证券研究中心

图 20: 品牌音响设备与非品牌音响设备出货量



资料来源: IHS、安信证券研究中心

中国是音响的制造和出口大国。2014年中国主要电子音响产品总产值约为2,774亿元,比2013年增长15.54%。该行业受国内外宏观经济情况影响较明显,波动较大,细分行业较多,情况差异大。随着人均收入的增长,消费能力的提高,音响产品逐步向二三级市场渗透,未来国内市场还有很大的需求空间。

图 21: 中国主要电子音响产品产值及增长趋势



资料来源: 中国产业信息网、安信证券研究中心

图 22: 中国主要电子音响产品出口额及增长趋势

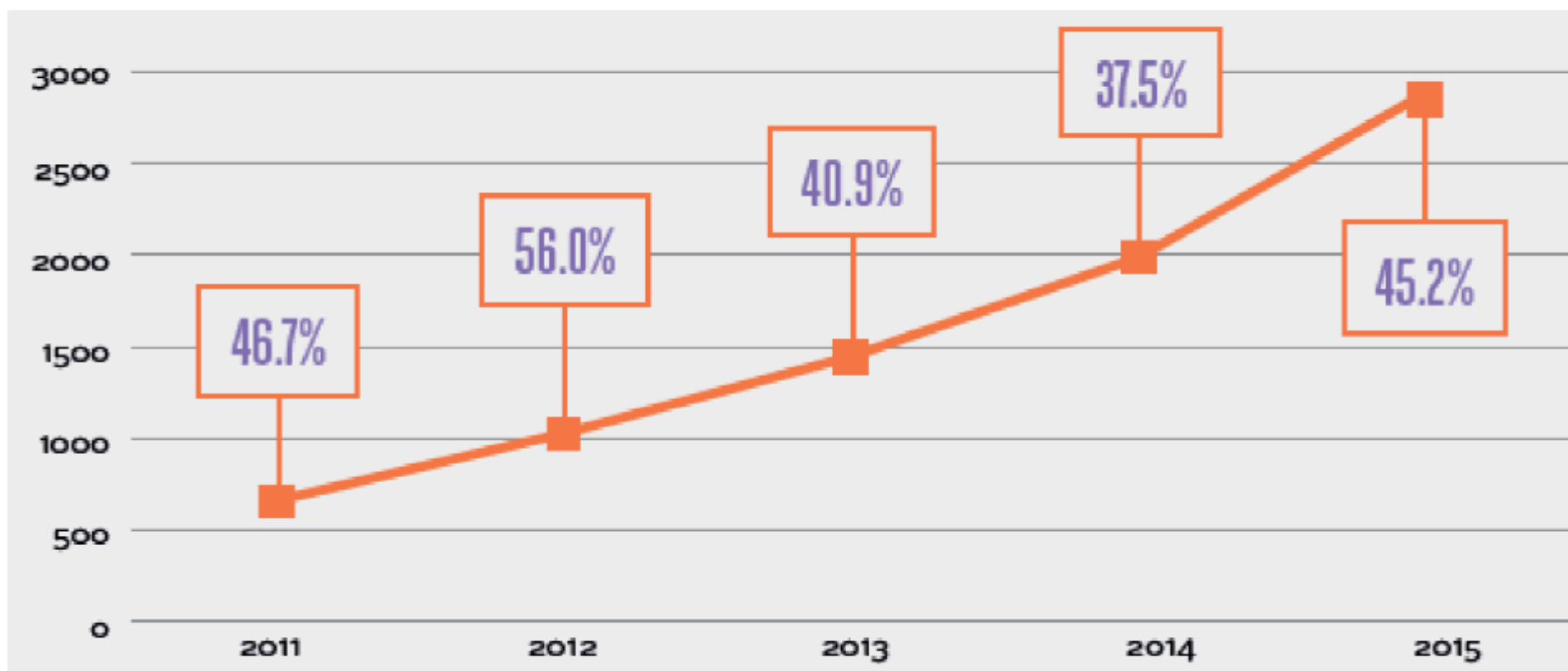


资料来源: 中国产业信息网、安信证券研究中心

从数据看流媒体已成音乐产业增长最快一环，付费的人群呈现跨越式增长。2015 年增长率达到 45.2%，总额达 29 亿美元，在过去 5 年间增长了约 4 倍。得益于智能手机用户的大规模增长，高品质音乐订阅服务的增加，以及相关乐迷向正版音乐服务的迁移，流媒体占全球音乐行业收入的比例从 2014 年的 14% 增长到了 19%。流媒体的收入占据了数字音乐行业收入的 43%，并且很快就要超过下载服务所带来的收入（45%），成为数字音乐的主要收入来源。近几年，付费订阅服务呈现明显增长的态势，据估计目前有 6800 万人正在为订阅音乐付费服务。这个数字在 2014 年的时候还只是 4100 万人，而在 2010 年的时候仅有 800 万人。



图 23: 2011-2015 年流媒体规模及其增长率



资料来源: IFPI、安信证券研究中心

从产业了解看,今年下半年预计将有多家互联网厂商进入智能音响领域,行业大趋势确定,互联网巨头纷纷布局语音交互硬件产品,音响类产品成主流方案。在自然语言处理技术支持下,以 Amazon、Google 为首的互联网巨头纷纷进军语音交互硬件产品,我们预计苹果、腾讯、阿里、百度都将在音响领域布局,音响是声音的输入与输出最佳的载体,主要由扬声器、麦克风以及处理芯片构成。这类产品首先具有基础音频输出功能,可以通过语音交互控制音乐播放。而它们的核心能力则是低噪声环境下的语音识别能力,这一能力能够很好解决目前智能家居产品在交互信息入口方面的能力。因而这种音响类产品能够被方便地引入智能家居控制系统,辅助用户进行高效便捷的控制。

表 2: 巨头布局音响类产品情况

名称	公司	图片	功能
google home	谷歌		支持语音交互；控制闹钟、购物清单；控制音响、灯光等室内家电，通过手机应用控制音频播放、视频播放，连接汽车
Echo	亚马逊		支持语音交互；搜索、音乐播放、产品或服务订购，支持智能家居及第三方应用
叮咚	京东		支持语音交互；可以成为用户的音频助理，完成百科查询、讲故事、聊天等同时，该产品可以让用户通过语音控制接入京东微联的智能产品，成为智能家居的全新交互入口。
Go Smart	JBL		支持语音交互；通过语音识别可以直接的播放在线音乐资源
苹果智能音箱	苹果		内置 Siri 语音助手，能够作为苹果 Homekit 家居平台的控制入口，同时能够接入第三方应用。目前尚未有相应产品上线。

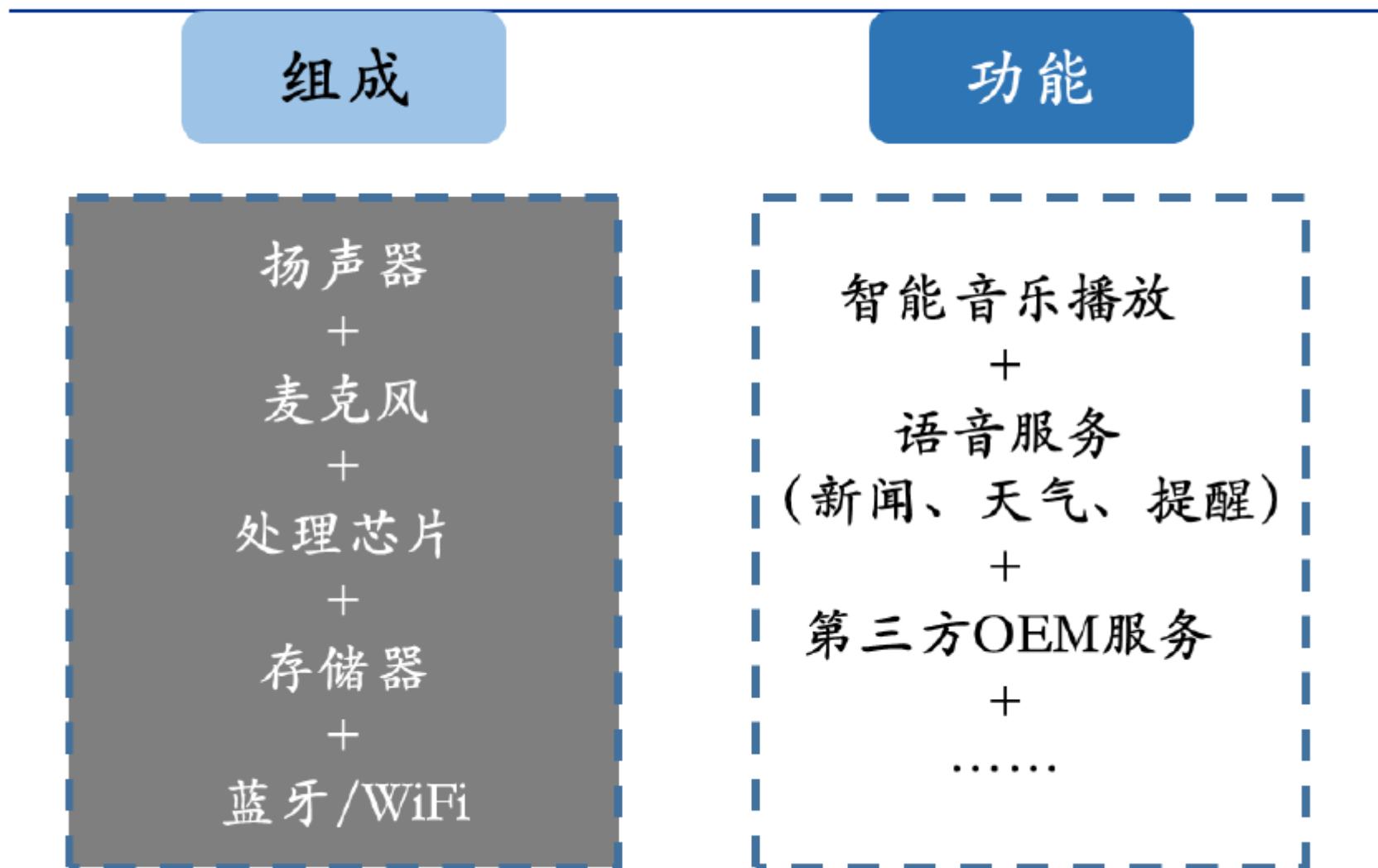
资料来源：安信证券研究中心整理

### 1.3.互联网的入口还在于声音的输入

随着 Echo 出现，all in one 音响类产品雏形开始显现，语音交互成为最直接的控制方式，Amazon Echo 的开机词 Alexa 是最佳代表，声音的输入是互联网入口的新价值所在。智能音响产品的构成包括扬声器、麦克风、处理芯片、存储芯片和 WiFi/蓝牙组件等部分，在音乐播放这一基础功能之外具有基于语音识别的多重功能，通常能够作为智能家居和物联网

