

스마트 스피커 시장 동향과 전망

산업기술리서치센터 강맹수 연구위원

I. 스마트 스피커 기술 개요

III. 향후 시장 전망

II. 스마트 스피커 시장 동향

스마트 스피커는 차세대 사용자 인터페이스인 음성인식 인터페이스를 활용하기 위한 최적의 플랫폼이다. 컴퓨터와 대화하듯이 명령을 내리고 다양한 기능을 활용할 수 있는 스마트 스피커는 음성인식과 합성, 자연어처리를 위한 고도의 인공지능 기술이 융합되고 방대한 통합 지식 데이터베이스 구축과 운영 기술이 결합된 서비스이다.

'18년 상반기 전 세계 스마트 스피커 판매량은 전년동기대비 197% 증가한 25.8백만대를 기록하였다. 아마존이 '14년 11월 세계 최초로 스마트 스피커를 출시하여 시장을 주도해 왔으나, 올해 들어 구글이 판매량 기준으로 세계 시장점유율 1위를 차지하고 있다. 최근 중국의 알리바바, 샤오미, 바이두가 스마트 스피커 시장에서도 점유율을 빠르게 높이고 있다. 국내에서는 SK텔레콤이 '누구'를 '16년 9월 출시한 이래 KT '기가지니', 네이버 '클로바', 카카오 '카카오미니'가 경쟁하고 있다.

전 세계 스마트 스피커 시장은 연평균 30% 내외로 성장하여 '22년에는 '17년 대비 약 3.5배 증가한 87.1억달러로 예상된다. 미국이 성숙시장으로 서서히 진입하고 있음을 감안 하더라도 올해 전 세계 스마트 스피커 판매량은 최소 5천만대에서 최대 8천만대에 이를 것으로 전망된다. 스마트 스피커 시장은 초기 보급 단계를 지나 점차 서비스 경쟁이 본격화 되는 시기로 접어들고 있다. 스마트 스피커 메이커들의 플랫폼 경쟁이 활발하게 전개되고 새로운 킬러 서비스와 다양한 제품들이 등장하는 등의 양상이 전개될 것으로 예상된다.

* 본고의 내용은 집필자의 견해로 당행의 공식입장이 아님

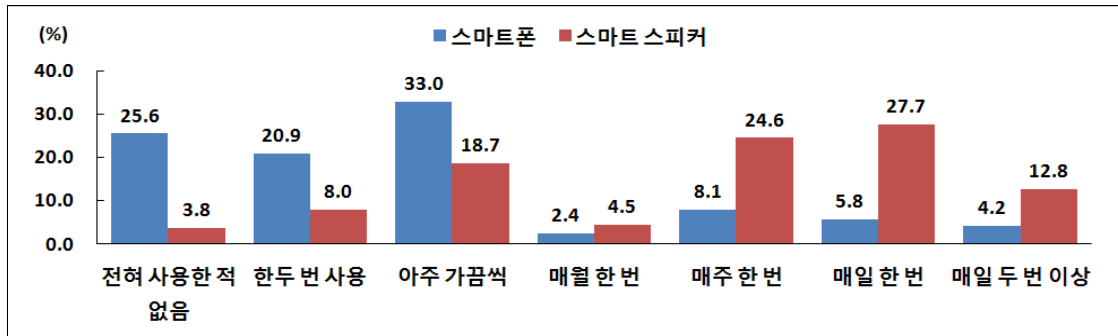
I. 스마트 스피커 기술 개요

□ (등장 배경) 스마트 스피커는 차세대 사용자 인터페이스의 대표 플랫폼

- ‘음성’ 인터페이스는 ‘터치’의 한계를 극복하기 위한 대안
 - 스마트폰과 함께 등장한 ‘터치’ 인터페이스는 기존의 입력 장치인 키보드나 마우스, 물리적 버튼, 스타일러스펜의 한계를 극복
 - ‘터치’ 방식은 별도의 물리적 입력 장치 없이 누구나 직관적으로 스마트폰을 사용할 수 있도록 만든 혁신적인 기술
 - ‘터치’ 방식은 누군가에게 문자를 보내거나 스마트폰의 설정을 변경하려고 할 때 여러 단계를 거쳐야 하고, 사용 중에는 스마트폰에서 눈과 손을 떼 수 없는 단점
 - ‘음성’ 인터페이스는 여러 단계를 거치지 않고 필요한 기능을 바로 사용할 수 있으며 ‘터치’에 비해서도 훨씬 더 직관적인 방식
- 스마트 스피커는 음성 인터페이스를 활용하기 위한 최적의 플랫폼
 - ‘음성’은 사람과 사람이 대화하는 것과 같이 사람이 컴퓨터와 대화하듯이 명령을 내리고 다양한 기능을 활용할 수 있는 가장 자연스러운 인터페이스
 - 애플이 '11.10월 아이폰에 음성인식비서 ‘Siri’를 탑재한 이래, 스마트폰에서 음성 인식 인터페이스를 사용할 수 있게 되었으나 실제 활용은 거의 없음
 - 스마트폰에서는 대부분의 경우 음성보다 터치가 더 빠르게 원하는 결과를 얻을 수 있고, 사용자들이 음성인식 인터페이스 사용 과정에서 프라이버시 노출을 꺼리기 때문¹⁾
 - 시장조사기관 Voysis가 '18.5월 미국 성인 1,200명을 대상으로 진행한 조사에 따르면 스마트 스피커 사용자의 65% 이상은 최소 1주일에 한 번 이상 음성인식 인터페이스를 사용하는 반면,
 - 스마트폰 보유자의 약 80%는 음성인식 인터페이스를 한 번도 사용하지 않거나 아주 가끔씩만 사용하는 것으로 나타남

