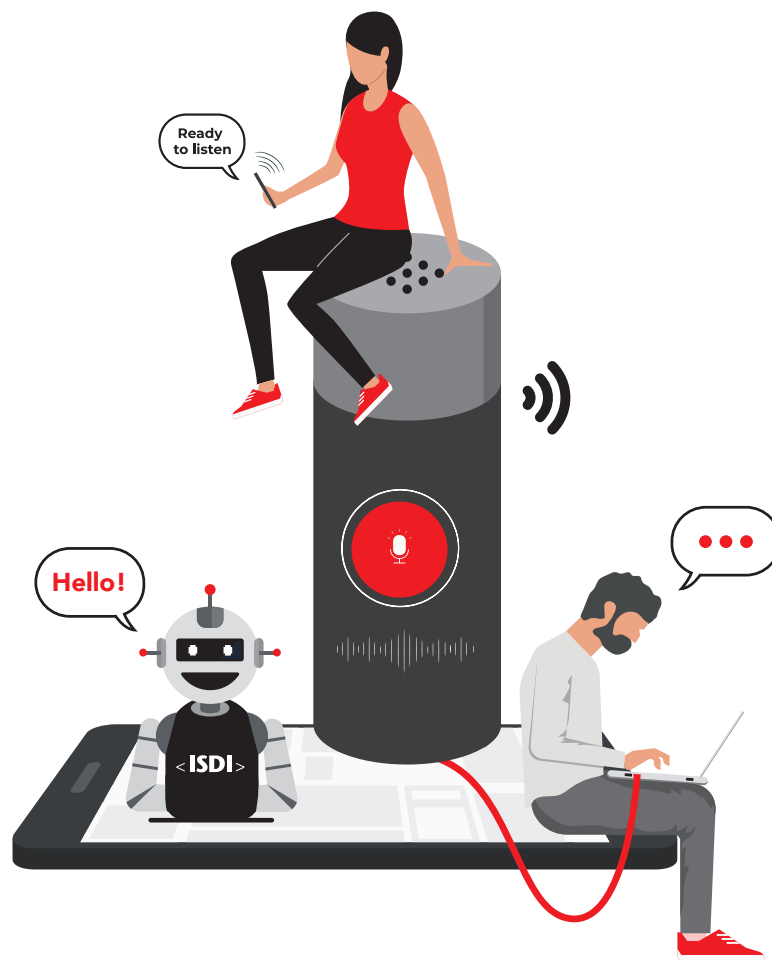


EL USO DE LA **TECNOLOGÍA** **DE VOZ**

EN ESPAÑA



webedia.

< ISDI >

ÍNDICE

INTRODUCCIÓN - ABSTRACT	<06>
RESULTADOS	<09>
PROYECCIONES	<53>
CONCLUSIONES	<63>

SOBRE EL ESTUDIO



INFORME DE:

ISDIGITAL.DATA - Febrero de 2020.



METODOLOGÍA

Análisis poblacional: 1.029 entrevistas online.

Respuesta de la tecnología: 2.800 preguntas en 6 verticales.



OBJETIVO:

Obtener una radiografía de los hábitos de consumo y comportamiento de los españoles en el uso de la tecnología de voz, así como la respuesta actual de los principales players.



AGENCIAS COLABORADORAS:

Análisis poblacional: PUNTO DE FUGA para ISDI.

Respuesta de la tecnología: WEBEDIA.

< > Introducción - Abstract

Usos, percepción y retos de la tecnología de voz en España

La tecnología de voz parece estar altamente implementada entre nosotros: más de la mitad de la población española utiliza ya comandos de voz para realizar acciones a través de diversos dispositivos inteligentes.

Sin embargo, cuando se analiza en profundidad los usos que se dan a esta reciente tecnología encontramos un panorama aún con múltiples retos para su desarrollo, sobre todo comparado con su potencial.

El Estudio ISDI-WEBEDIA sobre el uso de la Tecnología de voz en España, ofrece una visión profunda e innovadora en este campo aunando, por primera vez:

1. La observación comportamental del usuario de dispositivos de voz en smartphones, en el coche y altavoces inteligentes, identificando el grado de penetración, uso por dispositivo, aplicaciones, frenos y perspectiva de uso futuro.
2. Análisis de la respuesta tecnológica de 2 de los principales altavoces inteligentes del mercado (Google Home Mini y Amazon Echo Dot) para conocer la precisión y fuentes de información utilizadas por cada dispositivo.

Resultados principales del estudio

El 54% de la población española utiliza actualmente la tecnología de voz para realizar alguna acción en sus aparatos inteligentes. Además, un 13% del segmento de no usuarios se muestra abierto a su uso en un futuro próximo.

Como **beneficios principales** los usuarios actuales señalan la practicidad a la hora de buscar información (67%) y realizar multitareas (65%), así como el ahorro de tiempo que supone (62%).

El dispositivo tecnológico en el que más se utiliza la tecnología de voz es el Smartphone (87%), seguido del dispositivo de voz del coche (22%) y del altavoz inteligente (21%).

De forma global escuchar música, buscar información y realizar llamadas son las **tareas principales** en las que el usuario aplica este tipo de tecnología, siendo el **split por dispositivo** el siguiente:

- Usos en smartphone: buscar información (61%), hacer llamadas (51%), poner la alarma (40%).
- Usos en coche: hacer llamadas (60%), escuchar música (42%), oír la radio (29%).
- Usos en altavoz inteligente: escuchar música (81%), oír la radio (48%), buscar información (45%).

Los **principales frenos** que se declaran a la hora de ampliar el uso de la voz a otras funciones más complejas son:

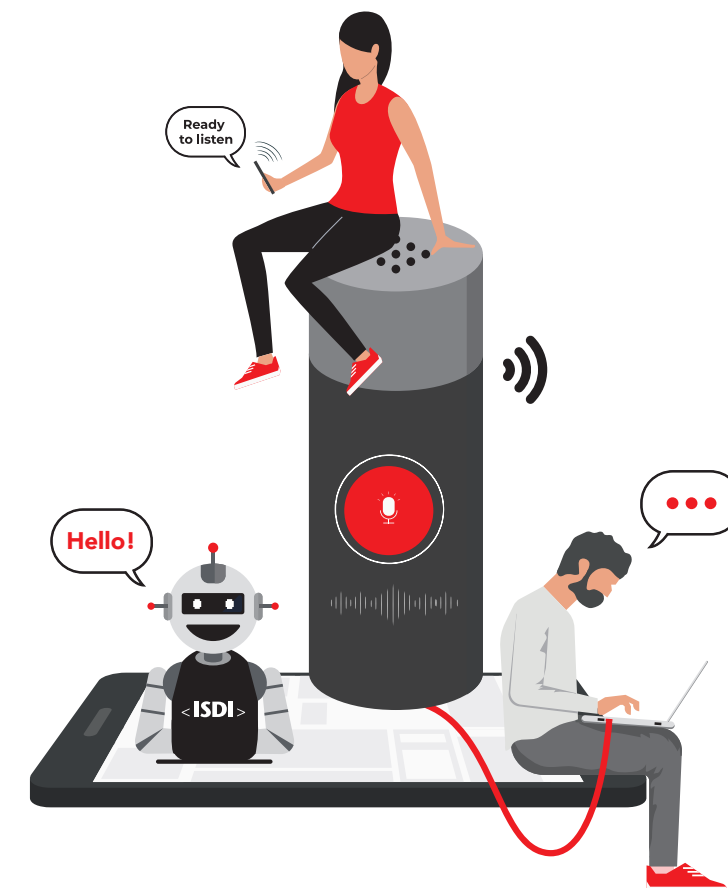
- La seguridad: el 64% de los usuarios no se sienten seguros con esta tecnología si tienen que realizar gestiones importantes.
- La privacidad: el 42% no se fía del trato que se le puede dar a sus datos.
- Fallos en su usabilidad y fiabilidad: 3 de cada 4 usuarios han experimentado algún problema en el uso de la tecnología de voz. De una forma específica fallos al detectar la voz y respuesta incorrecta (26%).

Respecto a estos dos puntos concretos y tras el estudio de respuesta de los dos altavoces inteligentes, se observa que tanto la calidad en las respuestas está relacionada con 2 factores: fuente de donde se obtiene la información y sector al que pertenece la pregunta.

En concreto:

- De forma general el asistente Google Home Mini muestra una tasa 18% superior en calidad de respuesta que Amazon Echo Dot.
- En el caso del Google Home Mini el 57% de las respuestas provienen de una fuente nativa vs 43% procedente de contenido web externo. Finanzas (86%) y Entretenimiento (84%) son los sectores con las respuestas más precisas.
- En el caso del Amazon Echo Dot el 85% de las respuestas provienen de una fuente nativa vs 15% procedente de contenido web externo. Entretenimiento (84%) y Gran Consumo (53%) son los sectores con las respuestas más precisas.

RESULTADOS



<

>

Histórico en el análisis del uso de la tecnología de voz en España

El uso de la tecnología de voz en España es una tendencia creciente que se refleja en la aparición de múltiples estudios que buscan aportar visibilidad sobre esta área.