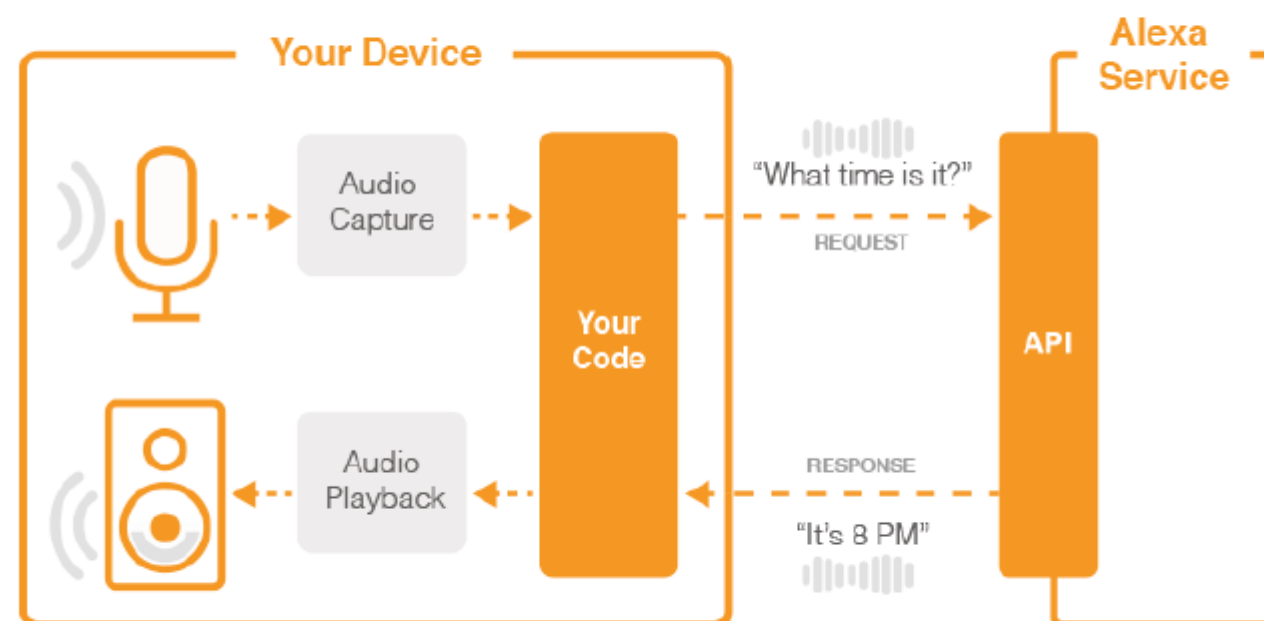
入口，并能够提供在第三方集成的 OEM 语音服务。工作模式方面，一个典型的工作流包括用户下达指令或提问、智能音响设备进行处理以及 反馈三个部分。反馈既可以是设备直接输出语音，也可以通过第三方设备进行图像或文字输出。

图 24: All in One 音响产品雏形渐显



资料来源：安信证券研究中心

图 25: Alexa Service



资料来源: Amazon、安信证券研究中心



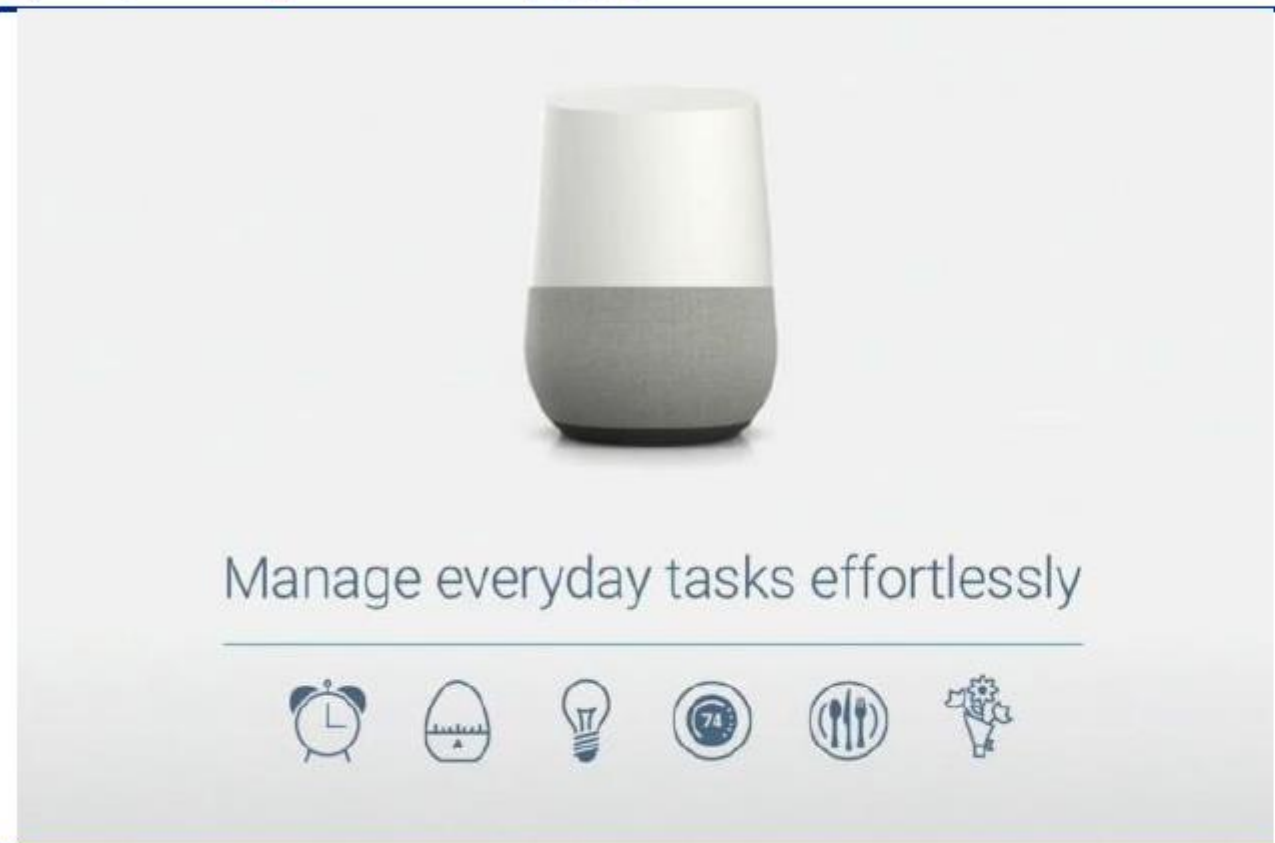
图 26: 智能音响类产品 workflow 模式



资料来源: 安信证券研究中心整理

Google 布局语音交互技术已久, 近期也开始参与硬件设备领域。在 2016 年度 I/O 开发者大会上, Google 发布了配有内置扬声器的语音交互设备 Google Home。通过 Google 语音助手 Google Assistant, 用户可以利用 Google Home 播放音乐, 查询天气、方向、航班等各类信息。

图 27: Google Home 示意图



资料来源: Google、安信证券研究中心

图 28: Google Assistant 语音助手辅助聊天



资料来源: Google、安信证券研究中心

相较亚马逊 Echo, Google 优势在于“语音技术+搜索数据”。一方面谷歌开发语音识别技术多年,另一方面还拥有十多年积累下来的庞大搜索数据,这意味着用户在进行语音搜索时能得到更好、更快以及更精确的结果。此外 Google 还可以提供配套的日历、电子邮件以及在线云存储服务 Drive 等服务。

整合 Nest, 语音交互入口服务智能家居。Google 早前于 2014 年以 32 亿美元收购智能家居设备制造商 Nest, 通过 Google Home 这一语音交互入口, Google 能够将 Google Assistant 与 Nest 设备进行融合, 将软件与硬件整合到一个中枢, 构成物联网生态圈辅助用户迅速访问不同服务。

图 29: Nest 智能家居平台



资料来源: Nest、安信证券研究中心

苹果也许会开发下一款智能音箱设备。科技博客 The Information 消息称苹果计划推出一款内置 Siri 语音助手的智能音箱，并且能够作为苹果 Homekit 家居平台的控制入口。这款设备配备了至少一个扬声器，支持 AirPlay，内置了网络连接功能。The Information 同时表示苹果将发布一款软件开发工具包，让第三方开发者能够将产品与 Siri 进行整合从而能够在第三方应用中使用 Siri。



图 30: Siri 此前已经与 yelp 等公司合作显示相关服务



资料来源: 互联网、安信证券研究中心

## 2.智能语音交互技术成熟，应用有望扎根音响 2.1.智能语音技术发展驱动语音交互快速渗透

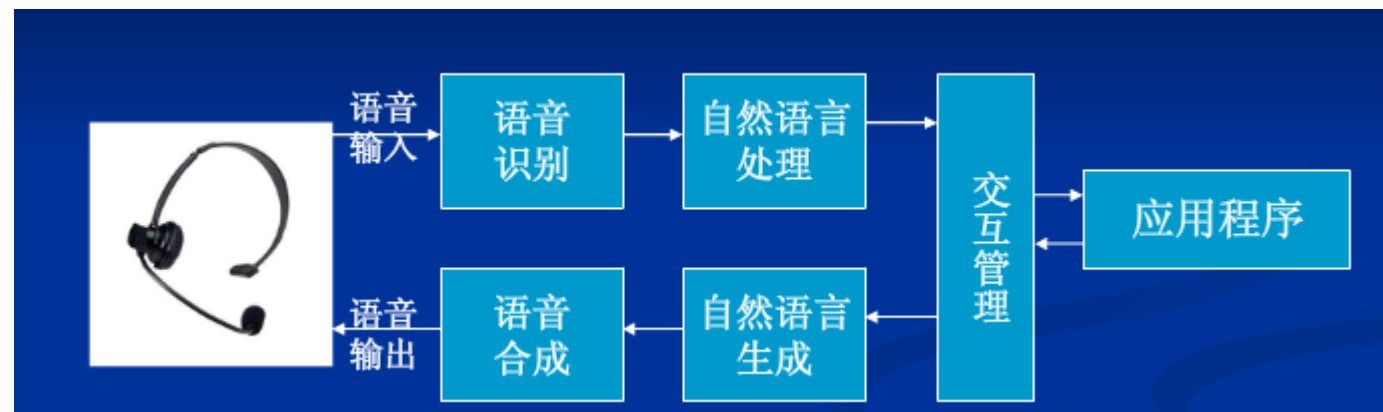
智能语音技术包括了识别和交互环节，融合了多种人工智能尖端技术，是真正的黑科技。在美国政府关于限制发放签证的“Technology Alert List”中，智能语音赫然在列，与核武器、火箭技术等同在“黑名单”中，可以说称得上是真正的黑科技。狭义上，智能语音技术主要是指语音识别，广义上，智能语音则可以分为识别和交互两大技术环节，其中交互环节又可以分为理解、搜索、生成三个子环节（服务、信息）。