1) 공공장소에서 스마트폰을 이용한 음성인식 인터페이스를 사용할 경우, 무엇을 검색하거나 실행하는지가 주변에 노출되는 반면, 스마트 스피커는 대부분 프라이버시가 보호되는 가정 내에서만 사용되는 차이점

〈그림 1〉 음성 인식 인터페이스 사용 빈도 비교



자료 : Voysis(2018)

□ (기술) 스마트 스피커는 다양한 분야의 인공지능 기술과 방대한 통합 지식 데이터 베이스를 구축하고 활용하는 기술들이 융합

- 스마트 스피커가 사용자의 말을 듣고 사람이 말하듯이 적절히 응답하려면 음성인식, 자연어처리, 음성합성을 위한 고난이도의 인공지능 기술이 필요
 - 음성인식은 사람의 말을 텍스트로 변환하는 기술로, 음성인식 인터페이스(VUI, Voice User Interface)를 사용하는 스마트 스피커의 기반 기술
 - 사람마다 서로 다른 발음, 상황에 따라 같은 단어도 다르게 발음되는 점 등의 다양한 문제로 인해 구현이 쉽지 않았으나 딥러닝의 발전으로 개별 단어 단위에서는 인식률 95% 이상을 실현
 - 스마트 스피커는 주변 소음이나 음악과 사람의 음성을 분리하기 위해 여러 개의 마이크를 내장하여 전처리 과정을 거침
 - 예를 들어 아마존 에코는 7개의 내장 마이크를 통해 주변 소음을 제거
 - 자연어처리는 사람의 말을 기계가 이해하도록 만드는 인공지능 기술로, 텍스트 문장에서 핵심어를 추출하여 명령 도메인과 구체적인 요청 사항을 구분
 - “BTS 히트곡을 틀어줘”라는 명령을 듣고 ‘음악’ 도메인에서 ‘BTS라는 뮤지션이 부른 노래’를 찾아서 ‘뮤직앱을 실행’하라는 뜻으로 해석하는 과정
 - 아직은 문장을 사람처럼 이해하는 것이 불가능하기 때문에 학습되지 않은 영역에 대해서는 대응하지 못하는 한계가 있음

- 음성합성은 텍스트 문장을 사람이 말하듯이 소리로 변환하는 기술로, 스마트 스피커가 사람과 자연스러운 커뮤니케이션을 가능하도록 구현

〈표 1〉 스마트 스피커의 기술 구성

구분	내용	해결해야 할 요소
음성 인식	<ul style="list-style-type: none"> • 사용자의 음성 명령을 인식하여 대기모드에서 활성 상태로 전환 • 음성 명령을 텍스트로 변환 	<ul style="list-style-type: none"> • 주변 소음과 음성 구분 • 2~3m의 원거리에서 음성 인식 • 불명확한 발음 인식 • 다중 사용자 식별
자연어처리	<ul style="list-style-type: none"> • 텍스트로 변환된 문장에서 처리해야 할 도메인 영역 식별 • 사용자의 명령 의도 파악 	<ul style="list-style-type: none"> • 모호한 표현 • 처리가능 도메인 범위 확보 • 일상대화 처리
음성 합성	<ul style="list-style-type: none"> • 텍스트로 생성된 응답 문장을 음성으로 변환하여 사용자에게 응답 	<ul style="list-style-type: none"> • 상황에 따른 음성 높낮이, 끊어 읽기, 발화 속도 설정 • 문장 내에서의 예외 발음 처리 (예 : 연음규칙 등)
통합 DB 구축 및 활용	<ul style="list-style-type: none"> • 방대한 지식 DB를 활용하여 적절한 응답 문장 생성 • 검색, 통합 DB, 추천기술 활용 	<ul style="list-style-type: none"> • 처리가능 도메인 범위 확보 • 커버리지 이슈 • 자연스러운 응답 문장 생성

자료 : 자체 작성

- 음성 명령에 적절한 응답을 하기 위해서 통합 지식 DB를 구축하고, 검색과 텍스트 요약 및 개인화 추천 기술들을 적용
- 통합 지식 DB는 사용자의 질의에 응답하기 위한 필수적인 기초 데이터의 집합이며 커버리지가 넓을수록 다양한 명령에 대응 가능
- SK텔레콤은 ‘누구’의 작동을 위해 약 2.5억 건의 트리플²⁾을 구축

2) 통합 지식 DB에 저장되는 데이터 형식으로 객체(entity), 속성(attribute), 값(value)으로 구성

〈표 2〉 통합 지식 데이터베이스 구축 사례

도메인	세부 카테고리 종류	Entity 수	Triple 수
음악	앨범, 노래, 가수	13,525,524	116,838,494
방송/영화	VOD, 인물, 방송, 코너, 채널, 편성표, 영화/상영관	1,069,387	25,313,520
교통	POI ³⁾ , 도로명, 행정명	11,315,317	90,301,885
오디오	오디오북, 팟캐스트	20,714	236,033
지식백과	위키피디아, 기네스북, 고사성어, 공공사전, 큐레이션, 우리말 사전, 시사 용어, 속담/학습	1,482,311	17,454,568
스포츠	평창올림픽	250	1,740
증권	KOSPI, KOSDAQ	2,569	3,448
음식	음식명, 상점명	27,646	143,407

주 : 2018.3월말 기준

자료 : SK텔레콤(2018)