Históricamente este tipo de estudios se han centrado en el análisis poblacional del uso de las tecnología de voz. Sin embargo, nunca se había analizado la calidad de respuesta que se ofrece desde los principales players, así como el impacto de ésta en la percepción del consumidor.

	Estudio 1	Estudio 2	Estudio 3	Estudio 4
Fecha	2018	2019	2019	2019
Ámbito	Mundial	Nacional	Nacional	Nacional
Muestra	1.000	2.088	1.053	150.000
Áreas	Interfaces de voz Consumidores y marcas	Altavoces inteligentes	Altavoces Inteligentes Contenidos audio	Consumo de medios

1. Ampliar **perspectiva** demográfica



Observación comportamental nacional y autonómica
Análisis de uso por dispositivo inteligente
Percepción de beneficios y desventajas

2. Analizar respuesta **tecnológica**



Análisis respuesta tecnológica altavoces inteligentes
Análisis procedencia de la información
Análisis sectorial

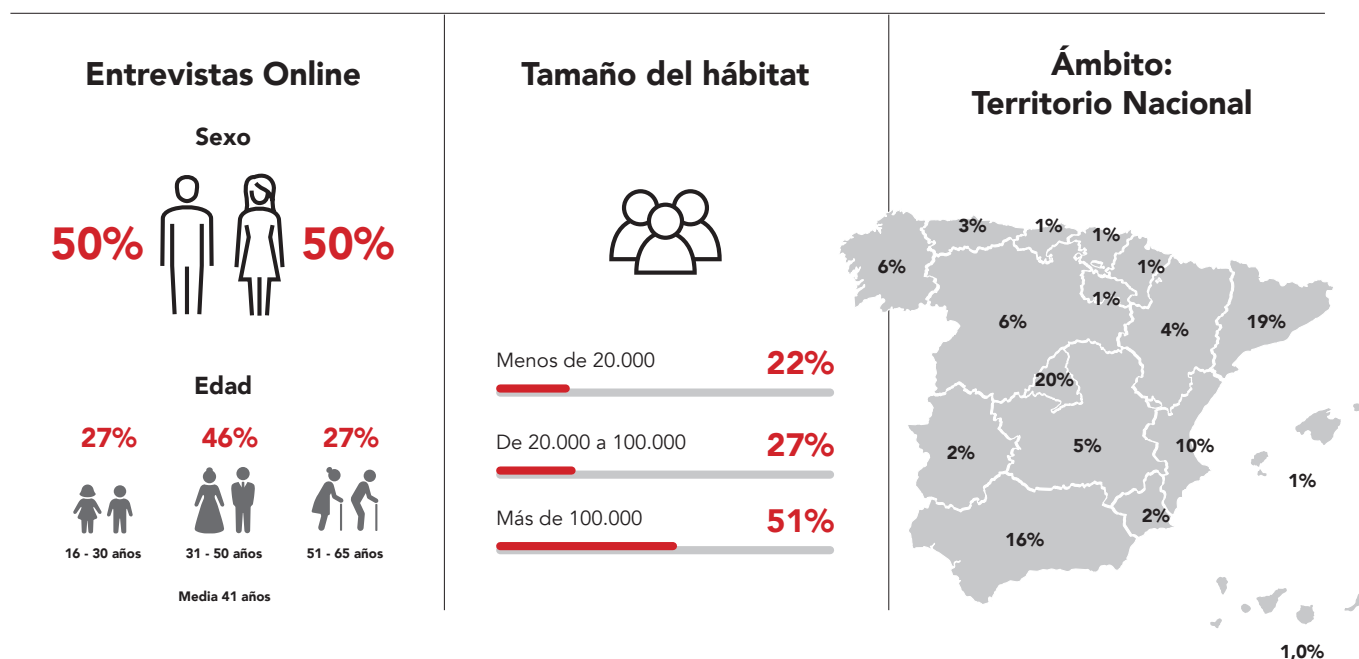
3. Proyecciones y **conclusiones**



Proyecciones de uso
Flujo de conclusiones
Oportunidades

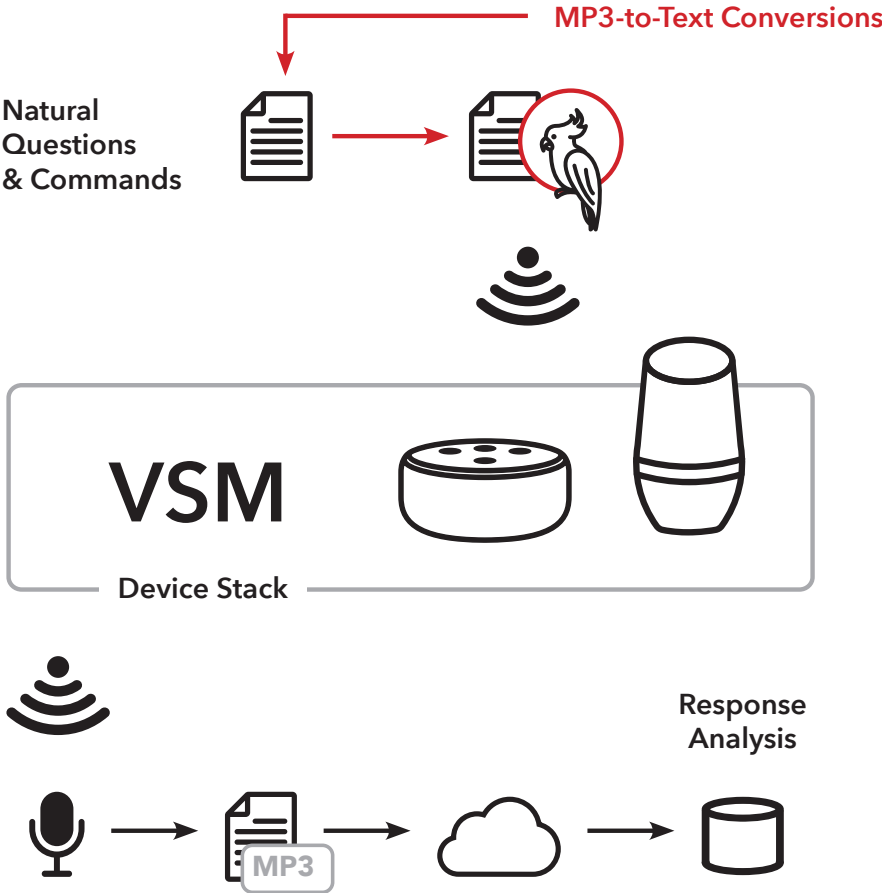
Se han realizado 1.029 entrevistas online a una muestra representativa de la población española en todo el territorio nacional. El 50% de los encuestados son hombres y el 50% mujeres. Por edades la distribución también es representativa: un 27% de personas de entre 16 y 30 años, un 46% concentrado en el intervalo de 31 a 50 años y un 27% entre 51 y 65 años. A nivel de CC.AA el mayor peso recae en Andalucía, Cataluña y la Comunidad de Madrid, con el 55% de los entrevistados. Así mismo, el 51% de los entrevistados reside en poblaciones de más de 100.000 habitantes.

1.029 entrevistas
a una muestra representativa de la población española



< > Metodología para el análisis de la **respuesta de la tecnología**

Tal y como vemos en el gráfico de metodología, se lanzaron más de 2.800 preguntas a dos de los principales altavoces inteligentes del mercado: Google Home Mini (Google) y Amazon Echo Dot (Amazon). Los resultados fueron analizados mediante nlp. Las preguntas se seleccionaron a partir de las consultas más frecuentes en 8 sectores verticales. Esta metodología está inspirada en un estudio anual que se hace en el mundo anglosajón por la consultora Perficient Digital, siendo la primera vez que se hace en nuestro país y en castellano.



< > El **interés** general por la Tecnología

Los españoles muestran un grado de interés por la tecnología bastante alto puntuándolo en 7.32/10.

A nivel autonómico Andalucía es la comunidad que presenta un interés más alto con un 7,56%. Se observa también que, mientras la edad no parece marcar diferencias en grado de conexión con la tecnología, sí se encuentran diferencias en cuanto a género: las mujeres capitalizan los segmentos de poca - media vinculación y los hombres los segmentos de vinculación media - alta.



Comunidad Autónoma



< > Penetración del **uso de la tecnología de voz**
en España

El 54% de la población española declara hacer uso de la tecnología de voz para realizar alguna acción con sus aparatos inteligentes. Si analizamos este dato a nivel autonómico Andalucía es la comunidad que mayor grado de interacción muestra (59%). La Comunidad de Madrid, Cataluña o la Comunidad Valenciana se sitúan por debajo de la media con un 51%, 50% y 49% respectivamente.