我们以度秘为例，分析完整的人机语音交互流程背后都有哪些人工智能技术在做支撑。当用户向度秘（duer）提出需求：“找一家附近好吃的餐馆”时，度密机器人助理第一步是要将用户的语音输入进行识别，转换成计算机语言，然后基于自然语言理解技术理解我们的需求，然后在通过搜索技术去寻找餐馆（比如大众点评上评价是5星的餐厅），最后再通过



语音合成或者自然语言生成（NLG）技术，将最后的搜索结果反馈给我们。上述流程仍然是一个简化版本，实际上，为了更好的去理解用户的意图，度密还必须学会发问，将模糊的用户意图具体化（本例中好吃的定义是抽象的）。

图 31：实现人机语音交互的流程图



资料来源：百度、安信证券研究中心

表 3：智能语音的核心技术概览

语音技术（识别）	功能	智能化技术（理解）	功能	大数据技术（搜索）	语音生成	功能
语音识别（ASR）	把人讲的话转成文字	自然语言理解（NLU）	了解用户的意图	各种垂直领域的的数据服务（天气、航班、餐馆、……）	自然语言生成（NLG）	把查询到的信息反馈给用户
声纹识别和认证	从众多候选人中确认发音人的身份或者认证发音人是否为某个特定人。在应用中，可以通过这些技术唤醒应用、解锁密码等。	多轮对话管理（DM）	通过提问来进一步明确用户意图	垂直搜索或通用搜索（OneSearch、神马搜索、钱包搜索等等）	语音合成（TTS）	把文字说出来
情感识别	通过用户讲话的语气语调、用词等线索，判断用户是否有不满或处于生气激动状态	场景感知（contextual awareness）	根据前面对话内容、当前地理位置、收到的邮件短信中的内容来更好了解用户意图	知识问答（基于知识库、知识图谱的问答）		
远场语音识别		个性化	通过用户的过往行为或选择，做模型的自适应，为用户提供更有开放式聊天针对性的服务			
		情感识别	通过用户讲话的语气语调、用词等线索，判断用户是否有不满或大规模机器学习模型的训练处于生气激动状态			
语音服务的数据闭环						

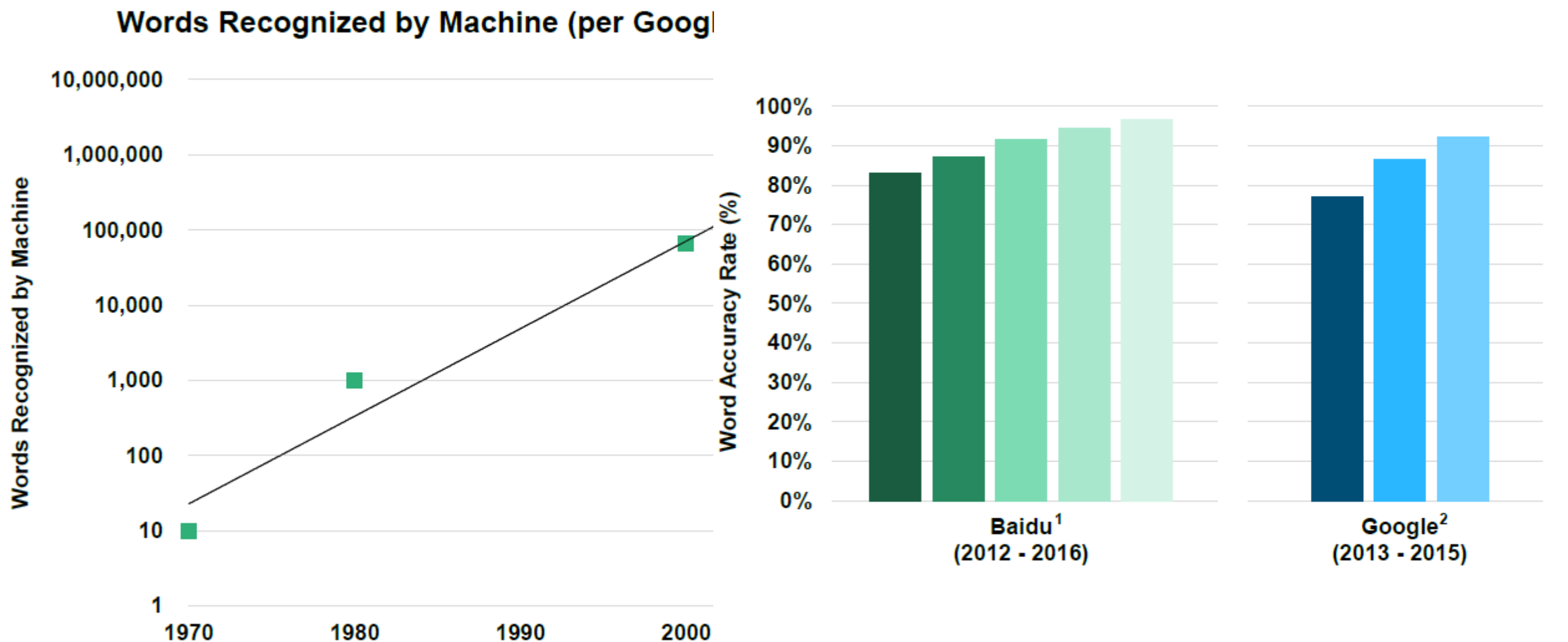
资料来源：阿里巴巴，安信证券研究中心整理

人机语音交互的应用关键在于准确率与响应时间，语音识别技术日益成熟，即将突破量变到质变的临界点。随着自然语言处理技术及硬件设备的发展，语音识别的准确率与响应时间持续得到改善，驱动语音交互方式快速渗透。其中语音识别技术经过 46 年的发展，已经能够达到 90%~95% 的准确率（低噪音环境下）。目前正在向 99% 准确率以及高噪音环境

应用发展。百度首席科学家吴恩达表示“现在语音识别可能已达到 95% 的准确度，但很多人没意识到 95%的准确度到 99%的准确度带来的不是量变是质变，是从你偶尔使用语音变到常常使用做到更自然。做到 99%准确，将彻底改变人与设备的交互。”

图 32：语音识别准确率持续提升

图 33：百度、谷歌及其他语音识别平台准确率均超 90%



资料来源：《Voice Technology and Research Lead》、安信证券研究中心

资料来源：Baidu, Google, VentureBeat, SoundHound、安信证券研究中心

**语音识别技术应用不断拓展：**早在二战期间，欧美的一些实验室已经开始进行语音识别相关的科学研究。1952年的贝尔研究所，Davis等人研制了世界上第一个能识别10个英文数字发音的实验系统。1960年英国的Denes等人研制了第一个计算机语音识别系统。而语音技术真正进入工业界则是在2000年左右，摩托罗拉和飞利浦等电子厂商在手机里加入了语音拨号软件。随着语音识别技术的不断成熟，其应用场景也在不断增加。2008年，谷歌把语音识别和搜索相结合推出了Google Voice Research，通过把声音转成文字，有文字之后再调后面的搜索引擎，整个识别引擎实现了正向反馈的机制，识别性能不断提高。2010年，苹果推出了革命性的个人语音助手siri。2014年，亚马逊推出了一款支持语音控制的Echo智能音箱，巨头们开始挖掘语音技术在智能家居场景中的应用潜力。得益于其在人机交互方面的优势，智能硬件、语音助手、语音搜索等领域均已开始加速上线语音识别技术。

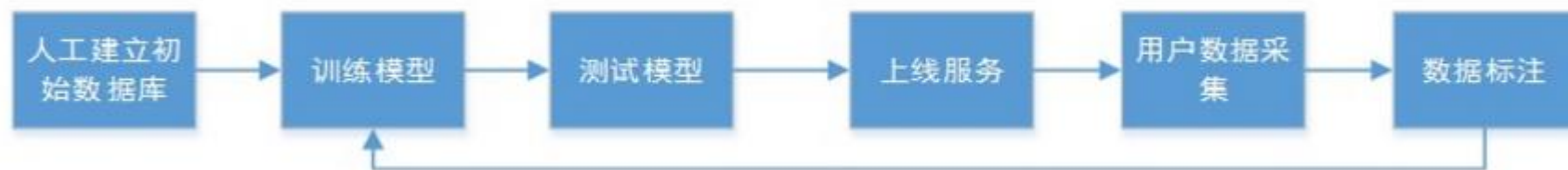
图 34：语音技术应用发展历程



资料来源: 安信证券研究中心整理

**发展优秀智能语音技术的关键——构建语音服务的数据闭环。**语音识别依赖于 science 和 engineer 的结合。Science 推动语音识别基本技术的升级, engineer 扩充语音识别的场景和语言。一个高准确率语音识别模型需要不断的训练提高识别准确率, 通过构建语音服务, 获取用户语音数据, 再通过海量语音数据优化迭代语音识别模型, 构建一个语音服务的数据闭环是发展语音技术的关键, 应用与技术双向促进将形成可持续发展的产业正循环。