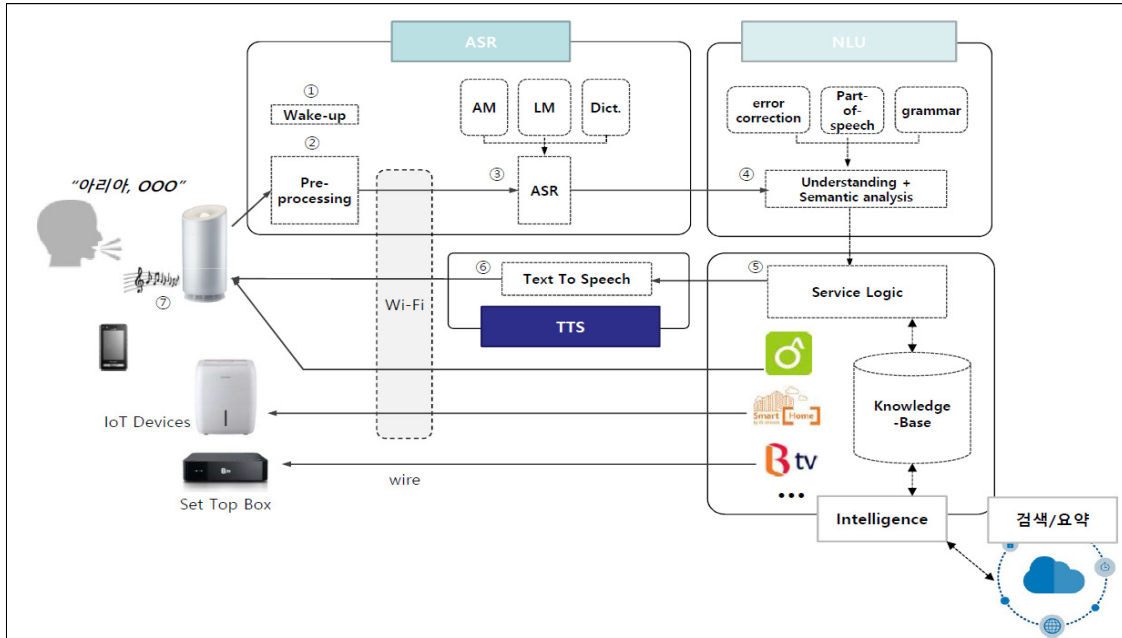
□ (작동과정) 스마트 스피커의 인공지능 플랫폼은 클라우드 상에서 작동

- 개별 기기는 Wake-up 명령⁴⁾ 인식과 음성 데이터 전처리만 담당
 - Wake-up 명령 처리를 제외한 음성인식을 위한 음향모델(AM), 언어모델(LM), 사전 등은 클라우드 상에서 구동하는 방식
 - 자연어처리 모듈(NLU, Natural Language Understanding), 사용자의 명령과 서비스 매칭, 음성합성(Text-to-Speech) 기능도 클라우드 상에서만 구현
- 새로운 기능을 클라우드에 등록하면 사용자가 별도의 업그레이드 절차 없이 이용 가능하기 때문에 서비스 확장에도 유리
 - 클라우드를 통해 다양한 외부 개발자와의 협업이 쉬워지고 스마트 스피커를 인공지능 플랫폼으로 활용하는 것이 가능해짐
 - 또한 클라우드를 이용함으로써 기기 가격을 저렴하게 유지할 수 있음

3) 관심지점(Point of Interest) : 전자 지도에 표시하는 주요 시설물, 역, 공항, 호텔 등의 좌표

4) Wake-up 명령은 대기 상태의 기기를 활성화는 “Alexa”, “아리아”, “지니” 등의 구문

〈그림 2〉 스마트 스피커 작동 프로세스 예시



자료 : SK텔레콤(2018)

Ⅱ. 스마트 스피커 시장 동향

1. 세계 시장

□ (시장규모) '18년 상반기 전 세계 스마트 스피커 판매량은 전년동기대비 197% 증가한 25.8백만대를 기록

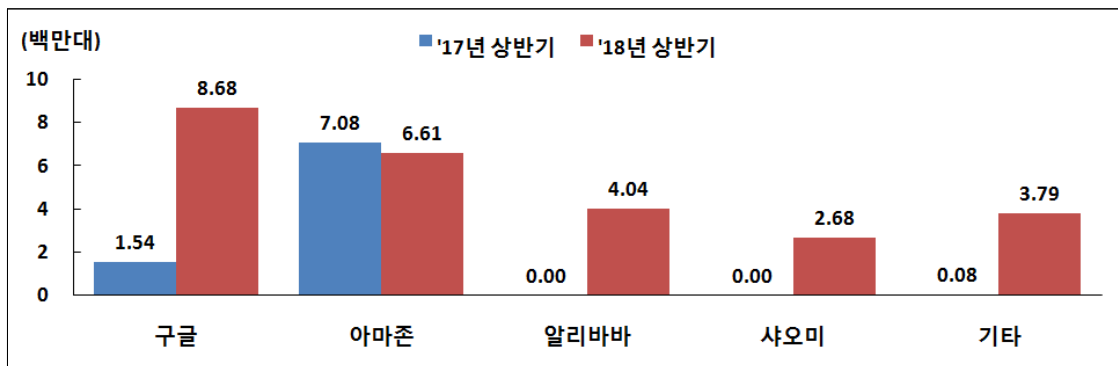
○ 구글과 아마존이 전 세계 스마트 스피커 시장을 주도

- 아마존이 '14.11월 세계 최초로 음성인식 스피커 'Echo'를 출시한 이래 '17년까지는 시장을 주도하였으나, '18년 들어 성장이 정체
- 구글은 '18년 상반기에 전년동기대비 464% 증가한 868만대를 판매하면서 아마존을 제치고 판매량 기준 시장점유율 1위로 부상
 - 전 세계 시장점유율('17.상→'18.상, %): 구글(17.7→33.7), 아마존(81.4→25.6)

○ '18년 들어 중국 IT 기업들의 약진이 두드러짐

- 알라바바의 'Tmall Genie'와 샤오미의 'Xiaomi Mini AI'가 '18년 상반기 각각 400만대, 268만대를 판매하여 전 세계 시장점유율 15.6%, 10.4%를 기록
- 중국 내 '18년 3분기 스마트 스피커 판매는 580만대이며 알리바바, 샤오미, 바이두가 전체의 87.9%의 점유율을 차지
 - 중국 내 판매량('18.3Q, 백만대) : 알리바바(2.2), 샤오미(1.9), 바이두(1.0)