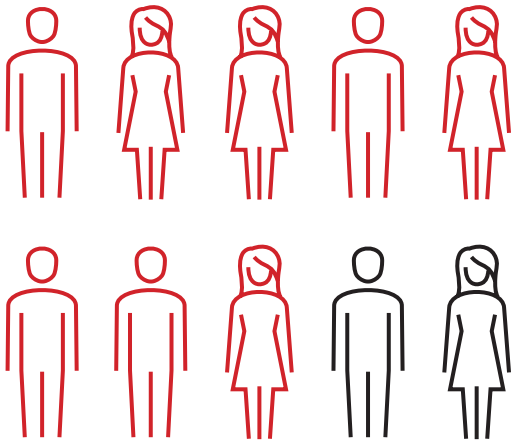


Comunidad Autónoma



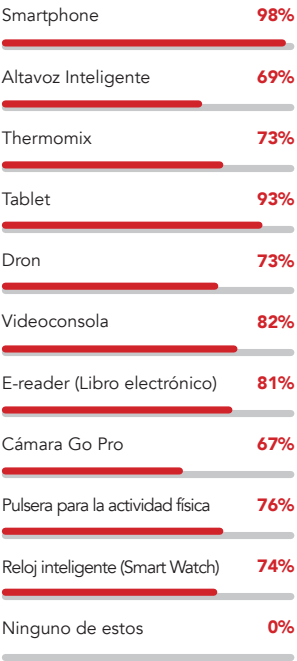
< > Presencia de **dispositivos tecnológicos**

Cuando analizamos la penetración de dispositivos tecnológicos en la población española el smartphone es sin duda el elemento más conocido (98%) y a su vez, el más poseído (95%). La tablet (93% - 67%) y la videoconsola (82% - 46%) ocupan el segundo y tercer puesto. En cuanto a los dispositivos específicos de tecnología de voz (altavoz inteligente), destaca la diferencia existente entre el porcentaje de conocimiento (69%) y la tasa de posesión (15%).



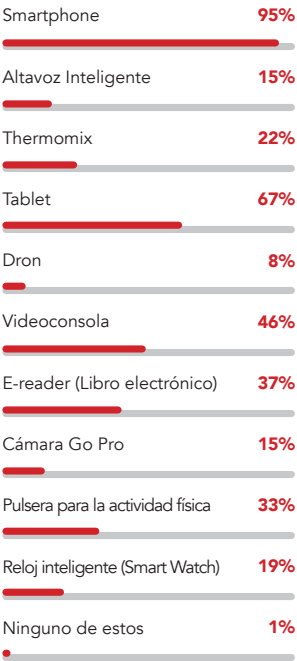
7 de cada 10 conocen lo que es un **"altavoz inteligente"**, **teniéndolo en posesión casi 2 de cada 10**

Conocimiento



Base total (n=1029)

Tiene

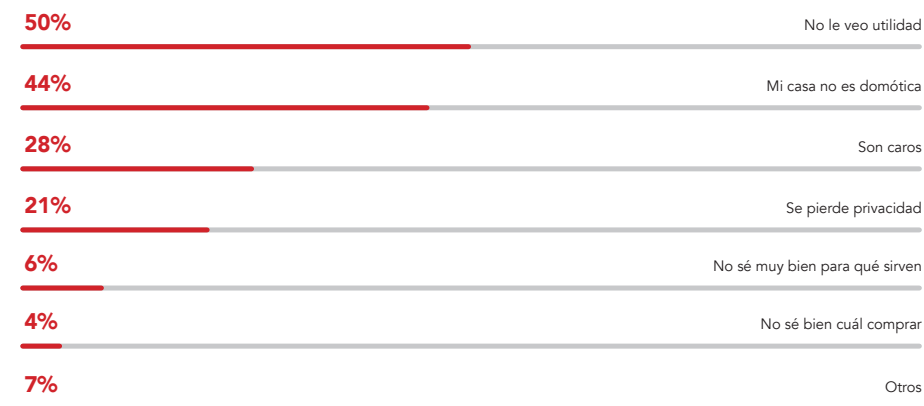


Base tiene algún dispositivo (n=1026)

< > La percepción de **propuesta de valor** de los **altavoces inteligentes**

Analizando de una forma más profunda la diferencia entre conocimiento y posesión en el segmento de los altavoces inteligentes se observa que la falta de utilidad (50%), el bajo nivel de domotización de los hogares (44%) y el precio actual de los dispositivos (28%) son los principales motivos que hacen que la penetración de este dispositivo se sitúe en el 15%.

Motivos no dispone de un altavoz inteligente



(Base no tiene altavoz inteligente pero sí lo conoce=553)

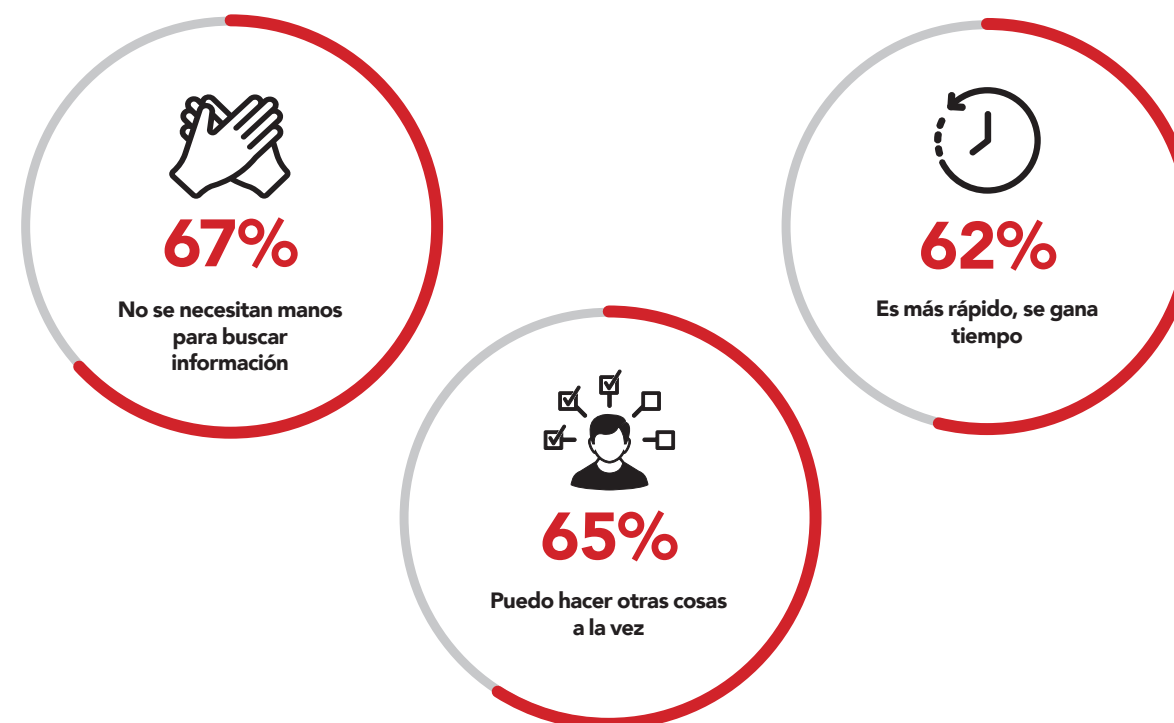
15% de la población posee un altavoz inteligente

La falta de una **Propuesta de Valor** en torno a la tecnología de Voz, está justificando un freno en la compra de altavoces inteligentes

< > **Beneficios** percibidos por el usuario de tecnología de voz

En cuanto a los beneficios que percibe el usuario de la tecnología de voz la practicidad de la interacción verbal (67%), la posibilidad de la realización de multitareas (65%) y el ahorro de tiempo (62%) son las principales ventajas declaradas para hacer uso de esta forma de interacción con sus dispositivos.