图 35: 标准的语音识别模型流程

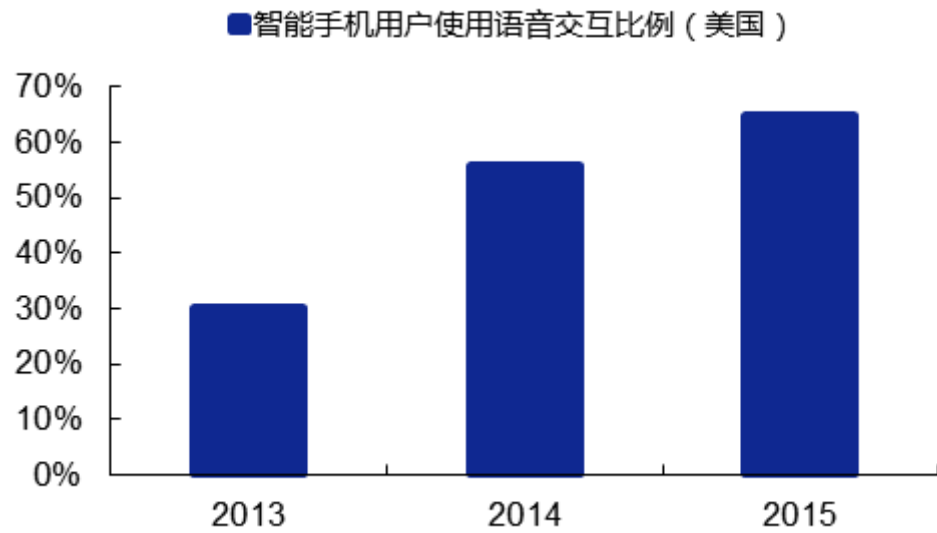


资料来源: 阿里巴巴、安信证券研究中心

**语言识别技术和硬件发展驱动语音交互渗透率迅速提升。**根据研究机构 Thrive Analytics 统计, 2015 年美国智能手机用户使用语音交互的比例较 2013 年提升了一倍以上, 而他们的用户调查同时显示软件/技术的改善是选择使用语音交互的最主要原因。

图 36: 2015 年 65% 的美国智能手机用户在使用语音交互

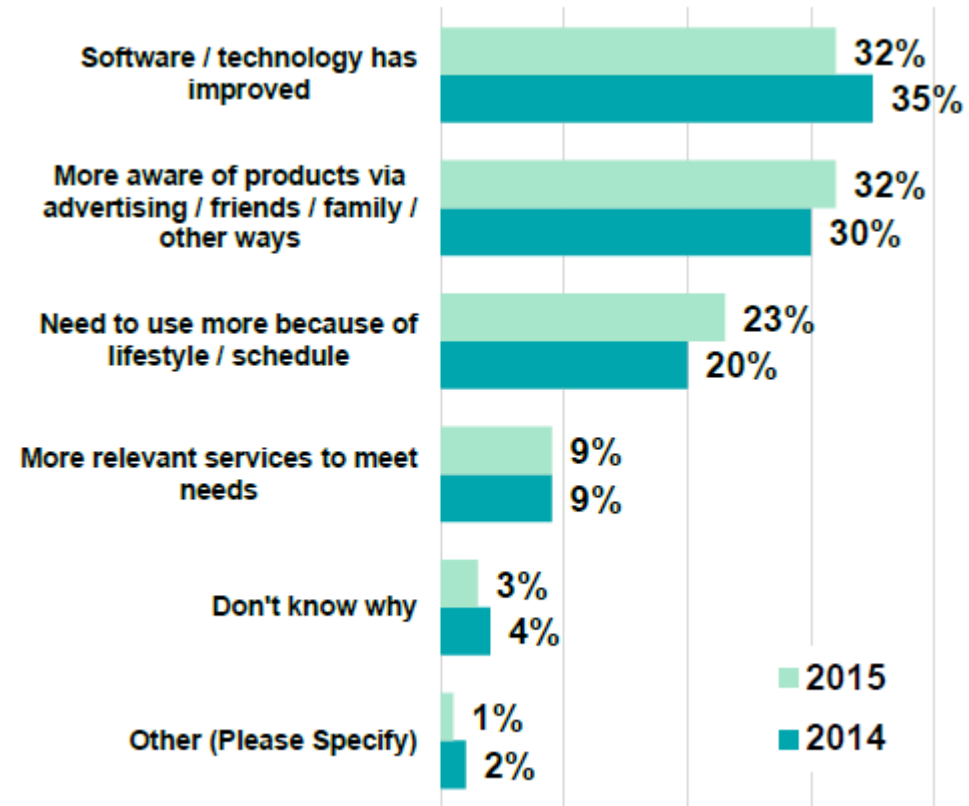
图 37: 技术提升是选择语音交互的主要原因



资料来源: Thrive Analytics、安信证券研究中心

在搜索服务方面，语音搜索在各平台占比迅速提升。根据 Google Trends 统计，自 2008 年 iPhone 及谷歌语音搜索推出以来语音搜索增长超 35 倍；近年来百度语音的占比提升同样十分迅速，由于智能手机键盘较小，汉字输入难度较英文输入更高，因而大量用户选择使用百度语音识别功能以及文本转换语音功能进行搜索和接收信息。根据百度披露数据，自 2014 年 Q2 以来语音输入增长 4 倍以上，语音输出增长 26 倍以上。同时百度人工智能专家吴恩达预测，2020 年语音及图像搜索占比有望达到 50%。