〈그림 3〉 전 세계 스마트 스피커 판매대수



자료 : Canalsys(2018)에서 재정리

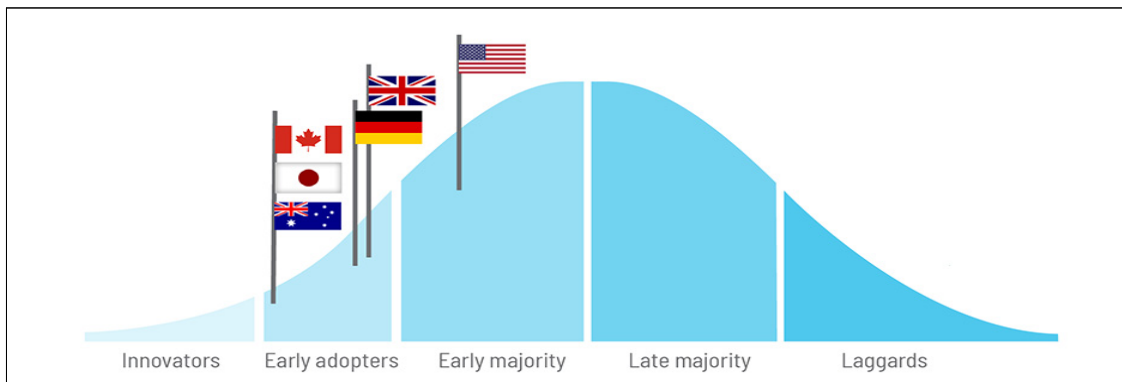
□ (보급률) 세계 스마트 스피커 시장은 초기수용 단계를 지나 초기대중화 단계로 발전하는 과정

- 미국은 이미 초기대중화 단계에 진입하였으며, 독일과 영국도 '19년에는 초기대중화 단계에 이를 것으로 예상
- 신기술이 확산되는 과정에서 보급률 16%까지는 초기수용 단계, 16~50%까지를 초기 대중화 단계로 구분⁵⁾
- 스마트 스피커 전문 조사기관인 Voicebot.ai의 '18.9월 조사에 따르면 미국의 보급률은 23%로 연말까지 최대 30%까지도 달성 가능할 전망
- 독일과 영국은 초기수용과 초기대중화의 경계선에 위치한 것으로 추정되며, 캐나다와 일본의 보급률은 약 10% 내외

5) 신기술 확산 과정 : 초기수용(Early adopter) → 초기대중화(Early majority) → 후기대중화(Late majority) → 말기수용(Laggard)

- 초기대중화 단계 진입으로 스마트 스피커의 품질 및 서비스 경쟁이 본격화
 - 아마존 ‘Echo’는 '18.8월 기준 5만개의 스킬⁶⁾이 탑재되고, 2만개 이상의 관련 제품이 출시
 - 스마트 스피커와 연계한 생활가전, 스마트홈 기기, 기존 스피커 등의 제조 기업들과 인터넷 서비스 기업들이 늘어나면서 생태계 경쟁이 심화되는 추세

〈그림 4〉 전 세계 스마트 스피커 도입 단계



자료 : Voicebot.ai

□ (주요기업) 뛰어난 인공지능 기술을 보유한 구글과 아마존의 양강 구도

- 아마존과 구글이 자체 음성인식 인터페이스를 바탕으로 전 세계 스마트 스피커 시장을 주도
 - 아마존은 ‘Echo’와 연동되는 2만개 이상의 제품을 확보하여 가장 광범위한 생태계를 구축
 - 구글의 ‘Google Home’은 뛰어난 인공지능 기술을 바탕으로 현재 유일하게 이중 언어를 동시에 지원 가능
 - 영어, 일본어, 독일어, 프랑스어, 스페인어, 이탈리아어 6개 중 2개 언어를 선택하면 해당 언어로 동시에 ‘Google Home’을 이용 가능
 - 애플의 ‘HomePod’은 스피커 음질이 우수하여 음악 감상에 장점이 있지만, 음성인식 AI인 ‘Siri’의 자연어 이해 성능이 구글에 비해 미흡

6) 스킬(skill)은 아마존 ‘Echo’를 이용해서 실행할 수 있는 기능으로 구글의 ‘Google Home’에서 이용할 수 있는 기능은 액션이라고 지칭

- 저렴한 가격이 가장 큰 경쟁 요소이며, 스피커 제조업체들은 음악 감상에 특화된 제품들로 틈새시장을 개척
 - 구글과 아마존은 적극적인 할인판매를 통해 보급률을 높이는 전략을 실행
 - 중국의 시장점유율 상위 3개사도 저가 정책과 할인 판매를 통해 시장에 안착
 - 바이두가 '17년 출시한 'Raven H'(244달러)는 실패했지만, 최근 출시한 'Xiao Du'는 89위안(약 1.5만원)으로 대폭 할인하여 중국 내 점유율을 확대
 - 미국의 소노스는 아마존, 구글, 애플의 음성인식 인터페이스를 모두 지원하고 음악 감상에 특화된 제품으로 고객을 확대

〈표 3〉 주요 스마트 스피커 기업

기업	제품	출시 시점	가격	주요 특징
아마존	Echo	2014.11월	80달러	5만개 이상의 스킬 보유
구글	Google Home	2016.11월	129달러	이중 언어 지원
애플	HomePod	2017.12월	349달러	뛰어난 음질
알리바바	Tmall Genie	2017.9월	73달러	알리페이 사용 가능
샤오미	Xiaomi Mini AI	2017.8월	45달러	샤오미 IoT 기기 호환
바이두	Xiao DU	2018.6월	36달러	뛰어난 검색 기능
소노스	Sonos One	2017.10월	199달러	아마존, 구글, 애플 호환

