

ResearchInChina
www.researchinchina.com

Automotive Voice Industry Report, 2023-2024

Dec. 2023

In OEM market, 46 brands install automotive voice as a standard configuration in 2023

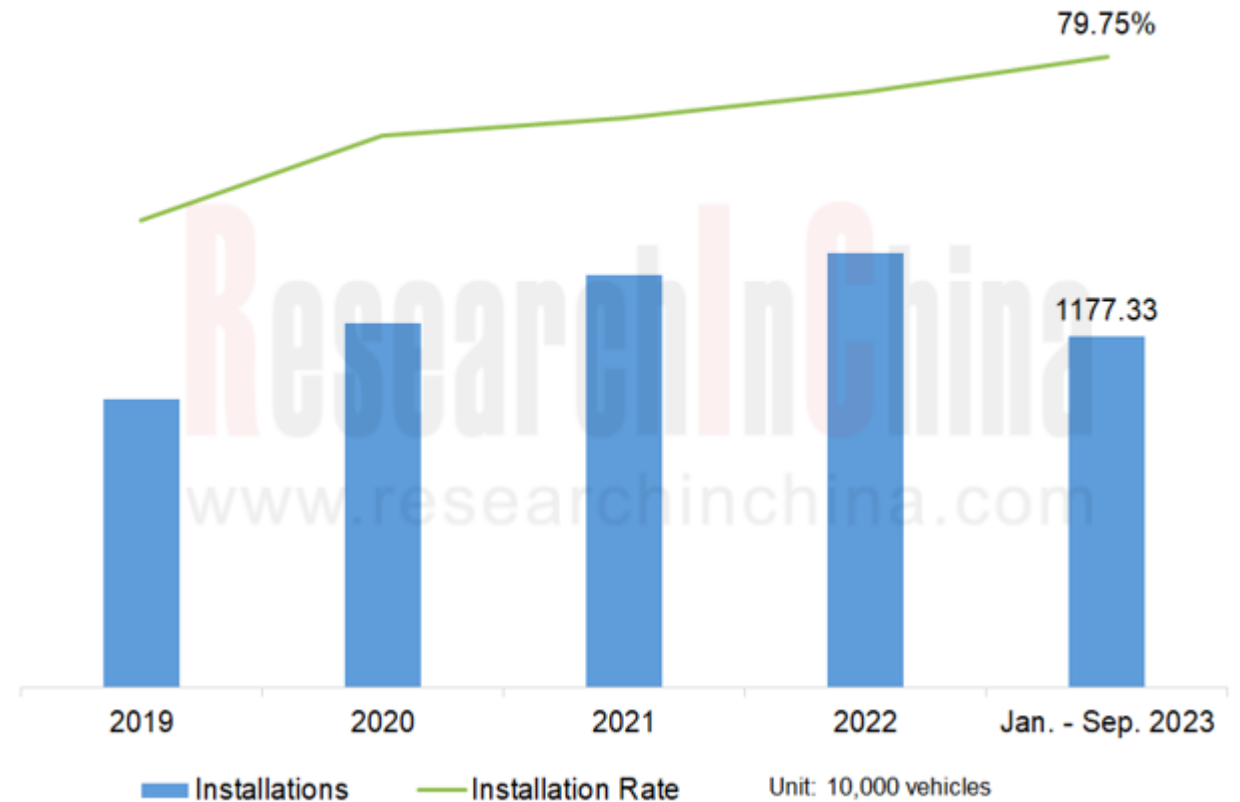
The automotive voice interaction market is characterized by the following:

1. In OEM market, 46 brands install automotive voice as a standard configuration in 2023.

From 2019 to the first nine months of 2023, automotive voice saw rising installations and installation rate. In the first three quarters of 2023, nearly 12 million vehicles were pre-installed with automotive voice, with the installation rate of nearly 80%.

In 2023, there are 46 passenger car brands boasting automotive voice installation rate of 100%, including AITO, Avatr, HiPhi, Rising Auto, ZEEKR, Voyah, Li Auto, Lynk & Co, Tank, NIO, and Xpeng. In 2023, over 20 million vehicles are equipped with automotive voice, with the installation rate higher than 80%.