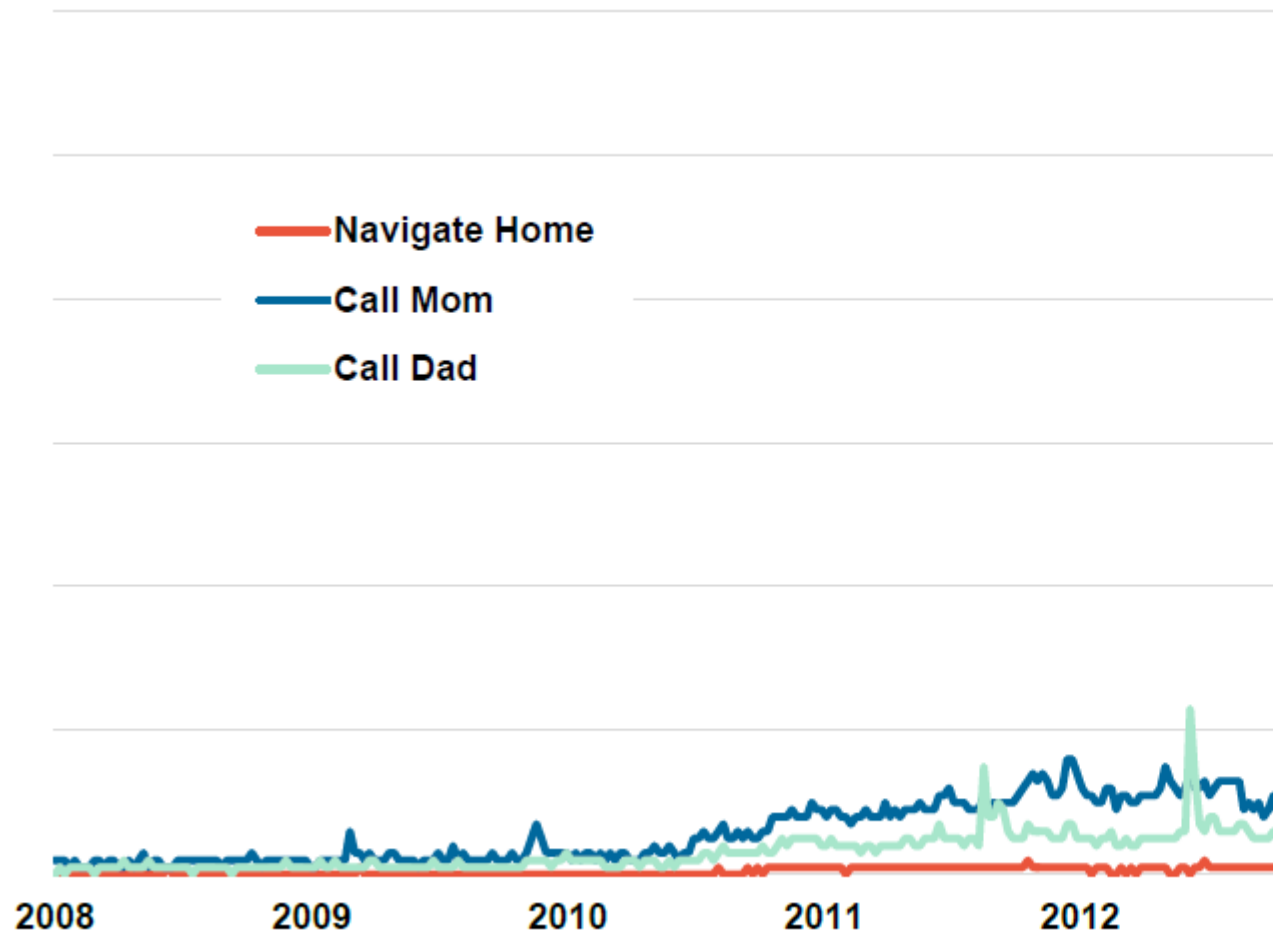
图 38: Google 语音业务增长迅速



资料来源: Thrive Analytics、安信证券研究中心

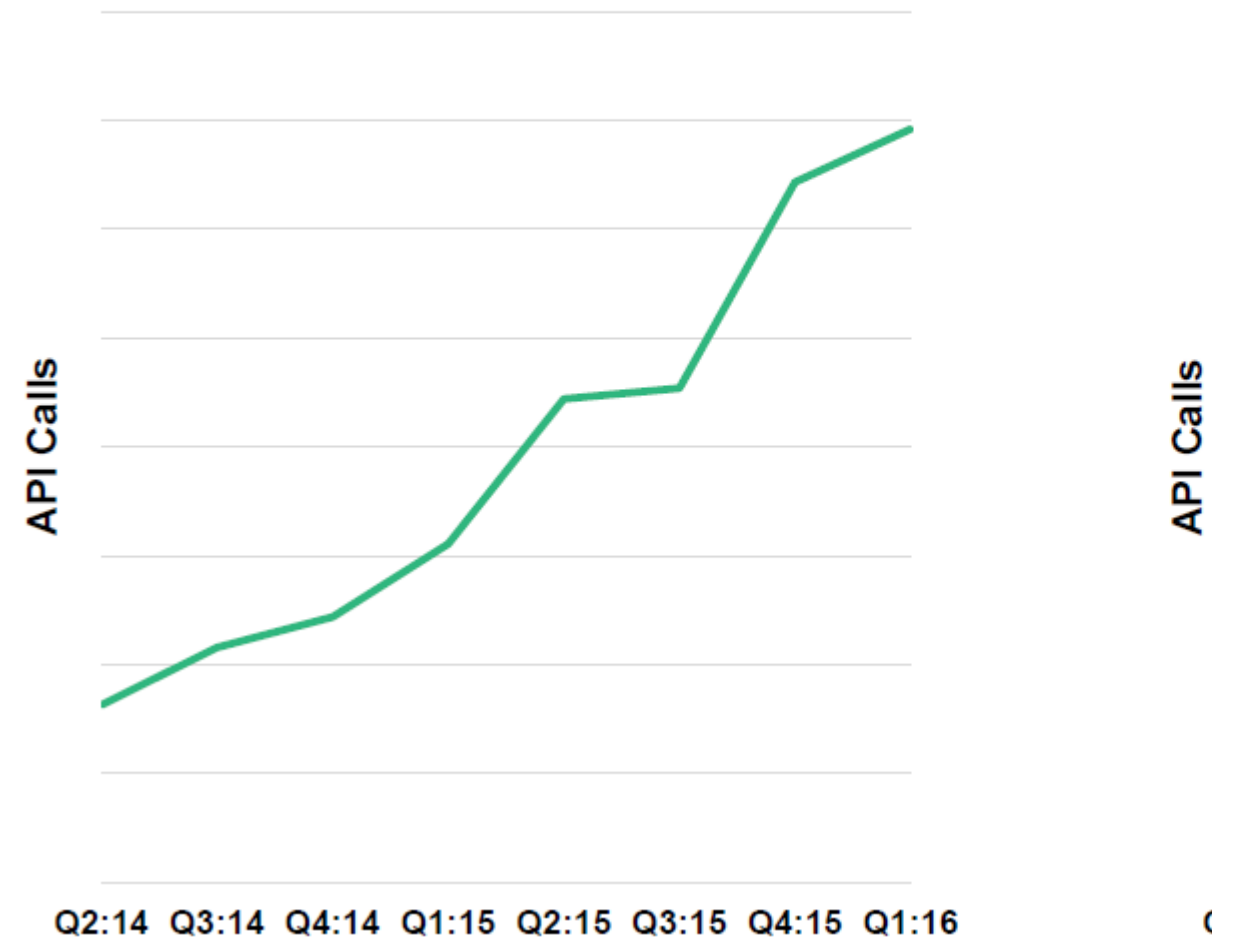
图 39: 百度语音输入两年来增长超 4 倍，输出超 26 倍





资料来源：Google Trends、安信证券研究中心

Baidu Speech Recognition Daily Usage by API Calls, Global, 2014 – 2016<sup>1</sup>

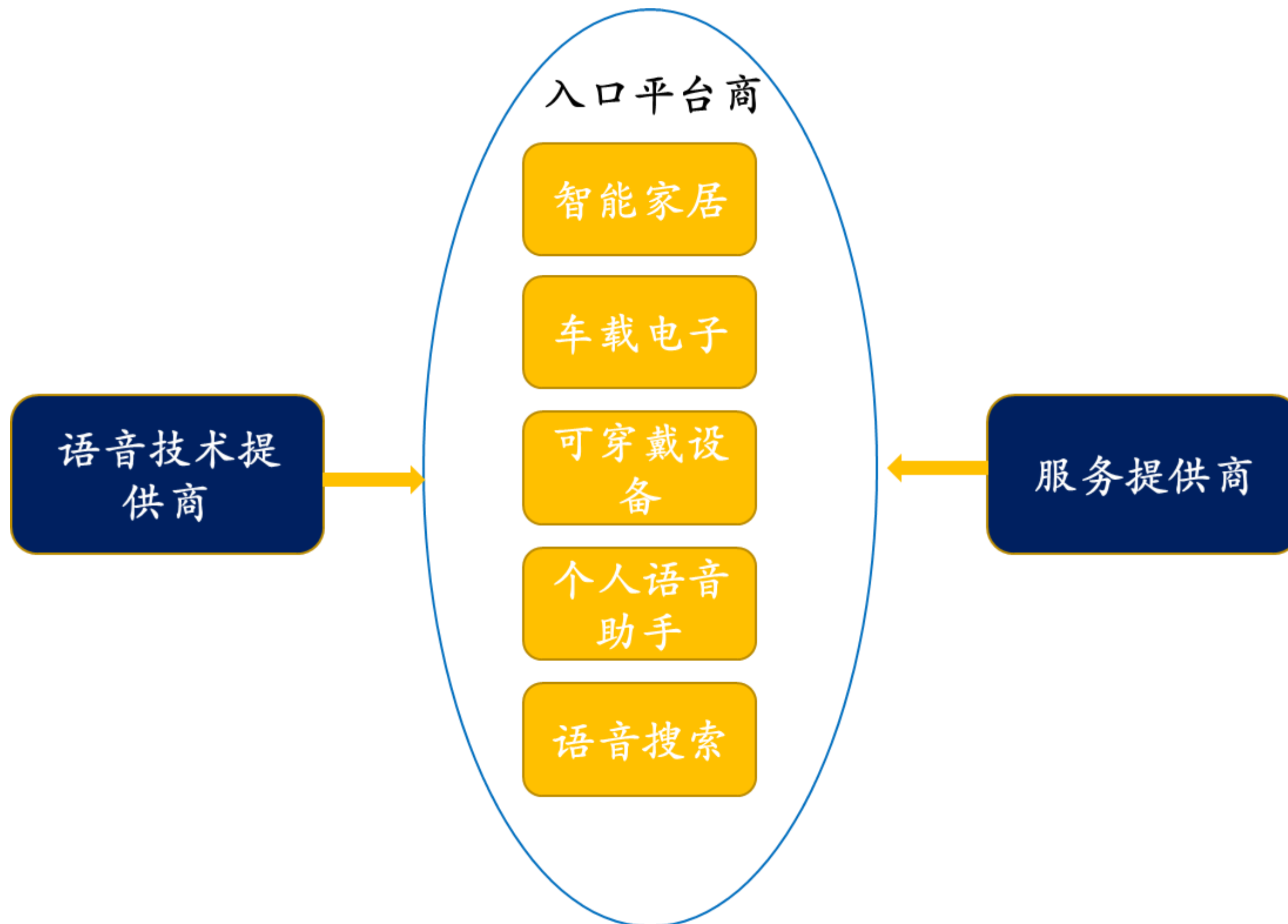


资料来源：百度、安信证券研究中心

## 2.2.智能语音产业链成熟，巨头布局完善

智能语音产业链分为三部分，分别包括智能语音技术提供商、后端服务提供商以及入口平台商。其中，智能语音技术提供商着重语音合成、语音识别等基础技术的研发，为下游的入口层提供基础技术支撑；以智能家居、个人语音助手、语音搜索为代表的入口平台则继续整合后端服务提供商（比如电商、本地生活等）为用户提供一站式的完整智能语音人机交互体验。

图 40：智能语音产业链

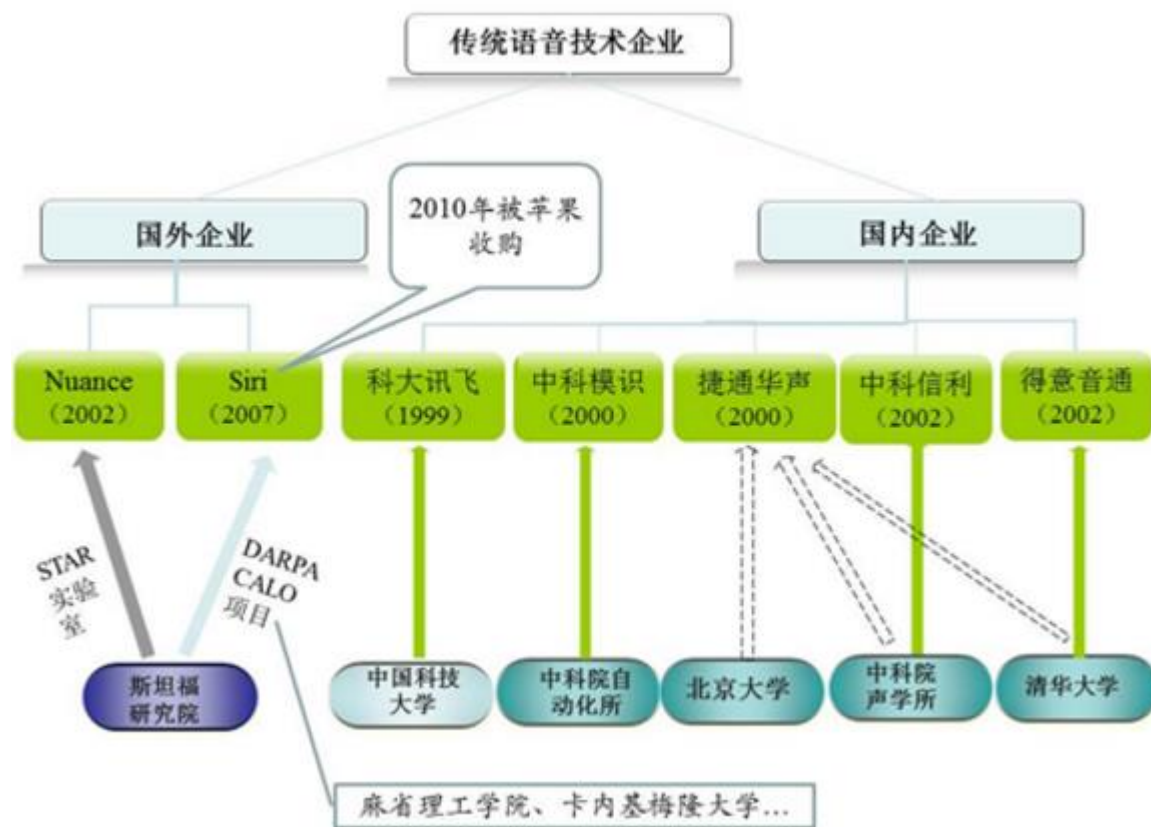


资料来源：安信证券研究中心整理

智能语音产业参与者可以分为两大流派：从科研实验室走出来的独立语音技术提供商以及希望抢占下一代入口的互联网巨头。作为科研驱动型行业，智能语音技术研究最早起源于贝尔实验室，斯坦福、卡内基梅隆等学校的研究为智能语音产业发展奠定了坚实的基础，苹果 Siri 的雏形便是源自于斯坦福研究院联合麻省理工学院、卡内基梅隆大学等多家机构

承担的美国国防高级研究计划局（DARPA）的 CALO 项目。Nuance 也是源于斯坦福 研究院的 STAR 实验室。国内的语音技术公司同样大多都脱胎于中国科学院声学所、中国科学院自动化所、中国科技大学、清华大学、北京大学等科研机构。而随着语音日渐成为人机交互的重要方式，互联网巨头们为了争夺下一个入口，在智能语音市场也展开了疯狂的“军备竞赛”，以亚马逊、谷歌、苹果、微软、百度、腾讯、搜狗为代表的巨头纷纷通过并购与自研推出自己的语音产品，加大对语音市场的争夺。