자료 : 자체 작성

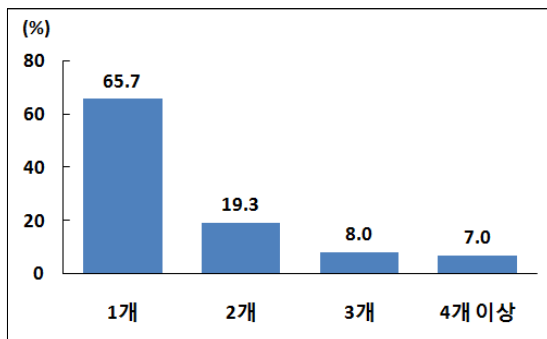
□ (가구당 보유대수) 스마트 스피커 사용자는 가구당 평균 약 1.5~2개를 보유

- 다양한 조사에서 스마트 스피커를 보유하고 있는 사람들의 가구당 평균 보유대수는 일관되게 1.5~2대
 - 미국과 영국에서 독립적으로 진행된 조사에 따르면 전체 사용자의 약 2/3는 스마트 스피커를 1대만 보유하고, 약 20%는 2대, 나머지는 3대 이상을 보유
 - Voicebot.ai의 미국 스마트 스피커 사용자 조사에 따르면 가구당 평균 보유대수는 약 1.8대
 - 영국의 여론조사업체 YouGov도 영국의 스마트 스피커 사용자들이 가구당 평균 최소 1.5대 이상 보유하고 있는 것으로 조사
 - 스마트 스피커는 장소를 이동하는 것이 불편하기 때문에 적극 사용자들은 필요한 장소마다 추가로 설치하기 때문

○ 스마트 스피커의 활용 영역이 넓어질수록 가구당 보유대수가 늘어날 전망

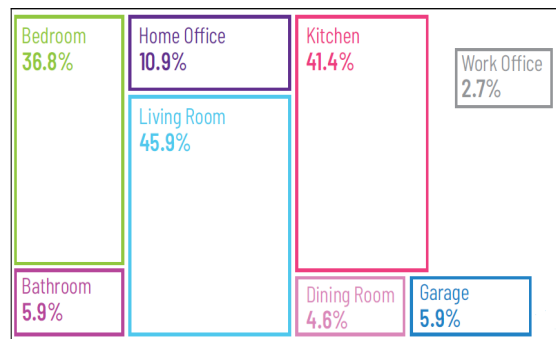
- 미국 가정 내에서 스마트 스피커를 가장 많이 설치하는 장소는 사람들이 많이 모이는 거실, 부엌, 침실 순
 - 거실과 침실은 개인적인 시간을 보내는 장소로 스마트 스피커를 통해 음악을 듣거나 정보를 검색하는 용도로 주로 사용
 - 부엌에서는 식료품 주문, 요리 레시피 검색, 요리 중 음악 재생 등에 이용
- 스마트 스피커의 개인별 추천 기능이 향상되면 사람들이 많은 곳에서 공동으로 사용하는 방식에서 개인별 보유로 변화할 가능성

〈그림 5〉 가구당 보유 대수



자료 : Voicebot.ai(2018.3)

〈그림 6〉 스마트 스피커 주 설치 장소



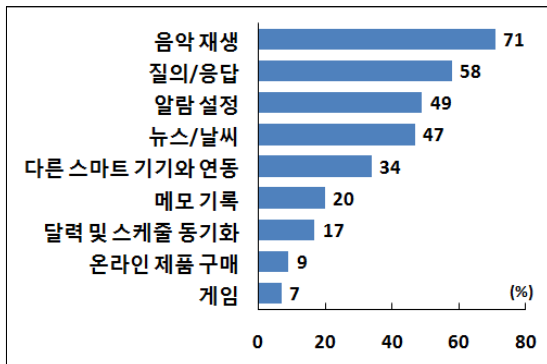
자료 : Voicebot.ai(2018.3)

□ (활용) 현재 스마트 스피커의 주된 사용 목적은 음악 감상 또는 아주 간단한 명령을 실행하는 것 등으로 제한적

- 미국과 영국 사용자 대상 조사에 따르면 사용자들은 스마트 스피커를 음악 재생, 날씨나 뉴스, 알람 및 일정 설정, 간단한 질의·응답의 용도로 사용
 - 스마트 스피커를 다른 작업을 하는 도중에 간단한 명령을 실행시키는 용도로 주로 사용한다는 의미
 - 국가별로 사용자의 이용 방식에 차별성이 없다는 사실은 처리할 수 있는 명령이 제한적이어서 스마트 스피커의 차별성이 뚜렷하지 않기 때문일 수도 있음
 - 현재의 스마트 스피커는 자연어처리를 위한 인공지능 기술의 한계로 단문 형태의 간단한 명령만 처리 가능

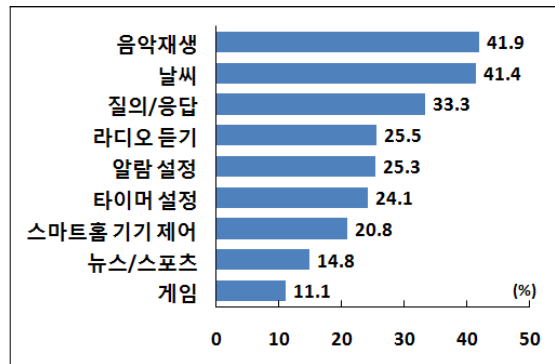
- 스마트 스피커 메이커들이 기대했던 것과 달리 온라인 쇼핑에 활용하는 비중은 아직 낮은 편
- 영국 사용자에게 대한 조사 결과, 스마트 스피커를 통해 온라인으로 제품을 구매한 경험이 있는 경우는 9%에 불과
 - 미국 사용자 대상 조사에서는 최소 한 달에 한 번 정도 주기적으로 온라인 쇼핑에 스마트 스피커를 활용하는 비중이 11.5%로 나타남
 - 온라인 쇼핑에서 결제나 제품 검색을 위한 보조 수단으로 스마트 스피커를 사용하면 구매 과정의 번거로움을 줄일 수 있기 때문에 점차 이용이 증가할 것으로 예상