Installations and Installation Rate of Automotive Voice, 2018-Jan.-Sep. 2023



Source: ResearchInChina

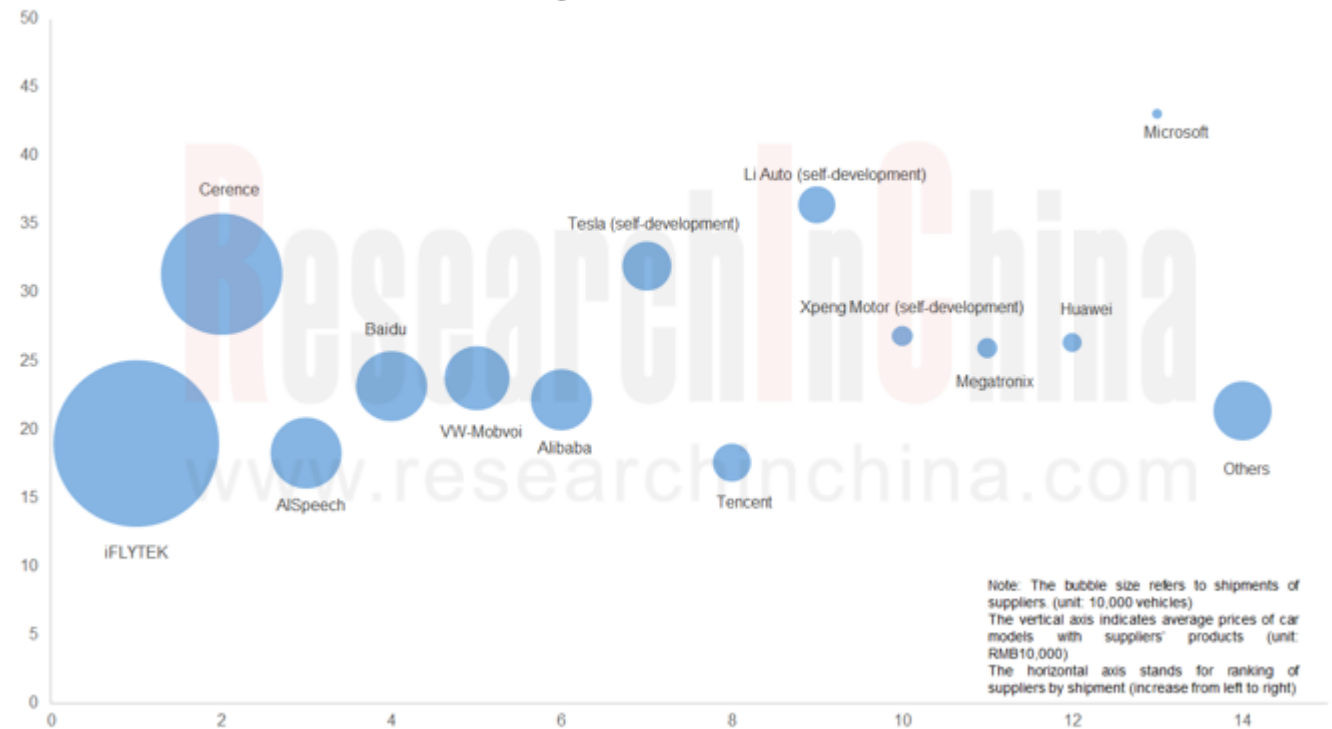
Automakers' self-development of voice facilitates the reshaping of the voice supply chain

2. Automakers' self-development of voice facilitates the reshaping of the voice supply chain.

OEMs' differentiated demand for intelligent automotive voice and their preference for independent development enable Tier 2 vendors in the conventional voice supply chain to cooperate directly with OEMs. Boundaries between upstream, midstream and downstream of the industry chain tend to blur. For example, the direct cooperation of automakers like GWM, ZEEKR and Wuling with AISpeech improves their installation and intelligence levels of intelligent voice.

The change in industry chain relationships makes the automotive voice competitive pattern change accordingly. By installations from January to September 2023, AISpeech that supported more than 150 models of over 30 automakers ranked third.

Competitive Pattern of Automotive Voice Suppliers, January-September 2023



Source: ResearchInChina

Note: The bubble size refers to shipments of suppliers (unit: 10,000 vehicles)
The vertical axis indicates average prices of car models with suppliers' products (unit: RMB10,000)
The horizontal axis stands for ranking of suppliers by shipment (increase from left to right)

See-and-speak function becomes a standard configuration, and advanced functions are available on cars.

3. See-and-speak function becomes a standard configuration, and advanced functions such as parallel instruction, cross-sound-zone inheritance, offline voice, and out-of-vehicle voice are available on cars.

In ResearchInChina's China Automotive Voice Industry Report, 2021-2022, "see-and-speak" was only installed by some emerging carmakers and leading Chinese independent brands, the longest continuous conversation duration was only 90 seconds, and dual-sound-zone recognition was still the mainstream solution.