图 41：传统语音技术企业概览



资料来源：中国语音产业信息网、安信证券研究中心

表 4：巨头纷纷布局语音技术

公司	时间	收购或者参股公司	功能
苹果	2010	Siri.inc	虚拟助手
	2013	Novauris Technologies	Novauris 是一种可识别整个短语的语音识别技术，这种技术并非简单识别单个词句，而是试图利用超过 2.45 亿个短语的识别辅助理解上下文
	2015	VocalIQ	能够帮助计算机与用户进行更为自然的对话。
谷歌	2011	SayNow	SayNow 可以把语音通信、点对点对话、以及群组通话和 Facebook、Twitter、MySpace、Android 和 iPhone 等等应用等整合在一起
	2011	Phonetic Arts	Phonetic Arts 可以把录制的语音对话转化成语音库，然后把这些声音结合到一起，从而生成听上去非常逼真的人声对话。
	2013	Wavii	Wavii 擅长“自然语言处理”技术，可以通过扫描互联网发现新闻，并直接给出一句话摘要及链接。
	2015	出门问问	拥有自主语音识别、语义分析、垂直搜索技术
亚马逊	2011	Yap	Yap 成立于 2006 年，主要提供语音转换文本的服务
	2012	Evi	加强语音识别在商品搜索方面的应用

	2013	Ivona Software	主要做文本语音转换
Facebook	2013	Mobile Technologies	公司的产品 Jibbiggo 允许用户在 25 种语言中进行选择，使用其中一种语言进行语音片段录制或文本输入，然后将翻译显示在屏幕上，同时根据选择的语言大声朗读出来。
	2015	Wit.ai	Wit.ai 允许用户直接通过语音来命令移动应用程序 (iOS、Android 等平台)、穿戴设备和机器人以及几乎任何你可以想到的智能设备。

资料来源：互联网资料，安信证券研究中心整理

### 2.3. “效率优势+场景拓展”决定语音交互发展趋势

回顾人机交互发展，实际上是一段不断改造机器解放人的历程。最早期的电脑，键盘是唯一的输入设备，随着图形界面 GUI 的出现，形成“键盘+鼠标”的组合，然而精准点击鼠标和敲击键盘仍然需要较高的学习成本。其后，设备终端的越做越小进一步解放用户，手机触屏的出现真正摆脱了键鼠这一中间介质，做到所触即所得。尽管如此，利用触控操作手机或平板，仍然需要用户的手和眼睛全程参与。

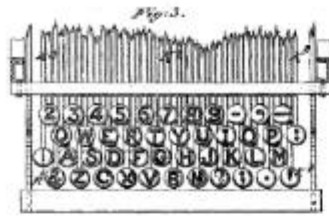
语音交互进一步解放感官，学习成本小。语音交互这一方式进一步解放手和眼睛，减少感官占用从而改善体验。此外直接通过语言信息进行交互，在空间上拓展了与智能设备的距离限制，这一点智能家居方面优势明显，能够真正实现随时随地“Always-Online”的智能体验。相比键盘、鼠标甚至触屏，语音作为交互的学习成本也要更小。

图 42：“Touch 1.0（键盘）àTouch 2.0（鼠标）à Touch 3.0（触屏）à 语音”发展历程





Punch Cards for Informatics  
1832



QWERTY Keyboard  
1872



Electromechanical Computer (Z3)  
1941



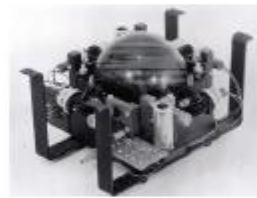
Electronic Computer (ENIAC)  
1943



Paper Tape Reader (Harvard Mark I)  
1944



Mainframe Computers (IBM SSEC)  
1948



Trackball  
1952



Joystick  
1967



Microcomputers (IBM Mark-8)  
1974



Portable Computer (IBM 5100)  
1975



Commercial Use of Window-Based GUI (Xerox Star)  
1981



Commercial Use of Mouse (Apple Lisa)  
1983



Commercial Use of Mobile Computing (PalmPilot)  
1996



Touch + Camera - based Mobile Computing (iPhone 2G)  
2007



Voice on Mobile (Siri)  
2011

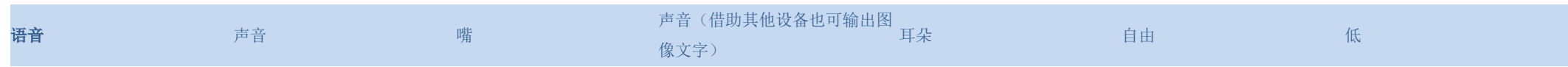


Voice on Connected / Ambient Devices (Amazon Echo)  
2014

资料来源：“History of Computer Interfaces”、安信证券研究中心

表 5: Voice In Voice Out, 语音交互进一步解放人类

交互方式	输入形式	输入器官	输出形式	接收器官	空间限制	学习成本
键盘	文字	手	文字	眼睛	大	高
鼠标	伪触摸	手	文字/图像	眼睛	大	高
触摸屏	触摸	手	文字/图像	眼睛	相对自由	较低



资料来源：安信证券研究中心

**语音交互方式除了能够改善用户体验，还具有明显的速度优势。**文字输入的世界纪录是 212 单词/分钟，而说话的最快纪录超过 600 单词/分钟。在正常人的文字及语音输入效率上这个比例仍然成立。在自然语言处理技术发展下语音识别的处理时间大幅减少，速度优势进一步扩大！此外语音识别还具有上下文驱动（基于历史问题/交互理解情景）、低成本（仅需麦克风、扬声器和处理 器）等特点。

图 43：语音交互特点

图 44：语音交互技术已经全面渗透到各项互联网应用

语音交互  
改善用户体验

简单  
快速  
个性化  
上下文驱动

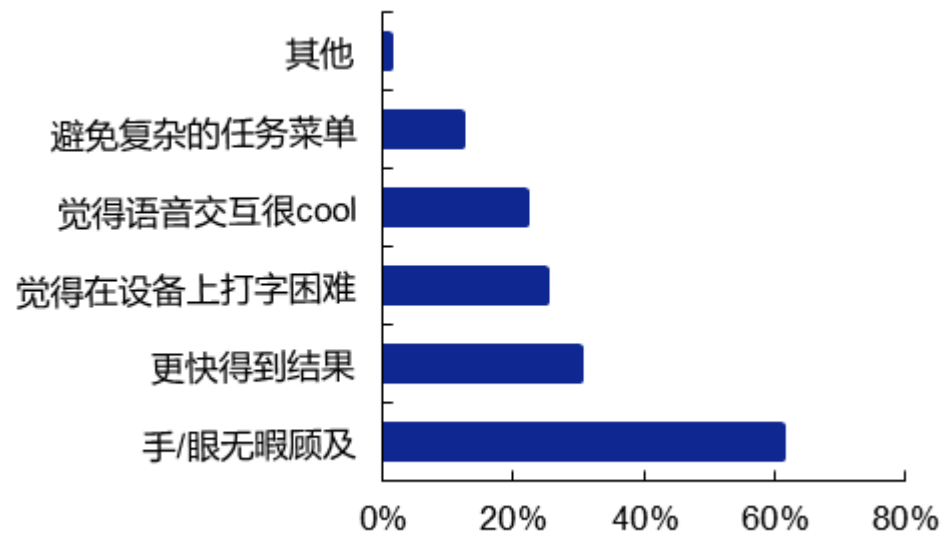
资料来源：《2016 Internet trend report》、安信证券研究中心