〈그림 7〉 주 사용 기능(영국)



자료 : KOTRA(2018.9) 재인용

〈그림 8〉 매일 사용하는 기능(미국)



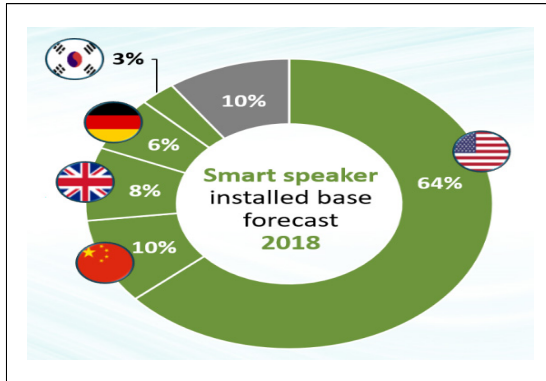
자료 : Voicebot.ai(2018.3)

2. 국내 시장

□ (시장규모) '18년 국내 스마트 스피커 누적 설치 대수는 3백만대로 예상

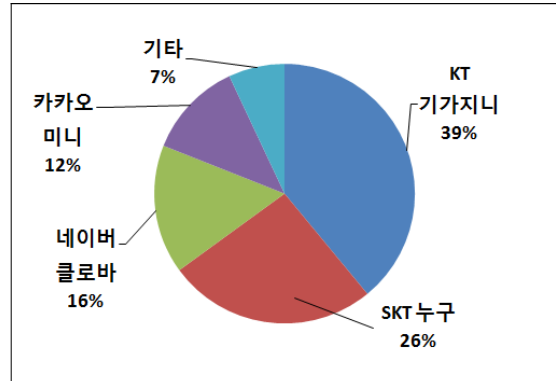
- 국내 스마트 스피커 시장 규모는 '18년말 기준 전 세계 5위권에 이를 전망
- 시장조사회사 카날리스는 '18년말 전 세계 스마트 스피커 누적 설치 대수가 약 1억대 이고, 한국은 그 중에서 약 3%를 차지할 것으로 예상
 - '17년말 기준, 국내 스마트 스피커 판매량은 약 100만대로, KT의 '기가지니'가 50만대, SK텔레콤의 '누구'가 40만대를 판매
 - 국내 시장조사회사 컨슈머인사이트가 '18.4월 1,415명을 대상으로 조사한 결과, 응답자의 39%가 KT '기가지니', 26%는 SK텔레콤의 '누구'를 사용

〈그림 9〉 스마트 스피커 국가별 점유율 전망



자료 : Canalsys(2018.7)

〈그림 10〉 국내 스마트 스피커 이용 브랜드



자료 : 컨슈머인사이트(2018.7)

□ (주요기업) 통신사업자인 SK텔레콤과 KT가 시장을 주도

- SK텔레콤이 '16.9월 국내 최초로 스마트 스피커를 출시한 이후, KT, 네이버, 카카오가 경쟁
- 통신 사업자는 IP TV 이용자를 중심으로 시장을 확대하고 있는 반면, 네이버와 카카오는 라인과 카카오 캐릭터를 활용하여 스마트 스피커를 차별화
- 다른 서비스와 연계한 프로모션을 통해 실제 구입 가격은 매우 낮은 편
 - ‘기가지니2’의 정가가 24.9만원이나 TV 결합상품으로 3년 약정 시, 월 1.5만원에 구입 가능하고, 기존 올레TV 사용자에는 무료 업그레이드도 제공
- 삼성전자는 음질을 강조한 스마트 스피커 ‘갤럭시홈’을 내년에 출시할 예정

〈표 4〉 국내외 주요 스마트 스피커

기업	제품	출시 시점	가격	특징
SK텔레콤	누구	2016.9월	14.9만원	인공지능 플랫폼 전략
KT	기가지니	2017.1월	24.9만원	TV 연동, 스마트홈 제어
네이버	클로바	2017.10월	12.9만원	음악재생, ‘배달의 민족’
카카오	카카오미니	2017.11월	11.9만원	카카오 서비스 이용
삼성전자	갤럭시홈	2019년 예정	미정	하만 AKG 스피커 6개 탑재

자료 : 자체 작성

□ (이용목적) 음악 감상, 날씨 확인 등에 주로 이용하며, TV 조작에 활용하는 비중이 높은 특징

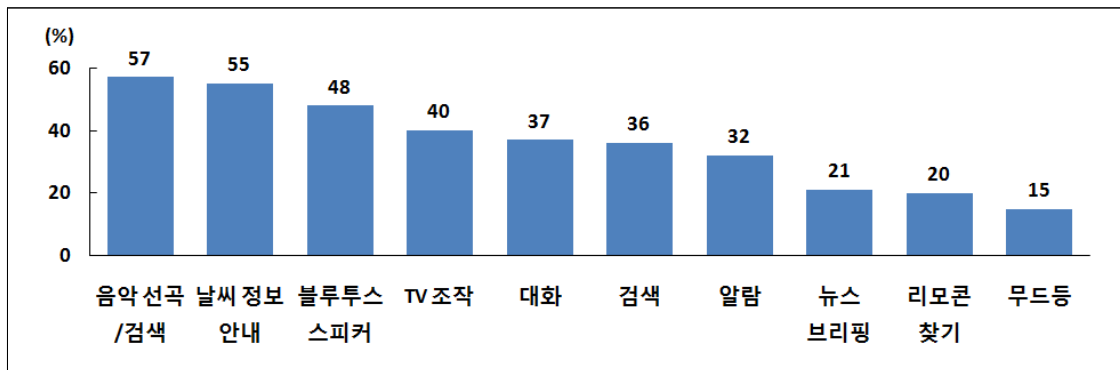
○ 사용자들의 이용 형태는 기본적으로 해외 사례와 크게 다르지 않음

- 음악 감상, 일기 예보, 간단한 대화, 정보 검색, 알람 설정 등을 주로 사용
- SK텔레콤과 KT가 자체 IP TV와 스마트 스피커를 연계하여 판매하면서 TV 조작과 관련된 기능을 많이 활용하는 것이 특징

