

# ¡Háblame!

## PROS Y CONTRAS DE LOS DISPOSITIVOS CONTROLADOS POR VOZ PARA NIÑOS, NIÑAS Y ADULTOS

David Kleeman

**El autor discute las ventajas y desventajas de la voz como medio de navegación de dispositivos inteligentes, teléfonos, juegos, juguetes, aplicaciones y experiencias del metaverso.**

Aún no estamos seguros de qué era “Sally”, pero estaba allí en nuestra lista de compras generada por Alexa. Claramente, alguno de nosotros ha usado la activación por voz para agregar algún ítem que necesitábamos del supermercado, pero “ella” había escuchado mal. Fue un buen recordatorio de que los altavoces inteligentes son solo tan inteligentes como su capacidad de

entender lo que queremos, y que puede ser un desafío con las voces de los niños y las niñas más pequeños/as. Los resultados pueden ser... alarmantes.

### EL ACCESO A LOS ALTAVOCES SE ELEVA ENTRE LOS NIÑOS Y LAS NIÑAS

La presencia de dispositivos con altavoces activados continúa creciendo en los hogares alrededor del mundo. A través de los 16 países encuestados por Dubit Trends<sup>1</sup> a partir de octubre de 2021, 46% de los niños y las niñas de 2 a 15 años de edad tenían acceso a altavoces intelligen-

tes, siendo que alrededor de un tercio tenían su propio altavoz inteligente. En los Estados Unidos, el acceso y la propiedad es casi 56%, alrededor del doble que en 2018.

Probablemente el número actual es mayor, debido a la cantidad de dispositivos hoy en día tienen control de voz, incluyendo a los teléfonos smart.

Ahora más del 50% de los niños y niñas de Estados Unidos de 2 a 15 años de edad tienen su propio teléfono smart, con alrededor del 43% de propietarios a través de otros países encuestados.

59% de los niños y niñas en Estados Unidos dicen que generalmente usan la búsqueda por voz para encontrar

lo que quieren mirar, jugar o escuchar; esto es arriba del 50% desde octubre de 2020. Un único altavoz inteligente es el dispositivo más usado para hacerlo, con 35%, seguido por el comando de voz del teléfono smart con 29% y luego el control remoto de la televisión 27%. Escuchar música es la aplicación principal de un altavoz inteligente, seguido por historias y podcasts y la búsqueda de información.



Ill. 1: Los dispositivos que se activan por voz tienen una presencia en las casas que continúa creciendo en todo el mundo

### EL NIÑO, LA NIÑA Y LA ACTIVACIÓN POR VOZ: VENTAJAS Y DESVENTAJAS

Claramente, el control de voz tiene algunas ventajas para los niños y las niñas, particularmente para aquellos/as que están pre alfabetizados y por lo tanto tienen dificultades para usar la búsqueda basada en el texto o juegos, o para aquellos/as que todavía están dominando las habilidades motoras finas de pantallas táctiles. Al mismo tiempo, existen limitaciones al habla como interfaz.

Los altavoces inteligentes tienden a dar una devolución de menor ayuda que los medios basados en pantallas, cuando la pregunta de un usuario joven no alcanza la marca. Generalmente, el dispositivo va a responder que no puede encontrar información que coincida con la solicitud, o va a devolver una solución basada en lo que “pensó” que escuchó. Con Alexa en casa, p. Ej., una solicitud sobre un particular servicio nuevo de streaming frecuentemente da como resultado una lista de canciones de una banda musical con un nombre vagamente similar.

Cuando se agrega una pantalla (con texto o con control por voz), se agregan algunas pistas adicionales sobre el contexto. Ahora le digo a Alexa, “Peppa Pig,” y responde, “Parece que estás preguntando sobre Peppa Pig; ¿te gustaría escuchar algo sobre eso?” Una respuesta afirmativa da como resultado una narración tipo wiki sobre el tema y los actores y actrices de la serie - no demasiado útil para alguien de 3 años de edad. Por el contrario, una búsqueda en Google de, simplemente, “Peppa Pig” arroja resultados de imágenes (productos y capturas de pantalla), links de videos, links de juegos, sitios web, descripciones y más. Tal vez sea abrumador para niños y niñas pequeños/as, pero al menos se les ofrece múltiples señales desde las cuales lanzar sucesivas aproximaciones. La interpretación del habla por IA

puede ser desafiante incluso para las voces de no niños/a.

Recientemente, un estudio académico grupal entre Estados Unidos e India (Ramesh y otros., 2022; resumido en Simonite, 2022) encontró que el 40% de más de 7.000 videos de YouTube de canales populares para niños y niñas contenían palabras “tabú” en algunas capturas generadas desde el programa de subtítulos automáticos.

Por supuesto, todavía es temprano para las interfaces de voz, y están mejorando a medida que los desarrolladores mejoran la inteligencia artificial. Cuanto mayor es la experiencia en la vida real que los programadores tienen para analizar – un rango más amplio de voces, acentos, palabras clave, sintaxis, frustraciones fuera de los objetivos, etc – la actuación de la IA será mejor, haciendo crecer sus bibliotecas internas masivas para determinar más rápidamente lo que un niño o una niña quiere.

### LA VOZ EN EL METAVERSO

Mientras que nos movemos hacia el metaverso (Kleeman, 2021), el control de voz será extremadamente interesante. En los mundos inmersivos, mientras que puede hacerse a través de controladores manuales, tener que usar una interfaz de texto para actividades más complejas o para moverse de un lugar a otro rompería la ilusión y „suspensión de la incredulidad.”

Estos mundos serán diseñados para juegos sociales, y por lo tanto vamos a querer comunicarnos con otros a través del habla.

En el metaverso, sin embargo, tendremos que trabajar más y no solamente nuestras palabras para generar el contexto que actualmente está faltando. En la realidad virtual, por ejemplo, la IA que se activa por el comando de voz podría complementar lo que oye con registro visual, la posición del cuerpo, y la inferencia del entorno. Meta recientemente demostró una tecnología

llamada “BuilderBot” que coloca ítems en un espacio virtual basados en palabras que dicen los usuarios así como también pistas del contexto del espacio virtual (Ghaffary, 2022).

Queremos estar seguros de que el metaverso es un espacio inclusivo y accesible, y que las voces van a tener un rol fuerte para llevar a cabo una experiencia satisfactoria para los usuarios con discapacidades y sin discapacidades. La voz puede ser el medio de navegación para aquellos/as que tienen una capacidad motora restringida y les impide usar controles manuales o al mover sus cuerpos del mundo real en espacios virtuales.

### ¡HÁBLAME!

¿Recuerdan cuando las tabletas fueron lanzadas al mercado, y los niños y niñas pensaron que el mundo entero estaba hecho de pantallas interactivas?

YouTube estaba lleno de videos protagonizados por niños y niñas tratando de deslizar y tocar libros y revistas como si fueran pantallas.

Tal vez en el futuro cercano, los niños y las niñas pequeños/as van a pensar que pueden controlar todo a través del comando de voz. Y, con más y más dispositivos controlados por voz – los altavoces inteligentes, teléfonos smart, sistemas hogareños inteligentes, controles inteligentes y juguetes, (también, supongo, inteligentes) – ¿van a estar equivocados? ■

## NOTA

<sup>1</sup> Dubit Trends se realiza en abril y octubre en los Estados Unidos (2.000 niños y niñas/padres y madres de 2 a 18 años) y en el Reino Unido (1.300 niños y niñas/padres y madres de 2 a 18 años), con 18 países más encuestados rotativamente (1.000 niños y niñas/familias de 2 a 16 años). La encuesta es en línea, representativa y con la identificación étnica, de ingresos, edad de los padres y de las madres, y más. El estudio está distribuido a los asociados globalmente, y se usa como fundamento de la investigación de Dubit y como estrategia de consultoría así como también para el desarrollo y la producción de juegos, aplicaciones y experiencias de metaverso.

## REFERENCIAS

Ghaffary, Shirin (2022). Mark Zuckerberg wants to build a voice assistant that blows Alexa and Siri away. *Vox recode* (February 23, 2022). Available at: <https://www.vox.com/recode/22948097/mark-zuckerberg-voice-assistant-metaverse-ai-announcem> [28.4.22]

Kleeman, David (2021). What is the “metaverse” & why does it matter? Available at: <https://staging.kidscreen.com/2021/09/14/what-is-the-metaverse-why-does-it-matter/> [28.4.22]

Ramesh, Krithika, KhudaBuksh, Ashiqur & Kumar, Sumeet (2022). “Beach” to “bitch”: Inadvertent unsafe transcription of kids’ content on YouTube. Preprint (February 17, 2022). Available at: [https://github.com/sumeetkr/UnsafeTranscriptionofKidsContent/blob/main/YouTube\\_Transcription\\_Final.pdf](https://github.com/sumeetkr/UnsafeTranscriptionofKidsContent/blob/main/YouTube_Transcription_Final.pdf) [28.4.22]

Simonite, Tom (2022). YouTube’s captions insert explicit language in kids’ videos. *Wired* (February 24, 2022). Available at: <https://www.wired.com/story/youtubes-captions-insert-explicit-language-kids-videos/> [28.4.22]

## EL AUTOR

David Kleeman es vicepresidente de Global Trends for Dubit, una consultora de estrategia e investigación localizada en Leeds, Inglaterra.



Traducción  
Vicky Romano