图 45：使用语音交互的主要原因（美国）

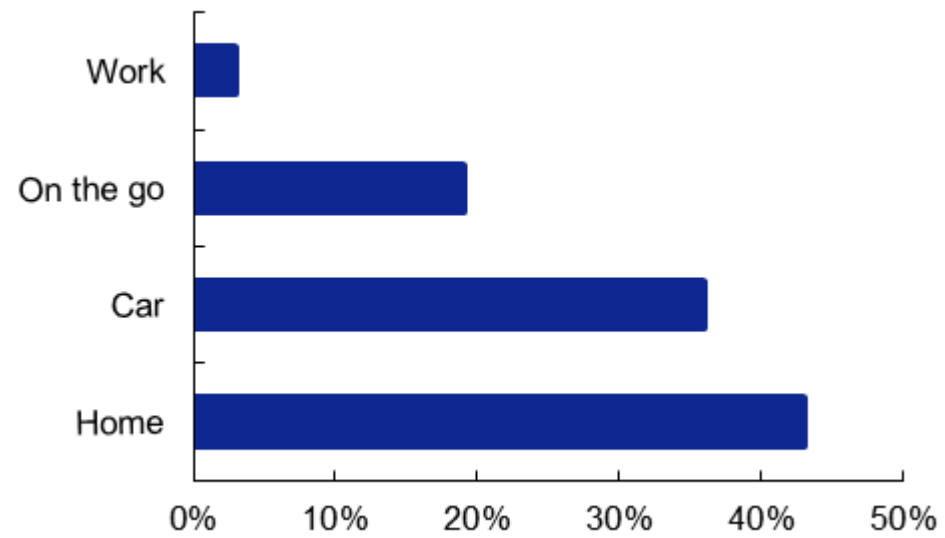


资料来源：互联网、安信证券研究中心

图 46：使用语音场景的主要原因（美国）



资料来源：《Intelligent Voice Assistants Research Report》、安信证券研究中心



资料来源：《Intelligent Voice Assistants Research Report》、安信证券研究中心

### 3. 智能音响不仅仅是互联网入口，更是物联网的入口 3.1. 智能音响其实是智能家居的一部分

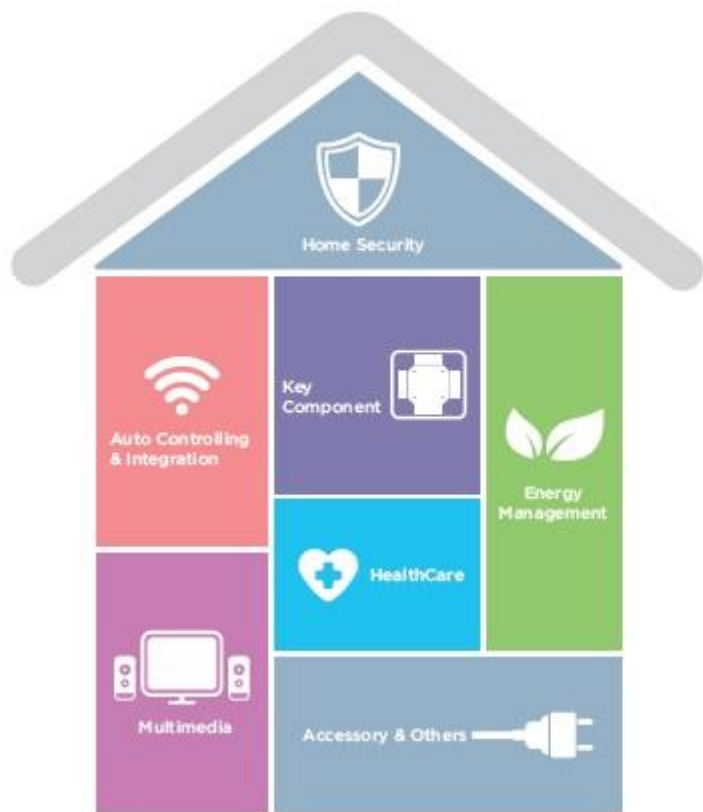
智慧家居将把家庭的安全、娱乐、饮食、健康等结合起来，用户面对的将不再是各种家电和移动设备的零散组合，而是一个有机的整体。这个有机整体需要有一种新的人机交互模式，从而使用户更加便捷、舒适地调用智慧家庭设备。因而，智慧家庭的发展势必将由被动的用户输入转化为主动捕捉用户的语言和动作，此时具备声音控制和播放的家庭音响可以发展成为智慧家庭的入口和中心。

互联网巨头纷纷布局语音交互硬件产品，音响类产品成主流方案。在自然语言处理技术支持下，以 Amazon、Google 为首的互联网巨头纷纷进军语音交互硬件产品。目前的主流方案为智能音响类产品，主要由扬声器、麦克风以及处理芯片构成。这类产品首先具有基础音频输出功能，可以通过语音交互控制音乐播放。而它们的核心能力则是低噪声环境下的语音识别能力，这一能力能够很好解决目前智能家居产品在交互信息入口方面的能力。因而这种音响类产品能够被方便地引入智能家居控制系统，辅助用户进行高效便捷的控制。

图 47：智慧家庭的范畴

图 48：智能家居概念图





-  **Home Security:** Video Surveillance, Alarm Systems Door Locks & Access Control, Video Door Phones
-  **Auto Controlling & Integration:** Intelligent Control Panels, Gateways/Hubs, Software Platform
-  **Energy Management:** Thermostats & Sensors, HVAC Control, HEMS, Smart Grid
-  **HealthCare:** Nursing Care, Exercise and Health Monitoring...
-  **Multimedia:** Speakers, A/V System, Home Theater...
-  **Key Component:** SoC/ISP, Networking IC/Chips
-  **Accessory & Others:** Sockets, Switches, Cables, Structured Wiring



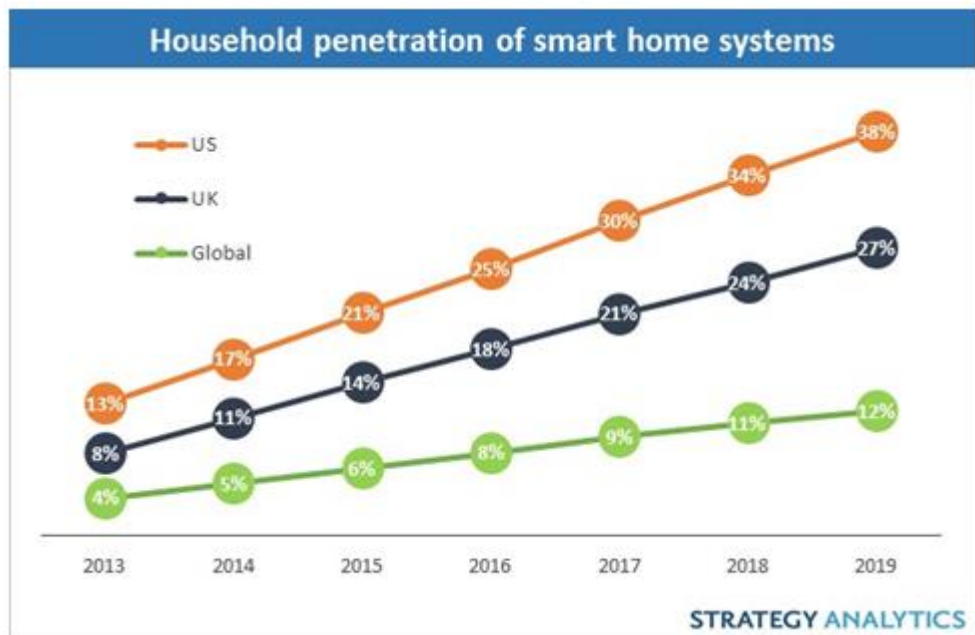
资料来源：互联网、安信证券研究中心

资料来源：互联网、安信证券研究中心

智能音响设备在智慧家居的带动下，有望发生革命性的变化。在 2019 年，智慧家庭的渗透率预计达 12%，市场规模超过 1500 亿美元，且随着智能化的普及，该市场仍然具有很大的增长空间。届时，作为入口的智能音响设备，计算能力将得到大幅度提升，“音响”属性将被削弱，家庭智能控制中心的作用会得到加强，并且伴随一波销售热潮占据智慧家庭的核心。

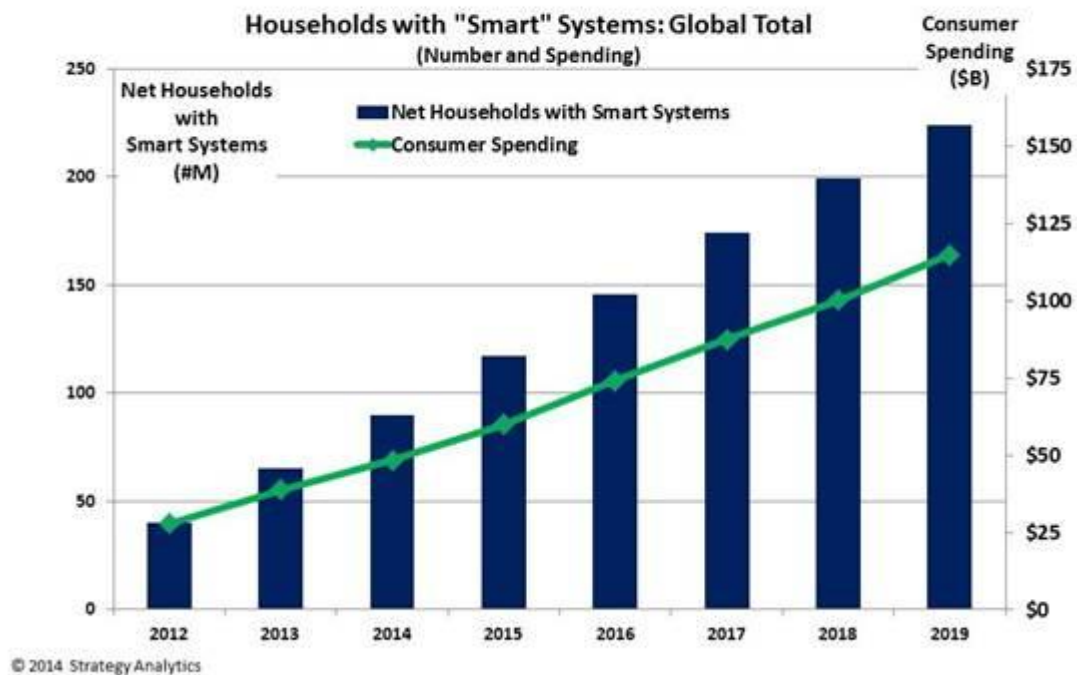
图 49：智慧家庭渗透率增长

图 50：智慧家庭市场价值增长



资料来源: strategy analytics、安信证券研究中心

### Global Smart Home Market will reach \$100 Billion by 2018



资料来源: strategy analytics、安信证券研究中心

## 3.2.全产业链的智能控制革命

从 Echo 到 DingDong, 智能音箱语音服务抢占未来智能家居的控制中心、核心交互入口。

自美国零售业巨头 2014 年发布智能音箱硬件 Amazon Echo 后, 国产网络零售企业京东不甘落后, 于 2015 年联手科大讯飞合作打造国产化智能音箱 DingDong。与 Echo 相似, DingDong 音箱的功能和应用领域也主要集中在音乐播放、语音讯息服务、智能家居控制等方面, 配套科大讯飞为其开发的语音交互服务, 可以通过语音指令控制符合京东微联标准的数百种智能家电。在国内智能家居生态尚未成熟的阶段, DingDong 智能音箱目前更多侧重“音箱”属性, 语音服务方面距离 Echo 还缺乏第三方应用协同和定制化应用, 目前京东发布国产化的智能音箱, 更多是看好语音交互模式的前景, 未来伴随生态搭建、硬件服务升级, DingDong 智能音箱语音服务将成为智能家居的控制中心、核心交互入口。