○ 스마트 스피커 이용 만족도는 50%를 넘지 않는 상황

- 컨슈머인사이트의 조사에 따르면 스마트 스피커에 만족하는 사용자는 평균 49%에 불과하여 소비자의 기대에는 미치지 못하고 있는 상황
- 가장 큰 불만족 이유는 음성 인식이 잘 되지 않는 것이었으며, 이는 한국어에 대한 음성인식과 자연어처리 기술 개발이 영어에 비해 더 까다롭기 때문

〈그림 11〉 국내 사용자의 스마트 스피커 활용 목적



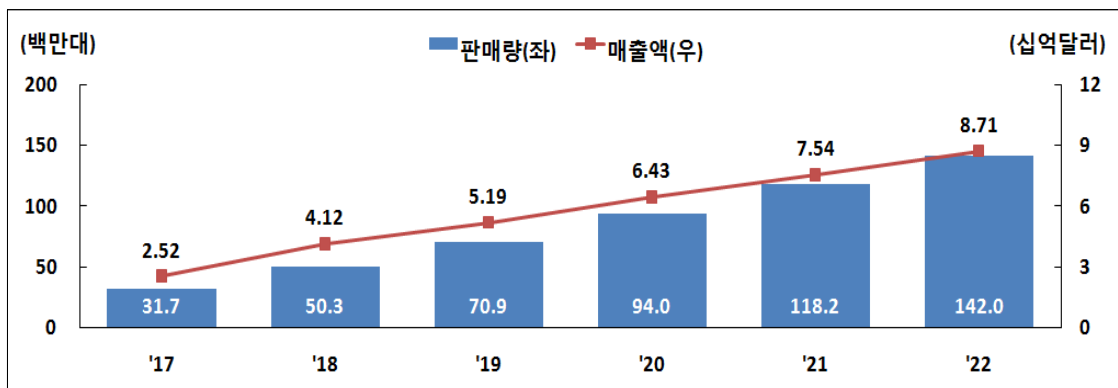
자료 : 컨슈머인사이트(2018.6)

Ⅲ. 향후 시장 전망

□ 전 세계 스마트 스피커 판매량은 '22년 142백만대에 이를 것으로 전망

- 스마트 스피커 시장 규모는 '22년에 '17년 대비 3.5배 증가한 87억달러로 예상
 - S&P Global Market Intelligence는 스마트 스피커의 유용성이 향상되고 수요가 꾸준히 증가한다면 '22년까지 연평균 30% 내외의 성장을 할 것으로 추정
 - 스마트 스피커 시장의 성장 속도가 매년 시장조사회사의 전망치보다 크게 높았던 점을 감안하면 실제 시장 규모는 훨씬 더 클 가능성이 높음
 - 다만, 평균 판매 단가는 '18년 81.9달러에서 '22년 61.3달러로 하락할 전망
 - 평균 판매가 예상(달러) : 81.9('18)→73.2('19)→68.4('20)→63.8('21)→61.3('22)
 - 스마트 스피커 시장이 초기 대중화 단계로 서서히 진입함에 따라 경쟁 심화로 인해 해당 판매가는 예상보다 더 빠르게 하락할 가능성
- 연간 판매량의 50%가 연말 쇼핑 시즌에 집중되는 특성을 감안하면 '18년 전 세계 스마트 스피커 판매량은 최대 8천만대에 이를 수도 있음
 - 3, 4분기 판매량이 2분기 판매량과 비슷하다고 가정하더라도 '18년 총 판매량이 5천만대는 무난히 넘을 전망
 - 가장 시장 규모가 큰 미국 시장이 서서히 성숙시장으로 진입하고 있고, 중국의 3분기 판매량이 전분기 대비 단지 1%만 증가한 점은 불안 요인