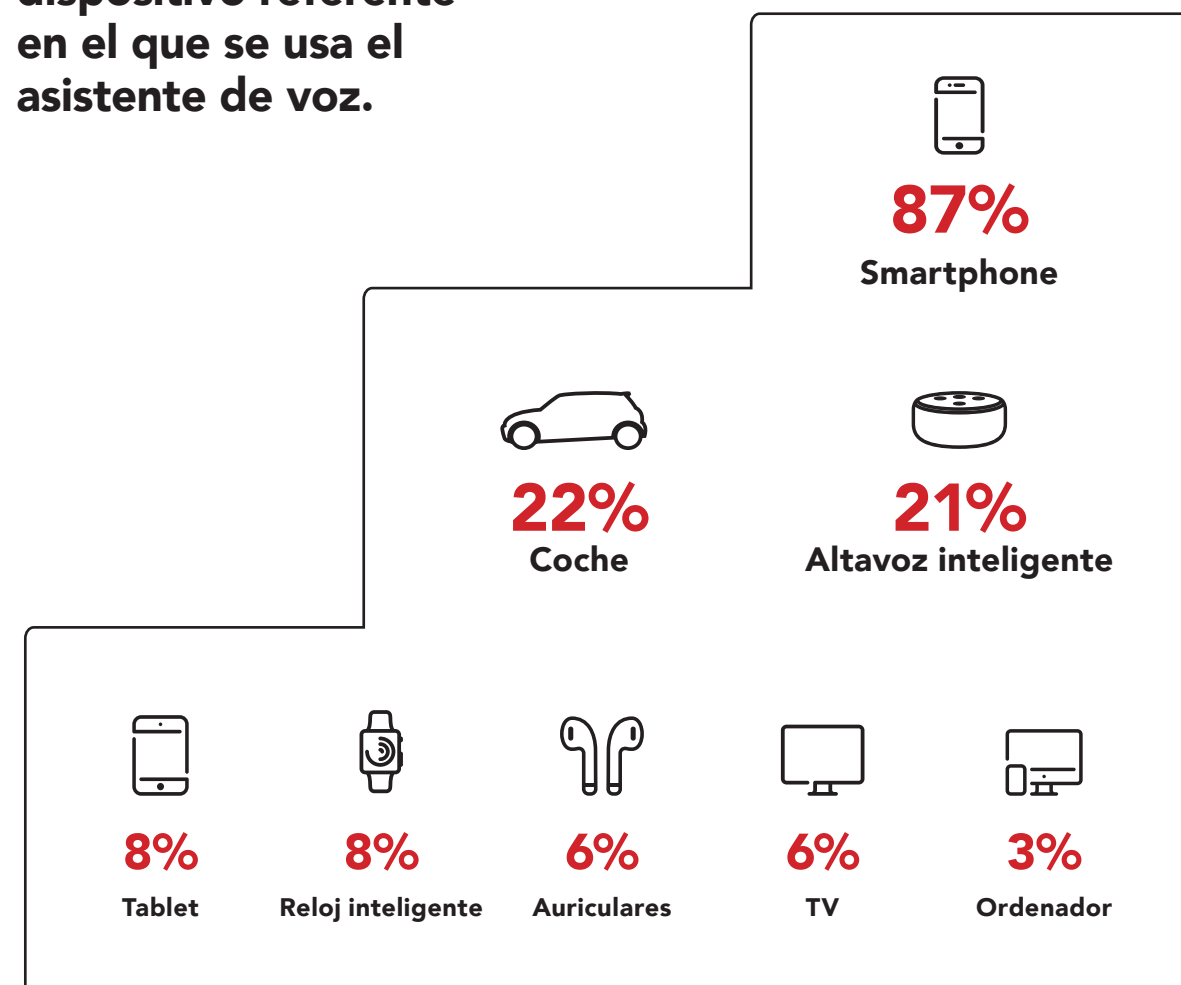
La comodidad y una notable ganancia de tiempo, las principales ventajas



< > Uso por **dispositivo** en usuarios de **asistentes de voz**

El Smartphone con un 87% es el dispositivo donde más se utiliza la tecnología de voz. Estas cifras bajan cuando se trata de otros dispositivos como el coche (22%) o los altavoces inteligentes (21%). Las tablets, auriculares, reloj inteligente, SmartTV y el ordenador tienen un uso residual por debajo del 9%.

El **SMARTPHONE**, el dispositivo referente en el que se usa el asistente de voz.



< > Posibilidades actuales de la aplicación de la tecnología de voz

La tecnología de voz nos ofrece a día de hoy múltiples aplicaciones que van desde acciones rutinarias sencillas (poner alarma, escuchar música, hacer llamadas, etc.) hasta acciones más complejas e integradas (domótica, gestión de servicios bancarios, compra de artículos, etc.). Antes de entrar en el análisis de la aplicación de la tecnología de voz por parte del usuario, es importante tener en cuenta el listado de posibilidades actuales de uso. De esta forma se puede comprender de una forma más precisa el grado actual de aplicación versus potencial.

- | | |
|---|--|
|  Buscar información (tiempo, tráfico, recetas cocina, la mejor ruta para hacer un viaje, etc.) |  Jugar a juegos |
|  Comprar productos (ropa, alimentación, del hogar, ocio, comida a domicilio...) |  Hacer llamadas – como altavoz o manos libres |
|  Gestión de servicios (bancarios, seguros, viajes, transporte, restaurantes, etc.) |  Enviar mensajes de voz |
|  Llamar a la atención al cliente de alguna compañía |  Acceder a mi calendario o agenda |
|  Escuchar música |  Controlar los dispositivos electrónicos de la casa |
|  Oír la radio |  De ayuda para mis hijos/as a la hora de estudiar |
|  Escuchar podcast |  Ayuda y compañía para una persona mayor (recuerda cuánto tiene que tomar medicinas, citas médicas, etc.) |
|  Poner la alarma |  Para que me recuerde diferentes cosas: la agenda, citas, rutinas, etc. |
|  Poner un timbre de recordatorio o temporizador (para controlar la elaboración de una comida, una actividad, etc.) |  Que me informe de las noticias del día |

<>

Usos actuales

de la tecnología de voz por dispositivo

En cierta medida, la aplicación que le da el usuario a la tecnología por voz está muy ligado a la propia esencia y usabilidad del dispositivo a través del que se ejecuta la acción: el smartphone se utiliza principalmente para búsquedas de información (61%), además de la realización de llamadas (51%). Por el contrario, la ejecución de comandos de voz a través de los altavoces inteligentes está muy enfocada a la escucha de música (81%), seguido de oír la radio (48%) y la búsqueda de información (45%). Así mismo, en el uso de la voz a través del asistente del coche la realización de llamadas (60%) y la escucha de música (42%) son las principales acciones que ejecutan sus usuarios.

	 Smartphone	 Altavoz inteligente	 Coche
Buscar información	 61%	 45%	20%
Enviar mensajes de voz	 51%	22%	23%
Hacer llamadas	 51%	28%	 60%
Escuchar música	 37%	 81%	 42%
Poner la alarma	40%	44%	6%
Escuchar la radio	14%	 48%	 29%

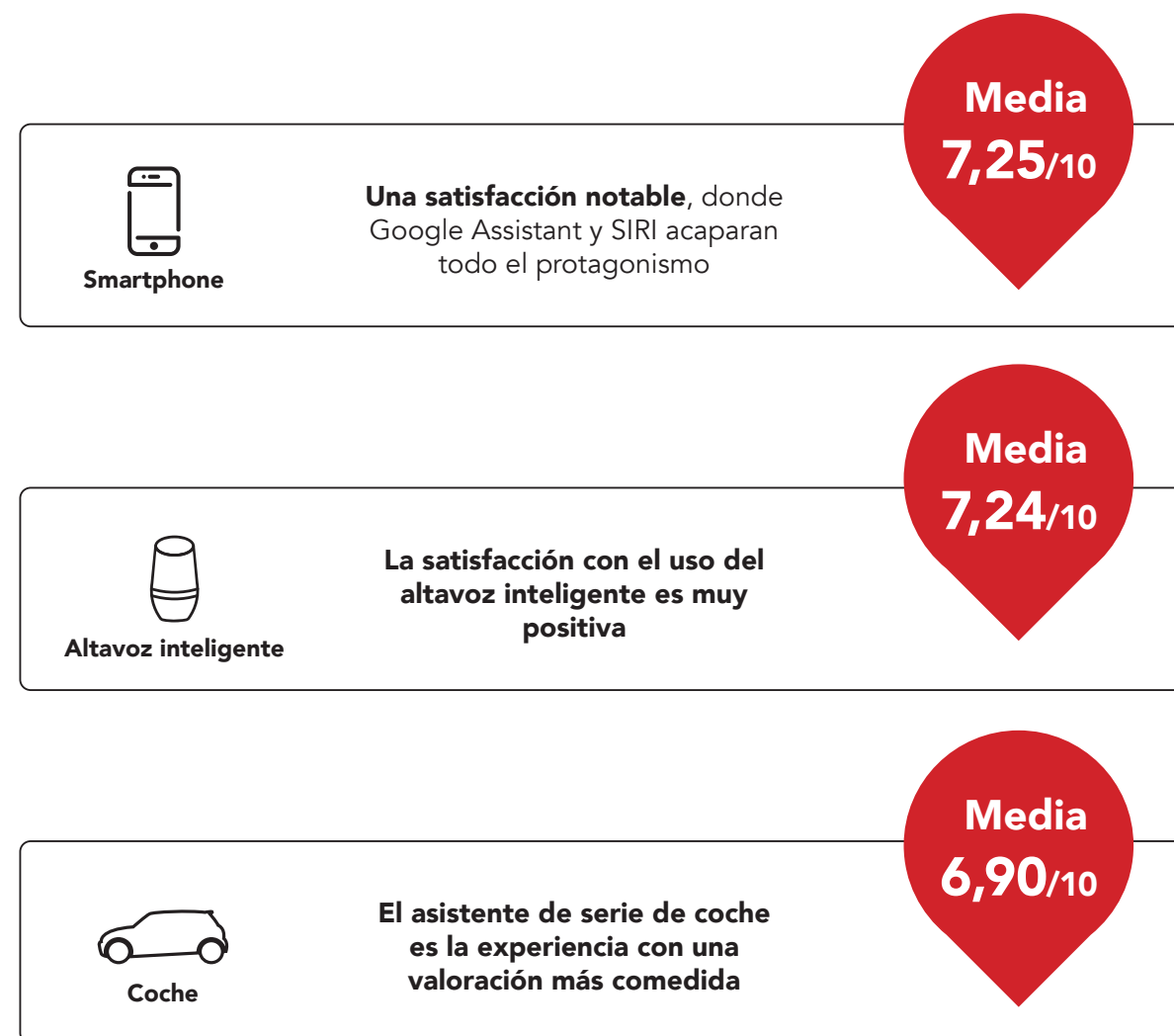
(Base asistente de voz en Smartphone=422)

Base tiene Altavoz Inteligente (n=183)

Base no tiene Altavoz Inteligente pero si lo conoce =553)

< > Valoración de la **experiencia** por parte del usuario

A día de hoy la experiencia en el uso de la tecnología por voz es muy favorable aunque no alcanza niveles sobresaliente. Este dato nos habla de una percepción por parte del usuario que aún tiene vías de mejoras abiertas. Por dispositivo la satisfacción es similar entre las experiencias vividas a través del Smartphone (7.25/10) y los altavoces inteligentes (7.24). Por el contrario, el uso de la voz en el coche genera niveles de satisfacción algo más comedidos (6,90/10).