图 51: DingDong 智能音箱实物及相关功能介绍

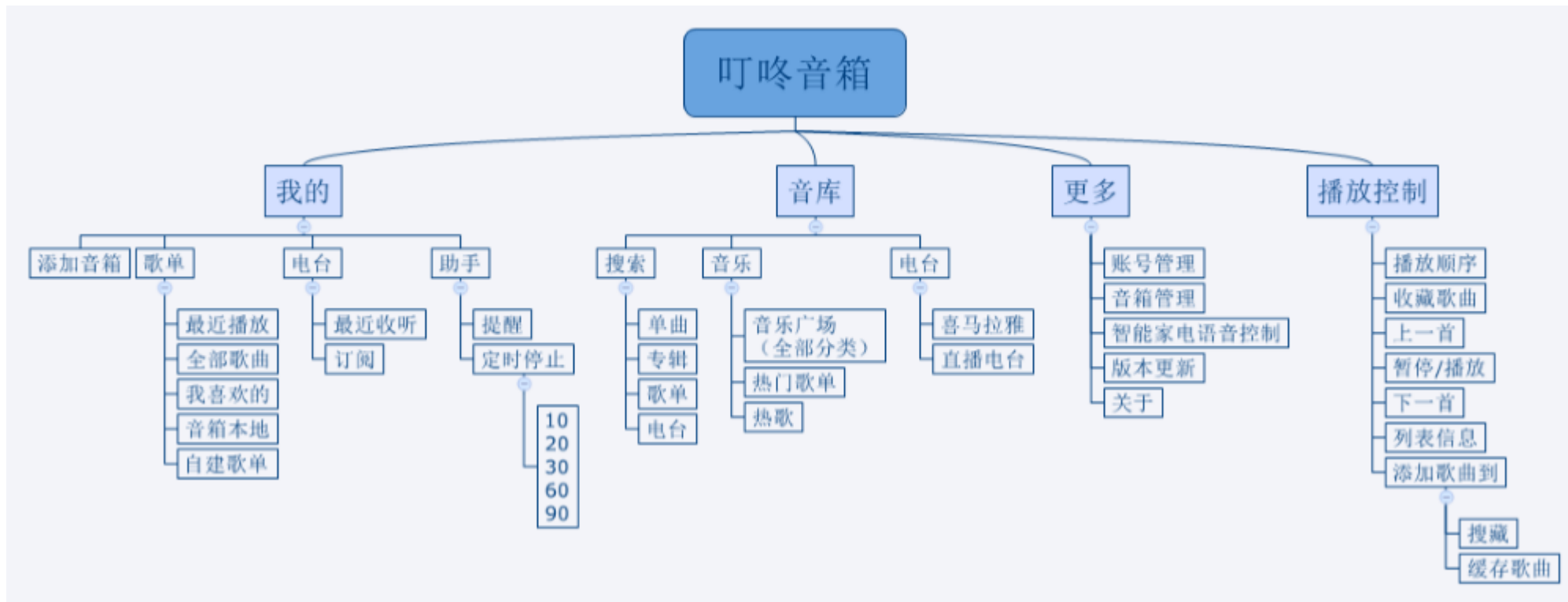


资料来源：京东、安信证券研究中心

既要智能交互也要音质表现，DingDong 智能音箱市场表现获用户肯定。不同于 Echo 在顶部设置 7 个麦克风收集语音信息的设计，这款价格仅 800 元的 DingDong 智能音箱在全方位 360 度都设置了音箱。同时，为了防止音频失真，在音箱之中进行了物理上的优化，加入了法拉第的短路环等一系列措施，在不影响语音交互的情况下，确保音质，尤其是对人声的还原。在保证语音智能服务的同时，不让音箱本身用户体验下降。根据京东下属研发机构提供数据，DingDong 智能音箱已经成为 wifi 音箱细分行业销量第一，目前日活跃用户与月活跃用户数保持在 30%到 40%之间，用户购买产品之后的次日留存率也达到了 75%，周留存率达到 55%，而用户每日的平均使用时间则达到了 2 小时。用户使用粘度反映 DingDong 为深度使用智能设备。

图 52：DingDong 音箱语音搜索服务信息架构





资料来源：互联网、安信证券研究中心

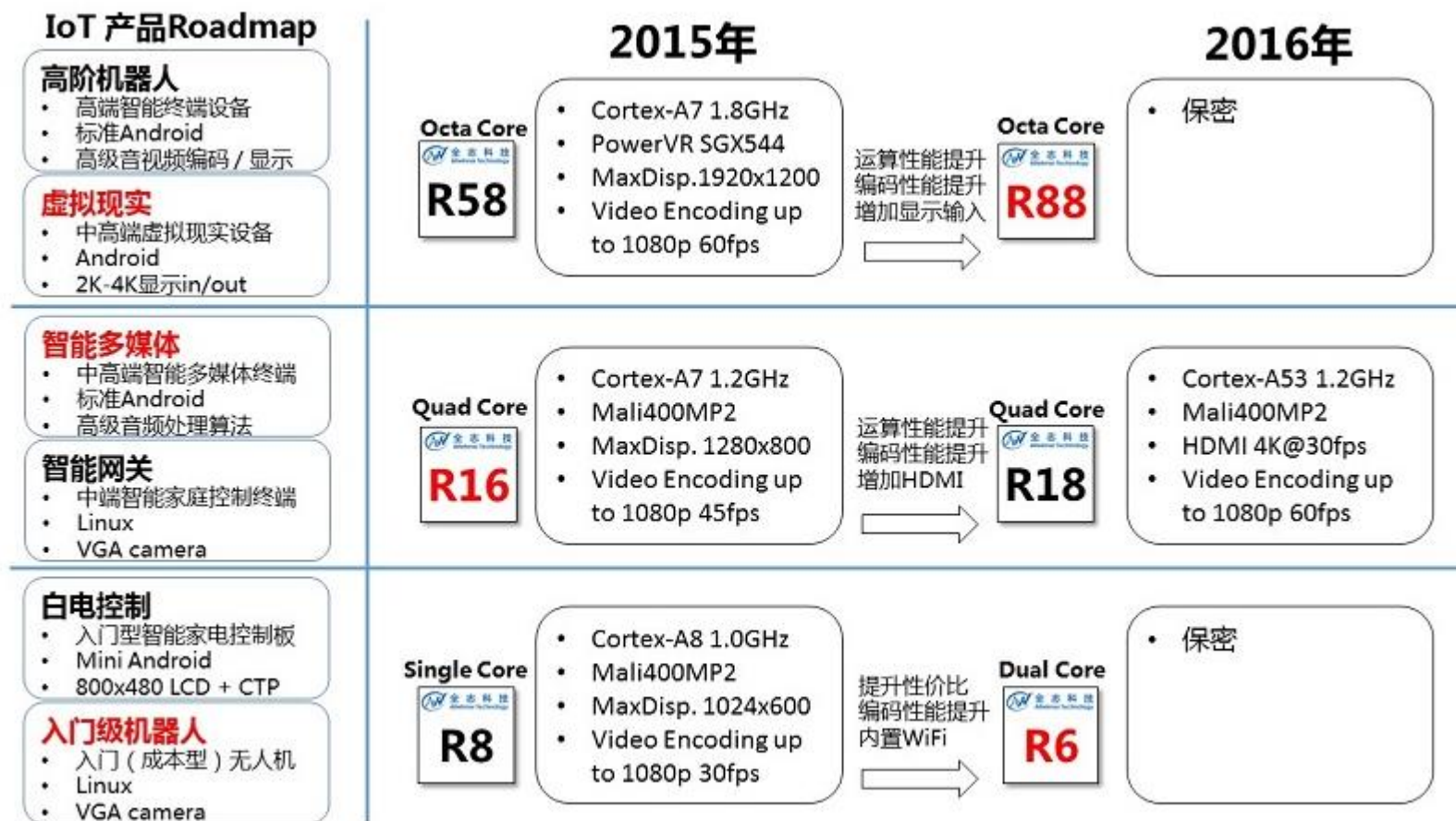
产品开发注重厂商强强合作，科大讯飞联手京东打造 DingDong 系列产品，未来将带动智能音箱全产业发展。与亚马逊研发 Echo 不同，京东在 DingDong 智能音箱产品研发方面一直采用厂商合作的模式。作为一家网络零售业巨头，京东更多采用厂商合作模式做智能硬件：在语音服务领域和科大讯飞设立合资公司，充分发挥其语音技术和人工智能技术优势；硬件开发过程利用全志科技的 R 系列智能硬件开发平台进行硬件开发；在音频领域，京东联合飞利浦、漫步者、猫王、索爱等影音硬件生产商及喜马拉雅、百度音乐等影音内容供应商成立了京东智能音频联盟。积极打造“JD+智能音频 解决方案”，考虑到传统的音箱企业拥有扎实的音频技术以及深厚的用户基础，智能云技术能很好的弥补传统音箱品牌在语音技术以及在线流媒体资源上的不足。京东最近就与老牌音频厂商哈曼联合推出的 JBL Go Smart 智能音箱，JBL 在麦克风拾音方面拥有独家的回声消除技术，结合 JD+远场语音识别和麦克风阵列，可以实现极高识别度的语音交互体验，并且支持 16 种方言的识别，语音、语意理解率达到 90%以上。

通过这种厂商合作开发模式，DingDong 系列产品可以集合语音服务、硬件开发、传统音响配件等多个行业优秀技术，共同打造具备下一代万联网核心设备要求的国产智能硬件产品，并通过蓝海战略迅速占领智能音箱国内市场。另外京东可以专注于配套智能硬件零售、应用服务平台等智能家居生态内容建设。未来 DingDong 系列产品将会带动智能音箱全产业发展的重大机遇。

DingDong 语音服务处理器全志打造，R 系列支撑物联网布局。全志科技支持 DingDong 的 R16 智能硬件平台对语音服务方面有完美的支持，还支持 AirPlay、DLNA、Qplay、Smart Link 等多种网络应用协议，方便开发者打造基于智能音箱的智能硬件产品。R16 智能硬件平台集成多媒体解码、液晶驱动、WIFI、蓝牙、录音，消噪，摄像于一体，提高了开发者的开发效率。由于全志科技的强项主要是在多媒体方面，所以从全志科技的物联网方案上我们不难看出，其主要针对的也还是侧重于多媒体性能的物联网产品。

图 53：全志科技 R 系列布局物联网





资料来源：全志科技、安信证券研究中心

## 4.投资建议

我们预计智能音响行业有望成为新的互联网和物联网入口，互联网厂商加速切入，从 CES 等展会观察到约三分之一的厂商在布局声学产品。音乐占据人们的时间，平均家庭收入 33,800 美元的家庭每周花 19 小时听音乐，音乐支出占到娱乐花费的 18%。音乐的商业模式在数字时代正在变化，内容收费需要与硬件结合。智能音响更多的价值也体现在家庭物联网的整合控制和人工智能的实际载体。

电子行业相关标的：

- 1、音响制造厂商：推荐国光电器，关注漫步者
- 2、消费电子芯片厂商：推荐全志科技，关注北京君正
- 3、声学元件厂商：推荐歌尔声学、信维通信

计算机行业相关标的：

1、软件平台处于产业核心地位：重点推荐科大讯飞、和而泰

2、家用服务机器人是产品最终发展形态：重点推荐参股 Jibo 智能家用机器人的东方网力

3、平台内容提供商受益新的商业模式：建议关注与图灵机器人合作的奥飞动漫、内容及版权提供商光一科技、安妮股份等

**传媒行业相关标的：**

1、音乐艺人：推荐苏宁环球，关注共达电声；

2、音乐演出&音乐综艺：推荐华录百纳、中南文化；

3、播映平台：推荐暴风科技，关注乐视网