In 2023, "see-and-speak" has become a standard configuration in emerging carmakers' flagship models, with up to 120-second continuous dialogue. Xpeng Motor has also introduced the "Full-time Dialogue at Driver's Seat" function (when turned on, it allows the driver to see and speak when looking at the center console screen, without needing to wake up the content on the screen). Meanwhile, four-sound-zone recognition has become a new mainstream solution, and Li Auto and Xpeng Motor also introduced six-sound-zone recognition solutions.

In addition, more advanced voice functions became available on cars in 2023.

Parallel instruction: support up to 10 actions in one instruction;

Cross-sound-zone inheritance: available on models of Xpeng, ZEEKR, and Li Auto (cross-sound-zone inheritance: when a person finishes an instruction, if other passengers want to continue, they can trigger this function by saying "I want too").

Offline instruction: more controllable content. Jiyue 01 supports all-zone, full offline voice. In offline state, Jiyue 01 still enables extremely fast interaction with occupants.

Out-of-vehicle voice: this function in Changan Nevo A07 allows for voice control on trunk, windows, music, air conditioning, pull-out/in, and other functions; this function in Jiyue 01 allows for voice control on car/parking, air conditioning, audio, lights, windows, doors, tailgate, and charging cover.

Comparison of Automotive Voice Functions between Some Brands

Comparison of Automotive Voice Functions between Some Brands

Brand	Voice Functions				
	See-and-speak	Continuous dialogue (s)	Sound source localization	Voice tones	Advanced voice functions
Xpeng Motor	√	20/30/60/120	Four-sound-zone	10	Up to 4 actions in one instruction; Multiple simultaneous dialogues Cross-sound-zone inheritance; full-time dialogue; Offline dialogue; speedy dialogue; Voice skill customization
Li Auto	√	20	Six-sound-zone	7	Cross-sound-zone inheritance; Voice + gesture recognition
AITO	√	10/20/30/60	Four-sound-zone	3	Up to 5 actions in one instruction; Multimode adaptive pause solution; Active services
ZEEKR	√	√	Four-sound-zone	11	Parallel instruction; cross-sound-zone inheritance
IM Motor	√	30/60/90	Four-sound-zone	/	Parallel instruction;
Denza	√	20	Four-sound-zone	3	Parallel instruction; offline instruction

Source: ResearchInChina

Voice interaction is the first stop for foundation models to get on vehicles in intelligent cockpit scenarios

4. Voice interaction is the first stop for foundation models to get on vehicles in intelligent cockpit scenarios.

The boom of ChatGPT allows the related foundation model technology to rapidly extend from AI to all other sectors. In 2023, foundation models gain pace in automotive industry, and quite a few automakers are exploring the opportunities to implement foundation models in intelligent cockpit, intelligent driving and other scenarios.

In intelligent cockpit scenarios, voice interaction is the first stop for foundation models to get on vehicles. In February 2023, Baidu released a Chinese version of ChatGPT - ERNIE Bot, and brands like GWM, Geely, and Voyah followed; in April 2023, Alibaba disclosed that AliOS intelligent vehicle operating system has been connected to Tongyi Qianwen foundation model for testing, and will later be applied by IM Motors; in August 2023, in Huawei HarmonyOS 4.0, intelligent assistant Xiaoyi was connected to Pangu model for the first time, mainly to improve capabilities of intelligent interaction, scenario arrangement, language understanding, productivity and personalized service.

Besides conventional Internet companies, voice providers as important foundation model players such as iFLYTEK, AISpeech and Unisound have also launched related products.

iFLYTEK Spark cognitive foundation model has six core capabilities: penetrative understanding of multi-round dialogues, knowledge application, empathic chat & dialogue, self-guided reply in multi-round dialogues, file-based rapid learning of new knowledge, and evolution based on correction opinions of massive users;

AISpeech DFM-2 is an industry language foundation model with generalized intelligence. In the field of in-vehicle interaction, AISpeech integrates Lyra automotive voice assistant with DFM-2, which significantly improves capabilities in planning, creation, knowledge, intervention, plug-in, multi-level semantic dialogue, and documentation, and supports multi-modal, multi-intent, multi-sound-zone, and all-scenario multi-round continuous dialogues.