< > Frenos y limitaciones para un uso más extensivo

Como es común en la pregunta de frenos frente al uso de la tecnología, la seguridad (64%) y la privacidad (42%), ocupan los primeros puestos en cuanto a limitaciones de uso. Sin embargo, y este dato si es específico de la tecnología de voz, 3 de 4 usuarios declaran que el haber experimentado fallos en el reconocimiento de comandos les supone uno de los principales inconvenientes para la realización de un uso más extensivo.



< > Frenos y limitaciones en el uso de altavoz inteligente

Si analizamos el factor “fallo” en el segmento de altavoces inteligentes encontramos que 3 de cada 4 usuarios declaran haber encontrado algún inconveniente durante su uso, siendo la dificultad en el reconocimiento de voz (26%) y la respuesta incorrecta del dispositivo (26%) dos de los tres factores declarados.

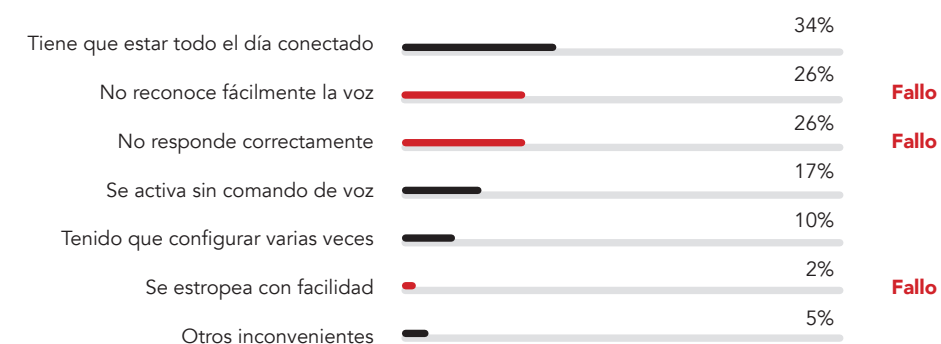
Ambos factores han sido observados de una forma concreta en nuestro análisis de respuesta de la tecnología por voz, como veremos a continuación.

ERROR EN RECONOCIMIENTO DE VOZ Y CALIDAD DE RESPUESTA PRINCIPALES FACTORES LIMITANTES

FALLOS EN FUNCIONAMIENTO



3 de cada 4 usuarios se han encontrado con algún inconveniente en su uso



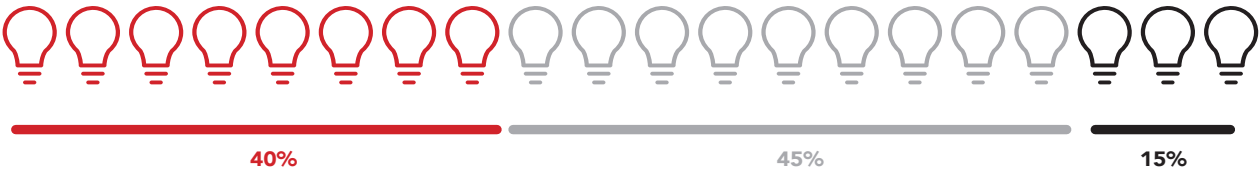
< > **Análisis de la calidad de respuesta en altavoces inteligentes**

En la realización del análisis de calidad de respuesta realizado en dos de los altavoces inteligentes con más penetración del mercado (Google Home Mini y Amazon Echo Dot) es Google quien presenta (+18%) una mejor calidad de respuesta. Al contar con toda la potencia de su buscador web, Google es capaz de ofrecer respuestas en muchos más casos de uso que Alexa. Se observa también que, de forma general, la aportación de respuestas (correctas o equivocadas) va aumentando versus el “no comprendido”, puesto que los asistentes mejoran su programación para tratar de responder cada vez más a tipos diferentes de consultas.

Google es un 18 puntos más “listo” que Alexa

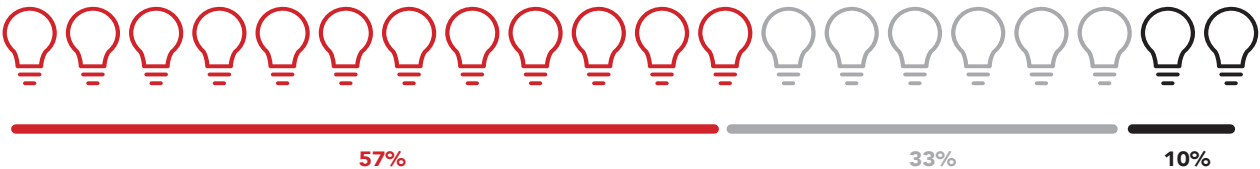


Calidad de las respuestas de Alexa



Google Home Mini

Calidad de las respuestas de Google

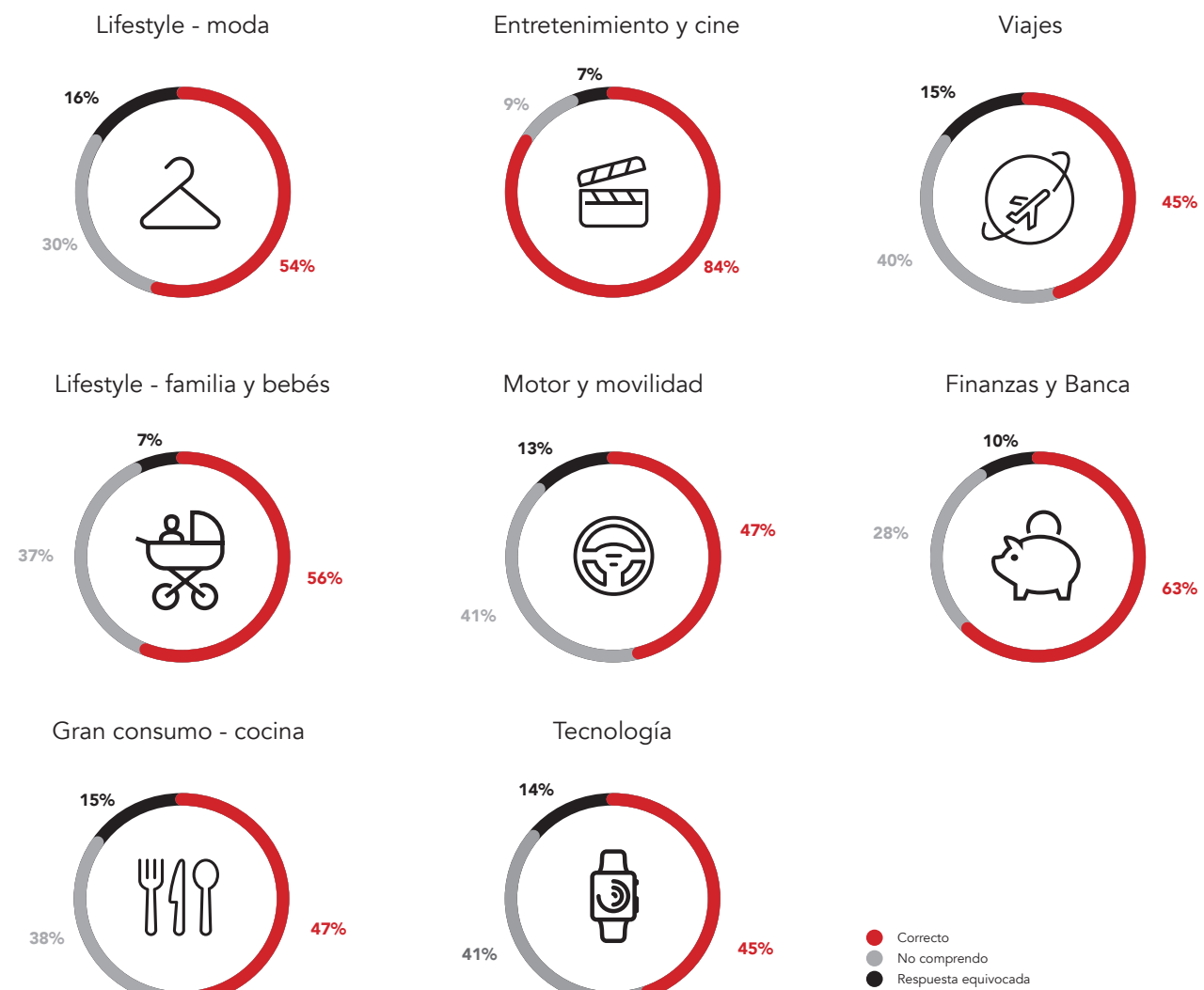


Correcto No comprendido Respuesta equivocada

< > Análisis de la **calidad de respuesta** por **sectores**

Si analizamos la calidad de respuesta entre los diferentes sectores vemos que los resultados dependen principalmente del carácter del contenido: por un lado entretenimiento dispone de multitud de datos públicos y estructurados. Por otra, "finanzas" es un vertical donde gran parte de las consultas más frecuentes son de tipo "glosario" o "búsqueda de términos" un caso de uso que ambos asistentes resuelven particularmente bien.