〈그림 12〉 전 세계 스마트 스피커 시장 규모 전망



자료 : S&P Global Market Intelligence(2018.6) 재인용

□ 스마트 스피커를 중심으로 AI 서비스 개발 플랫폼 구축이 활발해질 전망

○ 스마트 스피커는 최적의 AI 서비스 개발 플랫폼

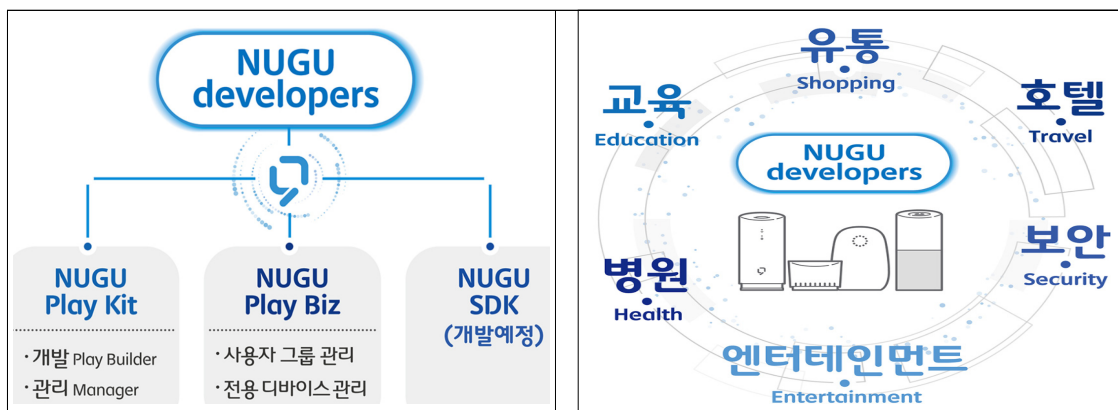
- 스마트 스피커에 접목된 다양한 인공지능 기술을 활용하면 새로운 서비스 개발 기간을 단축하고 시장에 빠르게 확산시킬 수 있는 장점
- 스마트 스피커를 위해 구축한 음성인식·합성과 자연어처리 모듈, 방대한 지식 DB를 활용하면 외부개발자들의 다양한 아이디어를 쉽게 구현할 수 있는 장점
- 클라우드를 통해 외부개발자들의 신규 서비스를 기존 스마트 스피커 사용자 전체에 손쉽게 확산시킬 수 있음
- 스마트 스피커 메이커들은 외부개발자들과 활발한 오픈 이노베이션을 함으로써 인공지능 생태계 경쟁력을 확보하려고 시도

○ 아마존, 구글 등의 해외 선도기업 뿐만 아니라 국내에서도 스마트 스피커 중심의 플랫폼 구축이 빠르게 진행되는 상황

- 아마존, 구글은 자신들의 생태계 내에 이미 1만개 이상의 호환 기기와 서비스 제공자를 확보
- SK텔레콤은 '18.10월 'NUGU Developers'를 일반에 공개하고 외부개발자 누구나 직접 인공지능 서비스를 개발할 수 있는 Play Kit와 Biz를 제공
 - 'NUGU Developers'를 통해 개발된 클래식 음악 무료 감상이 가능한 '클래식 매니저', 스타들의 행사 일정을 확인하는 '아이돌 캘린더' 서비스가 추가

〈그림 13〉

SK텔레콤의 'NUGU Developers'



자료 : SK텔레콤(2018)

□ 초기 보급 단계를 지나 스마트 스피커의 차별화가 부각되는 시기

- 스마트 스피커 메이커에 따른 차별성이 거의 없는 현재 상황을 개선하기 위한 새로운 서비스 개발이 시급
 - 음악 재생, 날씨나 뉴스 검색 등을 벗어나 소비자들이 스마트 스피커를 구입해야 할 이유를 제시할 수 있는 대표 서비스가 필요한 상황
 - 스마트 스피커를 활용한 양방향 커뮤니케이션이 가능한 외국어 교육 서비스, 음성인식 인터페이스를 활용하여 온라인 쇼핑이나 음식 배달, 금융 서비스 이용 편의성을 획기적으로 향상시킬 수 있는 서비스 등이 가능할 것으로 예상
- 외부기업들이 구글, 아마존, 애플의 음성인식 인터페이스를 활용하여 다양한 형태의 제품들이 등장
 - 음악 감상 목적으로 스마트 스피커를 활용하는 사용자를 타겟으로 전문 스피커 제조기업이 스마트 스피커 메이커와 협업하는 경우가 늘어날 것
 - JBL, 하만 카돈, 소노스 등과 같은 스피커 전문 제조기업들이 음악 감상에 특화된 다양한 스마트 스피커를 출시
 - 음성인식 인터페이스와 화면이 결합된 스마트 디스플레이가 시장점유율을 높여 나갈 것으로 기대
 - Lenovo의 ‘Smart Display’, 아마존 ‘Echo Show’, JBL ‘Link View’, 구글 ‘Google Home Hub’, LG ‘WK9’ 등이 이미 출시된 상황
 - 외부기업들은 단일 플랫폼에 종속되기보다 서로 다른 음성인식 인터페이스를 모두 지원하는 형태로 제품을 개발하는 경향이 심화될 것
 - LG전자는 ‘썬큐’ 가전제품에 구글, 아마존, 네이버의 음성인식 인터페이스를 지원하고, Sonos도 구글, 아마존, 애플을 모두 지원

참고문헌

[국문자료]

산업은행 미래전략연구소(2017.4), “인공지능 기반 음성인식”, 산업기술경쟁력맵
정규준(2018.6), “인공지능 서비스 누구 기술 소개”, SK텔레콤
최미란(2017.1), “자연어처리 기반의 인공지능 기술 현황”, 한국정보통신기술협회
최지혜·이선희(2017.5), “음성인식 AI 비서 시장의 현황과 시사점”, 정보통신정책연구원
카카오 AI Report(2018.10), “세상을 바꿀 변화의 시작, 음성 인터페이스와 스마트 스피커”
한국무역협회(2018.9), “중국 스마트 스피커 산업 동향”
한국인터넷진흥원(2018.4), “스마트스피커 출시 주요 해외기업들의 개인정보보호 대응 동향”
KOTRA(2018.9), “영국 스마트홈 기기 시장의 성장”

[영문자료]

Canalys(2018.4), “Google beats Amazon to first place in smart speaker market”
_____ (2018.8), “Global smart speaker shipments grew 187% year on year in Q2 2018, which
China the fastest-growing market”
_____ (2018.7), “Smart speaker installed base to hit 110 million by end of 2018”
_____ (2018.11), “Baidu heats up China’s smart speaker market, beating LingLong to enter top
three in Q3 2018”
PWC(2018.3), “Consumer Intelligence Series : Prepare for the voice revolution”
S&P Global Market Intelligence(2018.6), “Technology, Media & Telecommunications Smart
Speaker Take Off”
Voicebot.ai(2018.3), “Smart Speaker Consumer Adoption Report, March 2018”
_____ (2018.11), “Smart Speaker Consumer Adoption Report, November 2018”
Voysis(2018.6), “Voice Shopping Consumer Adoption Report”