Foundation Model Product Layout of Major Providers

Foundation Model Product Layout of Major Providers

Provider	Foundation Model	Release Time	Highlights	Cooperative Automaker
iFLYTEK	Spark Cognitive Foundation Model	May 2023	<ul style="list-style-type: none"> Six core capabilities: penetrative understanding of multi-round dialogues, knowledge application, empathic chat & dialogue, self-guided reply in multi-round dialogues, file-based rapid learning of new knowledge, and evolution based on correction opinions of massive users; "1+N" architecture: "1" is a base platform for general cognitive intelligent foundation model algorithm R&D and high-efficiency training, and "N" is dedicated foundation models applied to education, medical care, human-computer interaction, office, and other fields. 	Chery (first installed in STERRA ES), etc.
AISpeech	DFM-2	Jul. 2023	<ul style="list-style-type: none"> General intelligent industry language foundation model for vertical fields, with three key features: scenario specialization, model specialization, and experience specialization; Integrate with DUI platform to improve five core capabilities: user-oriented individualized multi-personality interaction, DocDFM - reliable knowledge query based on reading comprehension, accurate inference and decision enhanced by external information sources, automatic process planning and execution for complex tasks, and general semantic understanding based on deep cognition; Launch of foundation model application platform, linking DUI platform and DFM-2 foundation model for rapid implementation and application of foundation model skills. 	Not disclosed yet

Unisound	Shanghai Foundation Model	May 2023	See details in the report
Mobvoi	Monkey.Mobvoi	Apr. 2023	
Pachira	Pachira Shensi	Oct. 2023	
Tencent	Hunyuan	Sep. 2023	
Baidu	ERNIE Foundation Model	2022	
Alibaba	Tongyi Foundation Model	2022	
Huawei	Pangu Foundation Model	Sep. 2023	

Source: ResearchInChina

Table of Content (1)

1 Overview of Automotive Voice Industry

- 1.1 Overview of Automotive Voice
- 1.2 Application Scenarios of Automotive Voice
- 1.3 Automotive Voice Technologies
- 1.4 Automotive Voice Interaction Architecture
- 1.5 Automotive Voice Common Interaction Functions
- 1.6 Automotive Voice Development Factors
- 1.7 Development History of Automotive Voice
- 1.8 Automotive Voice Industry Chain Evolution
- 1.9 Automotive Voice Industry Chain
- 1.10 Market Size Forecast (2023-2026)
- 1.11 Voice Providers Market Rankings
- 1.12 Other Voice Technologies

2 Automotive Voice Applications for OEMs

- 2.1 Voice Function Comparison of OEMs
- 2.2 Summary of Voice Development Models by OEMs
- 2.3 OTA Voice Functions of OEMs
- 2.4 Xpeng Motor
 - 2.4.1 Automotive Voice-enabled Benchmark Models
 - 2.4.2 Automotive Voice Functions
 - 2.4.3 Voice Technology
 - 2.4.4 Self-developed Voice Architecture
 - 2.4.5 Self-developed Voice Basic Capabilities
 - 2.4.6 Automotive Voice Partners
- 2.5 Li Auto
 - 2.5.1 Automotive Voice-enabled Benchmark Models

- 2.5.2 Automotive Voice Skills
- 2.5.3 Vehicle Control Functions
- 2.5.4 Self-developed Voice Technology
- 2.5.5 Foundation Model
- 2.5.6 Cockpit Interaction Planning
- 2.5.7 Automotive Voice Partners

- 2.6 NIO
- 2.7 AITO
- 2.8 Aion
- 2.9 Rising Auto
- 2.10 Jiyue
- 2.11 ZEEKR
- 2.12 IM Motor
- 2.13 Denza
- 2.14 Leap Motor
- 2.15 Neta Auto
- 2.16 Geely
- 2.17 GWM
- 2.18 Changan
- 2.19 Chery

3 Automotive Voice Providers

- 3.1 Summary of Automotive Voice Providers: Market Position & Technical Competitiveness & Foundation Model Layout
- 3.2 iFLYTEK
 - 3.2.1 Profile
 - 3.2.2 Intelligent Vehicle Business Performance

Table of Content (2)

- 3.2.3 Intelligent Vehicle Core Technology
- 3.2.4 Voice Interaction Full Link Technology
- 3.2.5 Automotive Interaction Development Plan
- 3.2.6 Text-To-Speech (TTS) Technology
- 3.2.7 Interaction Model
- 3.2.8 Application of Interaction Foundation Model in Intelligent Cockpit
- 3.2.9 Cockpit OS Enhanced by Foundation Models
- 3.2.10 Knowledge Graph of iFLYTEK Interaction Foundation Model
- 3.2.11 Interaction Foundation Model Core Capabilities
- 3.2.12 Interaction Foundation Model Enabling Automotive Human-Machine Interaction
- 3.2.13 Accumulation in Cognitive Intelligent Foundation Model Technology
- 3.2.14 "1+N" System
- 3.2.15 Multilingual Interaction System
- 3.2.16 Support for Automotive Minor Languages
- 3.2.17 Open Platform Voice Technology Support
- 3.2.18 Out-of-vehicle Voice Interaction System