ISDIGITAL.DATA

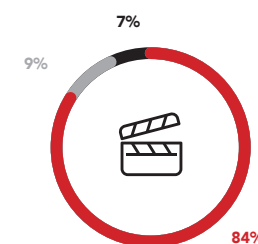


< > Análisis de la **calidad de respuesta** por **dispositivo y sectores**

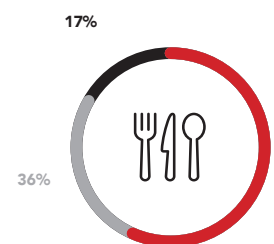
Alexa destaca en Entretenimiento - Cine gracias a su potente base de datos IMDB. En el caso de Gran Consumo - Cocina, las skills dedicadas ofrecen también una experiencia de usuario mejorada respecto a la versión nativa. Google por su lado aprovecha los contenidos mejor estructurados en su knowledge graph, que en este análisis han resultado ser Entretenimiento - Cine y Finanzas (snippets, preguntas frecuentes, etc.).

Alexa ofrece su mejor experiencia en Entretenimiento - Cine y Gran Consumo - Cocina

Entretenimiento - Cine



Gran consumo - Cocina

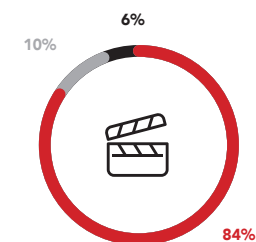


● Correcto
● No comprendo
● Respuesta equivocada

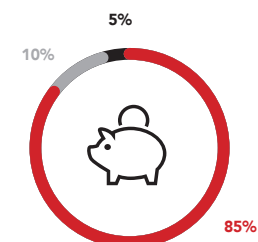
Precisión de las respuestas por sectores. Detalle Alexa.

Google por su lado destaca en Entretenimiento - Cine y Finanzas

Entretenimiento - Cine



Finanzas



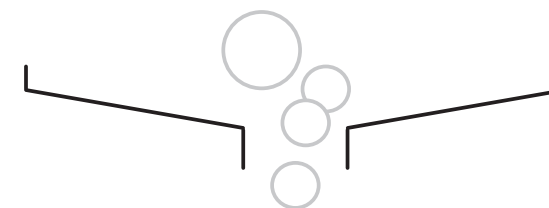
● Correcto
● No comprendo
● Respuesta equivocada

Precisión de las respuestas por sectores. Detalle Google.

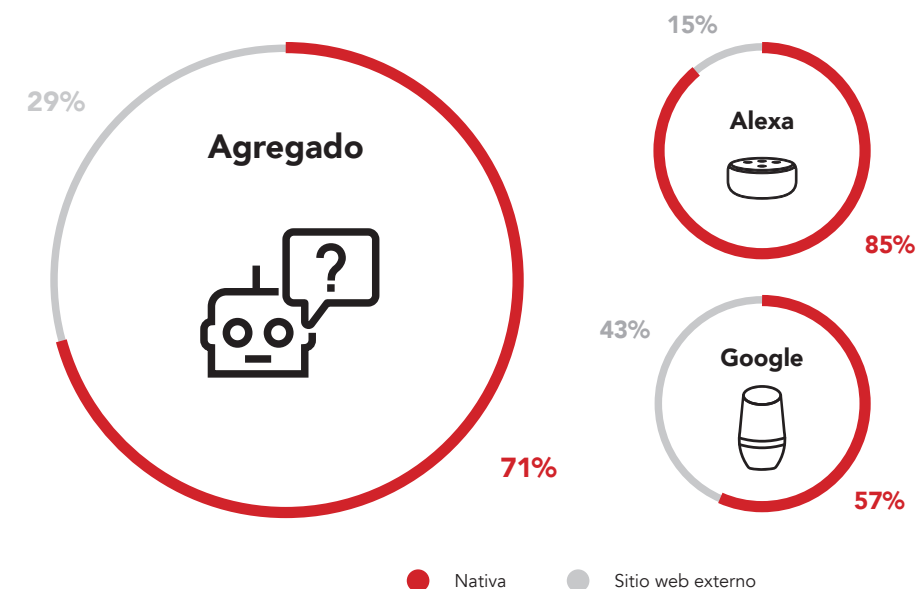
< > Origen de la información por dispositivo

Los usuarios buscan de una forma muy relevante información a través de sus asistentes. Las respuestas ofrecidas por los asistentes son por una parte nativas (fruto de la programación que hacen los equipos internos de cada plataforma) y por otra, de fuentes externas. En el caso de Google, que se está posicionando como el especialista en la búsqueda de información a través de voz, este aspecto es especialmente relevante cuando cerca de la mitad de la información que proporciona depende de fuentes externas.

El 29% de las respuestas ofrecidas por los asistentes vienen de una fuente externa



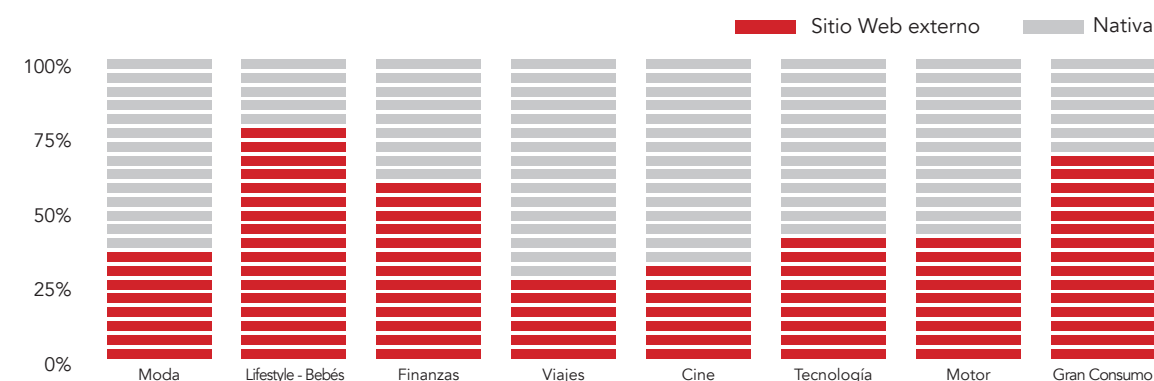
En el caso de Google, hasta un 43% de las respuestas dependen de contenido web



< > Origen de la información por dispositivo y sectores

Si observamos el reparto por verticales, vemos que algunas tienen una mayor dependencia de contenido externo en aquellos casos en los que hay menor información estructurada disponible para ofrecer de forma nativa.

Familia y Gran consumo son los verticales con mayor dependencia de contenido externo

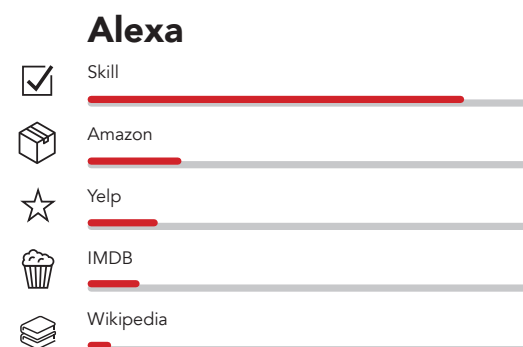


% respuestas nativas vs. fuentes externas por verticales. Agregado Alexa y Google.

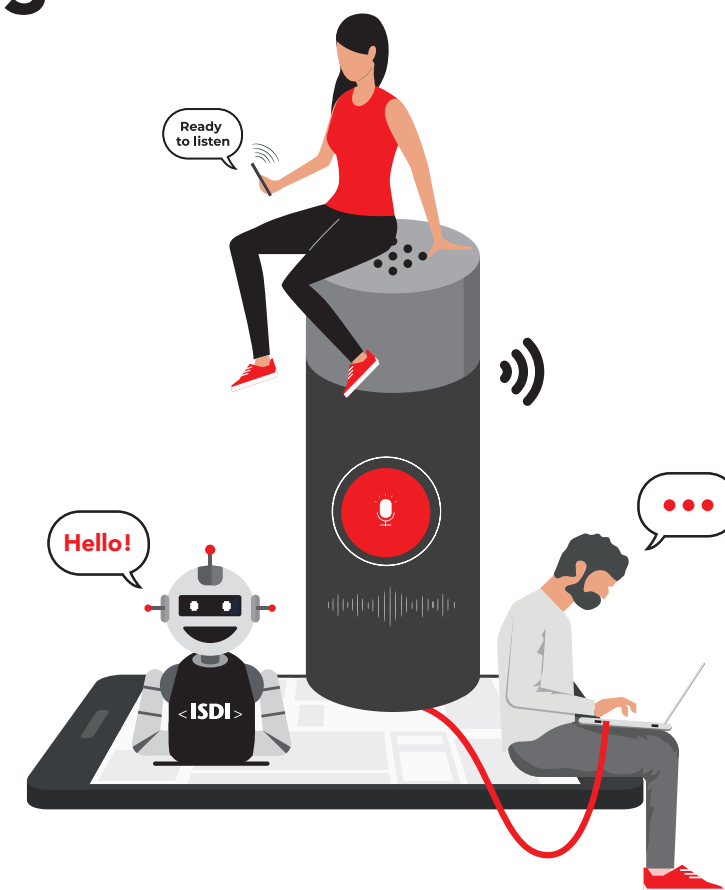
< > Partners para la obtención de información por dispositivo

¿Y cuáles son esas fuentes externas? En el caso de Amazon se trata en su mayoría de skill desarrolladas por terceros y sites participados o partners de la misma plataforma como IMDB o el propio ecommerce. En el caso de Google, destacan fuertemente las fuentes especializadas y blogs como fuente de información. A destacar, la fuerte presencia de Wikipedia como referencia para ambas plataformas.

Alexa depende sobre todo de **partners externos y skills**



PROYECCIONES

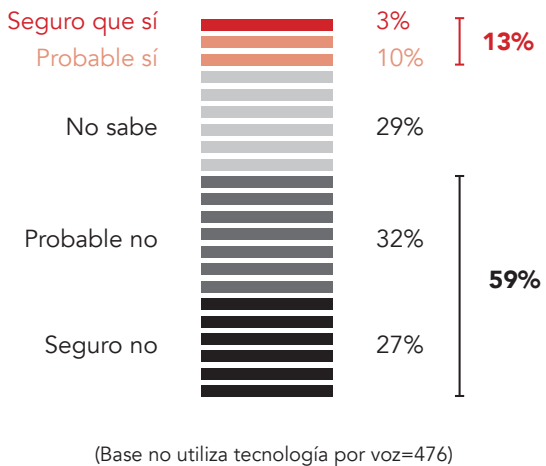


< > **Interés a futuro**
del no usuario por la tecnología de voz

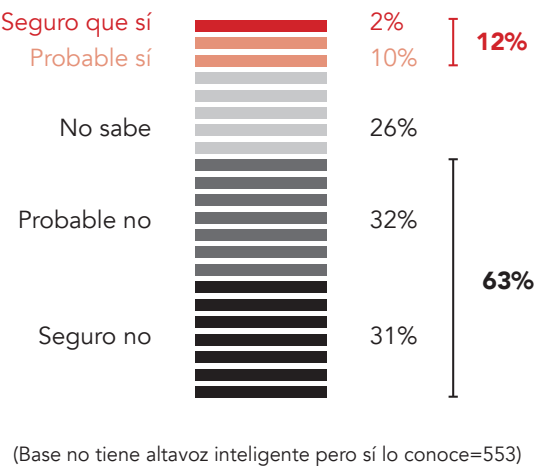
No se aprecia un grado de interés destacable por parte del no usuario actual a la hora de incorporar la tecnología de voz en su rutina a corto plazo: tan sólo el 13% lo considera. En el caso de los altavoces inteligentes se alcanza una intención de compra del 12% motivado principalmente por una baja percepción de utilidad, la ausencia de tecnología domótica en el hogar y el precio.

No se aprecia gran interés a futuro entre los actuales no usuarios de tecnología por voz

Intención utilizar un asistente en los próximos meses



Intención de compra de un altavoz inteligente



< > Proyección de la **aplicación de la tecnología de voz** entre los no usuarios

Referente al uso que proyectarían los no usuarios en el caso de incorporar la tecnología de voz a sus rutinas, encontramos que aunque son muy variados, los usos principales apenas difieren de los usos actuales declarados: tareas simples en las que predomina la comodidad y la capacidad para realizar multitareas.

Se proyectan usos muy variados, que **coinciden con los que llevan a la práctica los actuales usuarios**

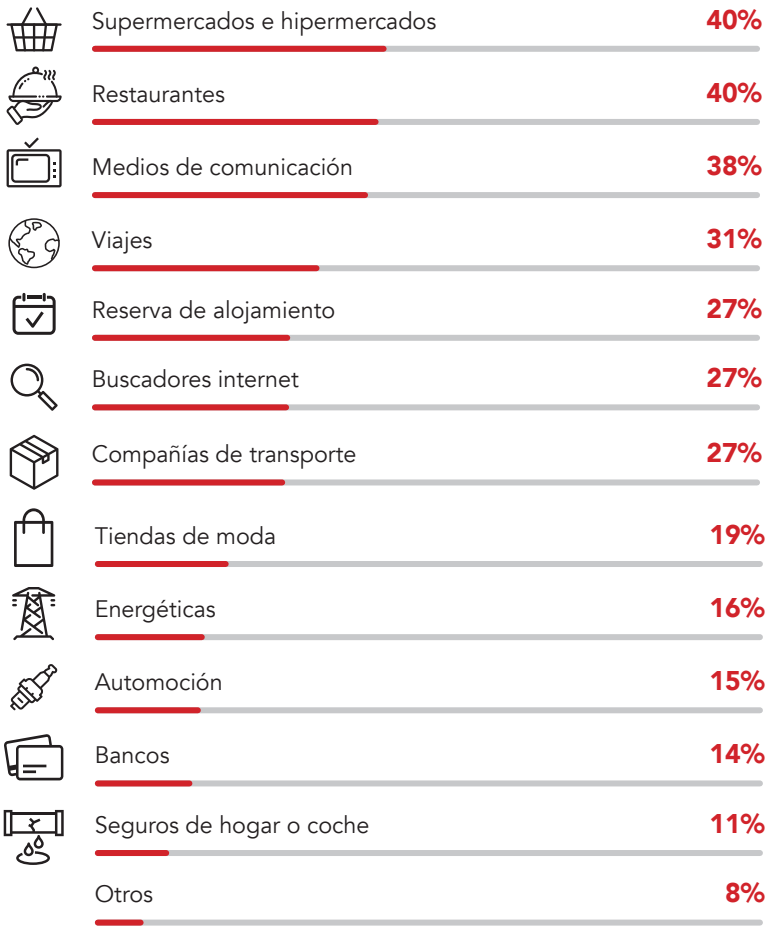


(Base no utiliza tecnología por voz=476)

< > Consideración futura
de la **aplicación de uso por sectores**

Supermercados, restaurantes y medios de comunicación son los sectores donde se ve un mayor potencial de uso futuro para la tecnologías de voz.

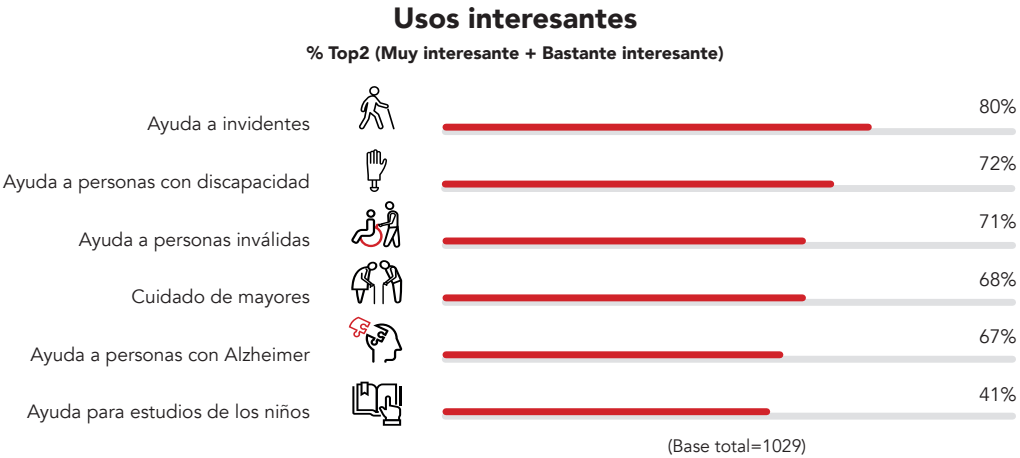
Gran consumo y medios de comunicación son los sectores en los que el usuario proyecta más crecimiento en la aplicación de la tecnología de voz