- 3.3 Cerence
 - 3.3.1 Automotive Voice Recognition Hardware Framework
 - 3.3.2 Vehicle-Cloud Integration Solution
 - 3.3.3 Core Technology
 - 3.3.4 ARK Main Content
 - 3.3.5 SSE
 - 3.3.6 Drive
 - 3.3.7 Automotive Voice Interaction + AI Solution
 - 3.3.8 Co-Pilot
 - 3.3.9 Biometrics
 - 3.3.10 ICC
 - 3.3.11 Out-of-vehicle Voice Interaction
 - 3.3.12 TTS
 - 3.3.13 Other Voice Solutions
 - 3.3.14 Product Development Roadmap (2023~)

- 3.4 AISpeech
 - 3.4.1 Profile
 - 3.4.2 Voice and Language Key Technologies
 - 3.4.3 "Cloud + Chip" Integration Strategy
 - 3.4.4 Customized Development Platform for All-link Intelligent Dialogue System: DUI
 - 3.4.5 Industry Language Model: DFM
 - 3.4.6 Intelligent Telematics Solutions
 - 3.4.7 Automotive Voice Assistant
 - 3.4.8 Intelligent Cockpit Products
 - 3.4.9 Cooperation Model Cases

- 3.5 Unisound
 - 3.5.1 Intelligent Automotive Solutions
 - 3.5.2 Foundation Model
 - 3.5.3 Voice Technology Capabilities
 - 3.5.4 TTS
 - 3.5.5 Automotive Voice Solution Business Models
 - 3.5.6 Core Technology
 - 3.5.7 Automotive Voice Chip
 - 3.5.8 Automotive Voice Solution Supporting

- 3.6 txzing.com
- 3.7 VW-Mobvoi

Table of Content (3)

- 3.8 Mobvoi
- 3.9 Pachira
- 3.10 Tencent

- 3.11 Baidu
 - 3.11.1 Core Voice Technology
 - 3.11.2 Voice Chip
 - 3.11.3 DuerOS
 - 3.11.4 DuerOS Empowered by Foundation Models
 - 3.11.5 ERNIE Foundation Model Enabled Cockpit Voice Interaction
 - 3.11.6 ERNIE Foundation Model Analysis

- 3.12 Alibaba
- 3.13 Huawei
- 3.14 Volcano Engine
- 3.15 Microsoft
- 3.16 VoiceAI

4 Automotive Voice Industry Chain

- 4.1 Platform Integration: PATEO
- 4.2 Platform Integration: Tinnove
- 4.3 Voice Processing Engine: SinoVoice

- 4.4 Voice Processing Engine: Megatronix
 - 4.4.1 Product Layout
 - 4.4.2 Automotive Voice SmartMega? VOS Module
 - 4.4.3 Automotive Voice Customized and Cooperation Modes
 - 4.4.4 Implemented Model Cases

- 4.5 Data Collection / Annotation: Haitian Ruisheng
 - 4.5.1 Voice Business
 - 4.5.2 Structure of Training Dataset
 - 4.5.3 Speech Services: Data Collection Services
 - 4.5.4 Speech Services: Data Annotation Services

- 4.6 Data Collection / Annotation: Testin
- 4.7 Data Collection / Annotation: DataBaker
- 4.8 Corpus: Magic Data
- 4.9 Chip: Horizon
- 4.10 Chip: ShensiliCon
- 4.11 Chip: Chipintelli
- 4.12 Voice Chip: Rockchip
- 4.13 Voice Chip: WUQi Micro
- 4.14 Voice Chip: LAPIS Semiconductor

5 Development Trends of Automotive Voice

- 5.1 Trend 1
- 5.2 Trend 2
- 5.3 Trend 3
- 5.4 Trend 4
- 5.5 Trend 5
- 5.6 Trend 6
- 5.7 Trend 7
- 5.8 Trend 8



Beijing Headquarters

TEL: 13718845418

Email: report@researchinchina.com

Website: [ResearchInChina](http://ResearchInChina.com)

WeChat: Zuosiqiche



Chengdu Branch

TEL: 028-68738514

FAX: 028-86930659