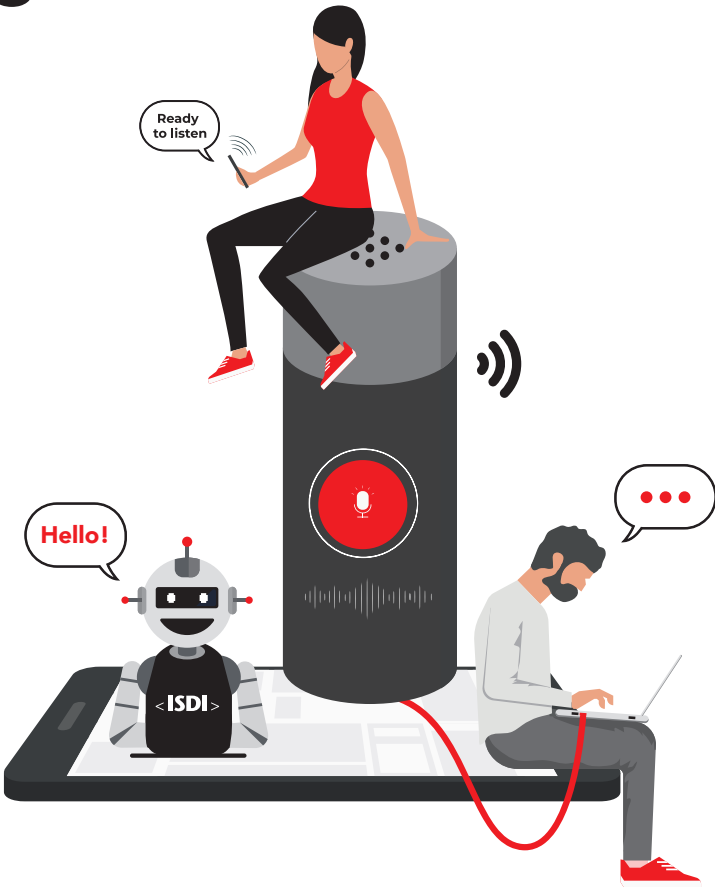
< > **Potencialidad** de la tecnología de la voz en el **ámbito social**

Cabe destacar que, de una forma espontánea, se observa de forma concreta una proyección del potencial de uso de la tecnología de voz en la mejora de la vida de personas pertenecientes a colectivos vulnerables. El segmento de invidentes, personas con discapacidad y personas de movilidad reducida ocupan los tres primeros puestos en la pregunta sobre usos interesantes a futuro de la tecnología de voz.

Buenísima aceptación de todos lo usos relacionados con la **ayuda a personas vulnerables**



CONCLUSIONES

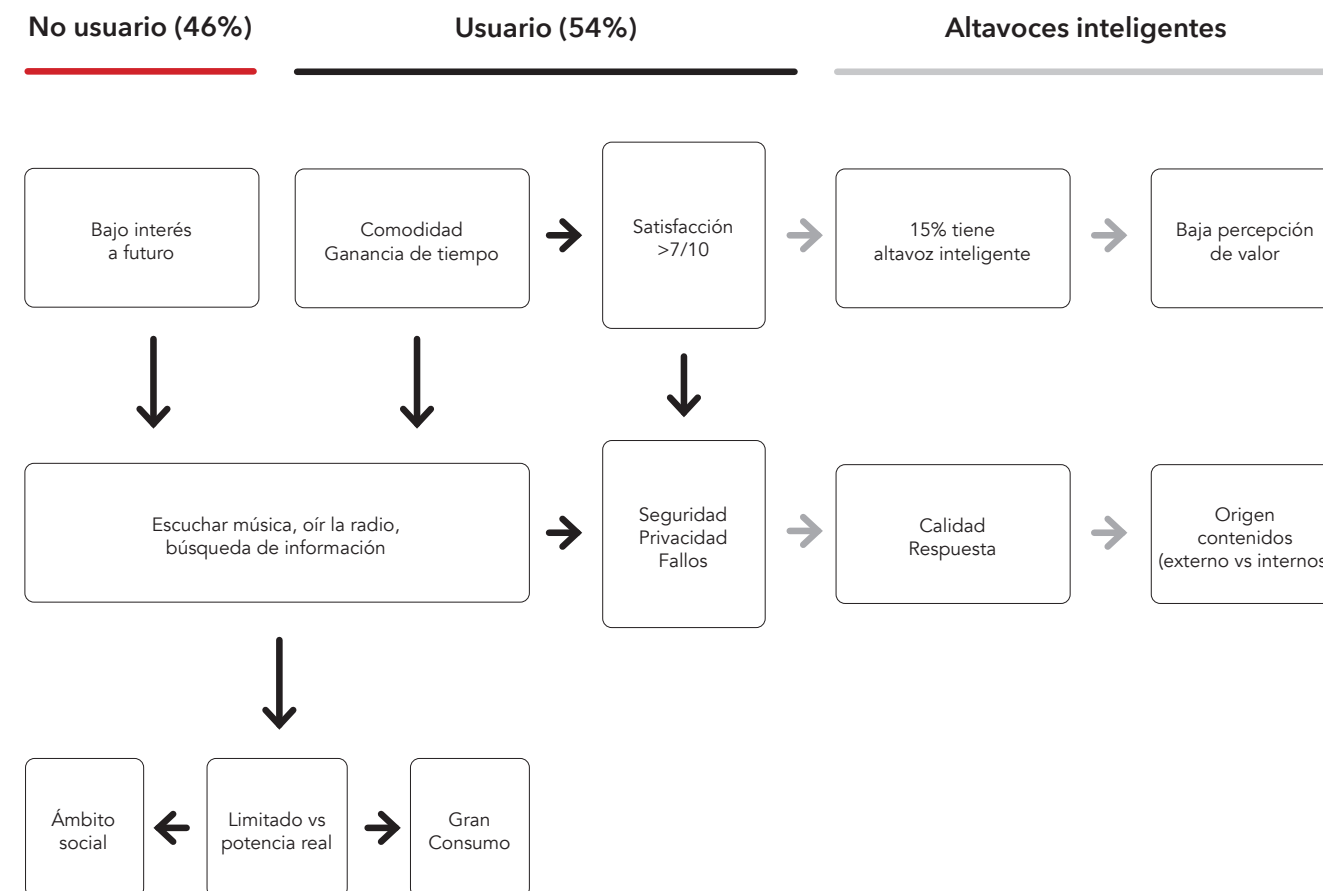


< ISDI > Conclusiones principales

1. La tecnología de voz en España es utilizada ya por más del 50% de la población, sin embargo, los usos son muy limitados comparados con el potencial que ofrece este tipo de tecnología.

Si bien es cierto que de los frenos que declara el usuario para un uso más extensivo, seguridad y privacidad son frecuentes cuando nos referimos al uso de la tecnología en general, los fallos en usabilidad y fiabilidad reflejan las limitaciones específicas que actualmente percibe el usuario en el uso de la tecnología de voz.

En el caso de los altavoces inteligentes, el usuario aún no percibe una propuesta de valor de suficiente utilidad: un 50% identifica como uno de los principales frenos a la compra de dispositivos de altavoces inteligentes el hecho de no apreciar en ellos ninguna utilidad diferencial con respecto a otros dispositivos que ofrecen esta tecnología de forma integrada.



< > Oportunidades para la mejora de respuesta

2. Desde la perspectiva de la tecnología se comprueba que la calidad de la respuesta es mayor en aquellos contenidos “nativos” de la tecnología de Voz, presentando problemas en aquellos casos que los contenidos se generaron desde otros orígenes. En este sentido, es particularmente relevante el posicionamiento estratégico que han tomado cada uno de los players. Mientras que Amazon apuesta por derivar todas las tareas no soportadas de forma nativa a assets propios dentro de su ecosistema (ecommerce Amazon, IMDB, skills...), Google se apoya en la fortaleza de su motor de búsqueda, haciendo que el posicionamiento en sitios especializados, verticales y blogs de la web gráfica sean clave para las marcas y organizaciones a la hora de obtener visibilidad en las respuestas del asistente.

3. Desde el territorio de los contenidos se observa que hay verticales que están dotadas de un tipo de estructura e información que las hacen más fáciles de interpretar por los dispositivos de voz. Es el caso de, por un lado, sectores como entretenimiento que disponen de multitud de datos públicos y estructurados, y por otro sectores como el financiero, en el que gran parte de las consultas más frecuentes son de tipo “glosario” o “búsqueda de términos”, (casos que los asistentes resuelven particularmente bien). En el caso específico de Google, las empresas tienen mas posibilidades de tomar acciones ya que la dependencia es casi total salvo en los verticales que cuentan con mejor estructuración de datos en buscadores (Viajes y Entretenimiento - Cine).



www.isdi.education

Tel: 900 814 144

Síguenos en:

