



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Grado en Marketing e Investigación de Mercados

Adopción de la inteligencia artificial en España y Taiwán: El caso de la tecnología de voz

Presentado por:

Meng Cheng Tsai

Tutelado por:

Carmen Antón Martín

Rebeca San José Cabezudo

Valladolid, 20 de septiembre de 2020

Índice

1. Introducción.....	1
2. Marco teórico.....	4
2.1. La tecnología de voz	4
2.1.1. AI como la base de la tecnología de voz	5
2.1.2. Limitaciones de la tecnología de voz: VUI vs GUI	5
2.1.3. Futuro de la tecnología de voz	7
2.2. Teorías de adopción de una nueva tecnología	8
2.2.1. Modelo de aceptación de la tecnología y sus derivados	8
2.2.2. Uses and Gratification theory U>	9
2.3. Propuesta de un modelo explicativo y desarrollo de los conceptos.....	10
2.3.2. Beneficio hedónico	10
2.3.3. Beneficio simbólico	10
2.3.4. Beneficio social	11
2.3.5. Riesgo percibido sobre la privacidad	11
2.3.6. Resistencia al cambio	11
2.3.7. Compatibilidad con el estilo de vida	12
3. Metodología.....	13
3.1. Plan de muestreo y selección de la muestra	13
3.2. Medida de las variables	13
3.3. Composición de la muestra.....	16
4. Análisis de resultados.....	17
4.1. Análisis factorial para construir las medidas de las variables.....	17
4.2. Regresión lineal: Análisis de los determinantes de la intención de utilizar la tecnología de voz	18
4.3. Regresión logística: análisis de los determinantes de la previsión de cambio en el uso.....	25
5. Conclusiones.....	30
6. Sugerencias para investigaciones futuras.....	33
7. Bibliografía	34
8. Anexos	37

Índice de los gráficos, tablas

Gráfico 2.1: Proceso de interacción con la tecnología de voz.....	4
Tabla 2.1: Escenarios ideales para la aplicación de la tecnología de voz	7
Gráfico 2.1: Modelo propuesto	12
Tabla 3.1: Variables y estadísticos descriptivos	15
Gráfico 3.1: Composición de la muestra (I)	16
Gráfico 3.2: Composición de la muestra (II)	17
Tabla 4.1: Resultados de la regresión lineal para la intención de uso.....	21
Gráfico 4.1: Ilustración de la interpretación de la previsión de cambio del uso	25
Tabla 4.2: Resultados de la regresión logística para la previsión de cambio en el uso ..	27
Tabla 5.1: Resumen de las diferencias entre España y Taiwan	32

Resumen

La tecnología de voz parece ser una tendencia de futuro según las predicciones, pero se conoce poco sobre los motivos para utilizarla atendiendo a cada una de sus funciones. En este trabajo se agrupan las funciones realizables con la tecnología de voz en cuatro niveles de uso: dar órdenes, buscar informaciones, dejarse aconsejar y entretenerse, y se construye un modelo con el fin de explicar la intención de utilizar la tecnología de voz en el futuro. La concepción de los factores explicativos de dicho uso está basada en las teorías de adopción de innovaciones y de usos y gratificaciones, así como en las propuestas del investigador. Para comprobar las hipótesis del trabajo, se ha utilizado una muestra compuesta por 244 individuos (70 españoles y 174 taiwaneses). Como resultados principales, se ha descubierto la importancia de la compatibilidad con el estilo de vida, la utilidad y el beneficio hedónico en las intenciones de uso. Por el otro lado, se nota que los motivos de uso son diferentes para cada función, lo que conduce a una posible división de las 4 funciones en dos bloques, uno compuesto por las funciones de dar órdenes y buscar informaciones, y otro por las de dejarse aconsejar y entretenerse. También se sugiere que la percepción de beneficio social hace que los individuos busquen nuevas posibilidades en el uso de la tecnología de voz, más allá de las cuatro funciones identificadas inicialmente. Por último, se observa que los taiwaneses tienen mayor intención de utilizar la tecnología de voz que los españoles, además, el factor “escenarios de pereza” tiene un efecto impulsor especialmente relevante para este público.

Palabras claves: Tecnología de voz, Adopción de la nueva tecnología, Inteligencia artificial, Compatibilidad con el estilo de vida

Código JEL: O39

Abstract

Voice technology seems to be a trend of the future according to the predictions, but little is known about the reasons for using it according to each of its functions. In this research, the functions that can be performed with voice technology are grouped into four levels of use: giving orders, seeking information, being advised and entertained, and a model is built in order to explain the intention to use voice technology in the future. The conception of the explanatory factors of said use is based on the theories of adoption of innovations and of uses and gratifications, as well as on the researcher's proposals. To test the hypotheses of the work, a sample composed of 244 individuals (70 Spanish and 174 Taiwanese) has been used. As main results, the importance of compatibility with lifestyle, usefulness and hedonic benefit in use intentions has been discovered. On the other hand, it is noted that the reasons for use are different for each function, which leads to a possible division of the 4 functions into 2 blocks, one composed of the functions of giving orders and searching for information, and the other for those of be advised and entertained. It is also suggested that the perception of social benefit makes individuals look for new possibilities in the use of voice technology, beyond the four functions initially identified. Finally, it is observed that the Taiwanese have a greater intention to use voice technology than the Spanish, in addition, the factor "laziness scenarios" has an especially relevant driving effect for this audience.

Key words: Voice technology, Adoption of new technology, Artificial intelligence, Lifestyle compatibility

Code JEL: O39

1. INTRODUCCIÓN

La tecnología de voz es una rama importante de la revolución de la inteligencia artificial, y hay consenso de que ha venido para quedarse (Galeano, 2019). Según Ovum, una empresa de investigación, en 2021 habrá tantos asistentes activados por voz como personas en el planeta, y según el informe elaborado por Data Bridge Market Research, se espera que el mercado mundial de la voz alcance los 677.021 millones de dólares en 2025, lo que supondría un crecimiento anual del 25,7% para el periodo 2018-2025 (Carreras, 2019).

La difusión de la tecnología de voz en el público general se evidencia con el creciente uso de dos de sus aplicaciones, el altavoz inteligente y la búsqueda por voz. El altavoz inteligente es el dispositivo con mayor frecuencia de uso para comunicarse con el asistente de voz. Según la investigación previa, el 44% de los poseedores lo usa varias veces al día (ReasonWhy, 2019). En el estudio realizado por Canalys y Strategy Analytics, en el último trimestre de 2019 los envíos de altavoces inteligentes llegaron a los 55,7 millones de unidades en todo el mundo, lo que supone un aumento del 44,7% con respecto al mismo trimestre del año anterior. A pesar de que el Covid-19 paró temporalmente a su producción, se prevé que 2020 seguirá siendo el año en el que se logrará el nuevo récord en cuanto a envíos y ventas de los altavoces inteligentes (MoneyDJ, 2020). Por otro lado, la búsqueda por voz también está experimentando un crecimiento increíble. Según Comscore, en 2019 el 20% de las búsquedas en Android eran a través de la voz y se estima que ese porcentaje ascienda al 50% en 2020 y que el 30% de todas las búsquedas se realizarán sin pantallas (Carreras, 2019).

En el ámbito español, el 45,7% de los navegantes habituales utilizan la tecnología de voz, y de ellos, 22% son considerados como usuarios activos con una frecuencia de uso igual o superior a "varias veces a la semana". El uso de asistente de voz en 2019 ha incrementado un 2,9% con respecto al año anterior y la función más utilizada es realizar búsquedas o preguntas en general. Sin embargo, la difusión de los altavoces inteligentes parece ser más lenta de lo esperado. Solamente el 7% de los navegantes poseen este tipo de dispositivo y únicamente el 6,4% manifiestan su intención de comprarlo. También cabe mencionar que el 40% de los poseedores del dispositivo lo utilizan para hacer control de domótica, siendo el control de la luz la opción favorita para el 76%

(AIMC, 2019; 2020).

A pesar de las predicciones optimistas sobre el crecimiento del uso de la tecnología, los análisis anteriores advierten que los usuarios prefieren utilizarla en lugares privados, ya que en caso contrario la presencia de otros generaría “presiones” (GrowthHK, 2020). En consecuencia, es más difícil percibir el aumento de uso porque no es fácilmente observable. El investigador del presente trabajo no es usuario habitual de este tipo de tecnología y tampoco percibe su difusión progresiva. Por tanto, tiene curiosidad sobre las afirmaciones que dicen que la tecnología de voz será una tendencia del futuro y ha querido hacer una investigación sobre el tema.

Objetivos de la investigación

Muchos estudios existentes meramente describen los comportamientos de uso de la inteligencia artificial (*Artificial Intelligence*, AI) por voz, tales como la frecuencia de uso, dispositivos a través de los que se usa o qué funciones se realizan con ella, etc. (AIMC, 2019; 2020, Galeano, 2019). Otros estudios se centran más en explicar el motivo de utilizar la tecnología a nivel general, sin entrar a diferenciar entre sus distintas funciones (McLean y Osei-Frimpong, 2019). El presente trabajo trata de avanzar sobre lo anterior y se plantea para conocer los motivos por los que una persona tiene mayor o menor intención de utilizar la tecnología de la voz para distintos usos. Como existen muchas funciones, para simplificar la investigación, se han agrupado todas las acciones realizables con la tecnología de la voz en cuatro grandes bloques o niveles de uso:

- Dar órdenes: controlar los dispositivos o aplicaciones (sustituye los clics físicos por la voz). En este nivel el usuario necesita cumplir ciertas tareas y ya tiene previsto cómo serán los resultados dados por la máquina. Algunos ejemplos son encender/apagar la luz, control de electrodomésticos, poner alarma/música, hacer llamadas, apuntar eventos en el calendario o escribir textos a través de la voz, entre otros.
- Buscar informaciones: acceder a información que el usuario no tiene. En este nivel el usuario necesita saber algo que desconoce en ese momento. Así puede buscar información de tiempo/clima/localización, resultados de

los partidos, búsqueda de una canción/película, búsquedas generales sobre cualquier tema, etc.

- Dejarse aconsejar: ceder al asistente de voz el poder de decisión, es decir, que seleccione por el usuario entre varias alternativas. En este nivel el usuario necesita tomar decisiones que le cuesta tomar, entonces delega la decisión a la máquina. Por ejemplo, puede pedirle que recomiende un restaurante, que elija una película, que seleccione la ropa para ponerse, etc.
- Entretenerse: pasar un buen rato mediante la interacción con el asistente de voz. Este nivel de interacción es más particular y a diferencia de los tres anteriores no tiene tanto valor utilitario. Algunos ejemplos son la conversación con el asistente, los juegos de voz o pedirle que imite la voz de personajes famosos o de animales.

Por otro lado, también se pretende explicar el paso desde el uso actual a la intención de utilizar la tecnología de voz en el futuro. Se analizan las intenciones de uso en los cuatro niveles en comparación con el uso actual. ¿Cómo es la intención de los individuos que actualmente están utilizando mucho la tecnología de voz, están muy seguros o no sobre su uso continuado? ¿Y para los que la están utilizando poco? En otras palabras, si los individuos prevén un cambio de su comportamiento en el uso de la tecnología en el futuro.

Por último, gracias a la condición que tiene el investigador, ha sido posible recoger muestras de dos países diferentes, España y Taiwán. Por tanto, otra aportación importante de este trabajo es la comparación entre públicos de estos dos países, es decir, si hay diferencias entre estas dos culturas de cara a las intenciones de utilizar la tecnología de voz.

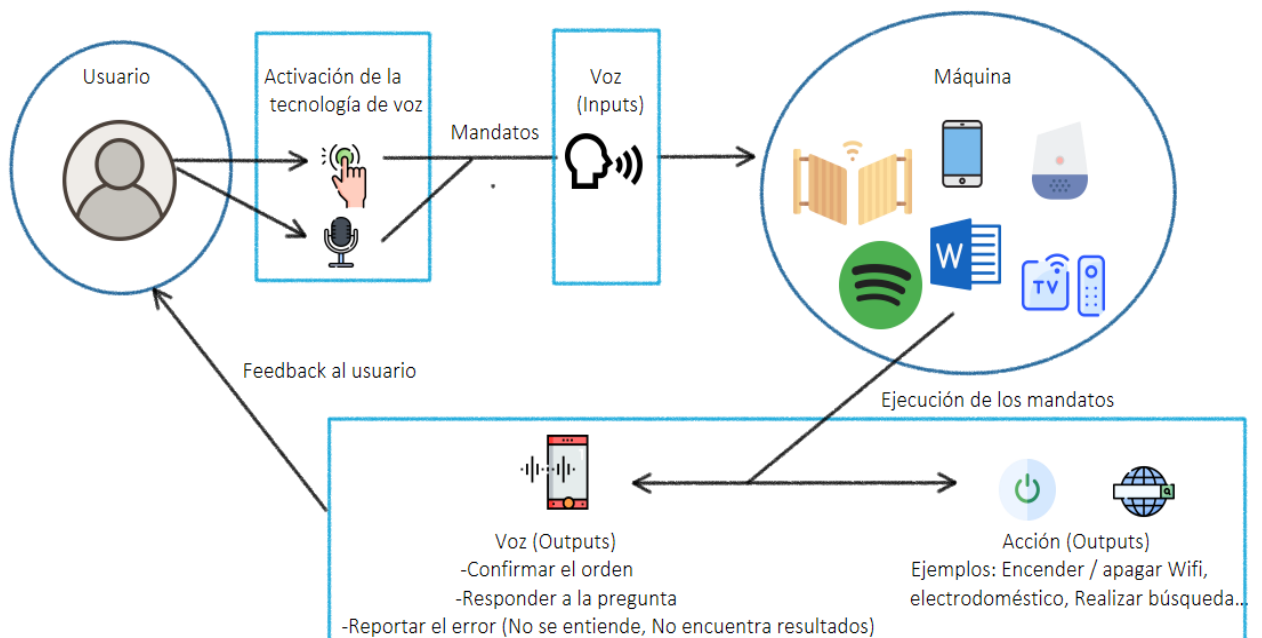
En resumen, el presente estudio tiene tres objetivos principales:

1. Explicar las intenciones de utilizar la tecnología de voz en cuatro niveles de uso, a partir de una serie de variables explicativas previamente concebidas.
2. Analizar las previsiones sobre el cambio del comportamiento de uso de la tecnología de voz, a partir de dichas variables.
3. Comparar la intención de uso de la tecnología de voz entre los dos países del análisis: España y Taiwán.

2. MARCO TEÓRICO

2.1. La tecnología de voz

Se refiere a la tecnología que permite al usuario interactuar con la máquina a través de la voz (Alan, 2019). También es conocida como control por voz, reconocimiento de voz o VUI (*Voice User Interface*). Para utilizar la tecnología, el usuario tiene que activar el “oído” de la máquina, o bien pulsándole (por ejemplo, hacer clic en el icono de grabación o activarlo con botones físicos), o bien llamándole (como Decir Hey Siri o Alexa). A continuación, ya se puede empezar a hablar a la máquina, que entonces procesa los mensajes verbales y ejecuta las acciones correspondientes. En la siguiente imagen se muestra gráficamente el proceso de interacción con la tecnología de voz (Gráfico 2.1.).



Icon made by Freepik, Chanut, Srip, FlatIcons, Pixel perfect, Vitaly Gorbachev from www.flaticon.com

Gráfico 2.1: Proceso de interacción con la tecnología de voz

La tecnología de voz está presente en muchos elementos de nuestro entorno. Se puede insertar en dispositivos o aplicaciones (por ejemplo, los asistentes de voz que están en los teléfonos móviles, los altavoces inteligentes o los coches); también engloban los programas de dictado por voz en los software de documentación (por ejemplo, el que incluye Word), en las plataformas de información (como Google, Youtube, etc.) o en las aplicaciones de mensajería

(como Whatsapp), entre otros.

El asistente personal es la figura que mejor ilustra el papel que desempeña la tecnología de voz. Así, tenemos ejemplos tan famosos de asistentes de voz como Siri y Alexa. Está claro que otras tecnologías también asisten a su usuario, pero ninguna de ellas sabe escuchar y hablar como hace la tecnología de voz, lo cual crea muchas posibilidades en cuanto a sus aplicaciones.

2.1.1. AI como la base de la tecnología de voz

La AI (*Artificial Intelligence*) representa una nueva revolución tecnológica. Podemos definirla como un sistema (ordenador) que según lo que percibe, reacciona de forma lógica para lograr resultados óptimos (Lee y Wang, 2017). Los dos rasgos de la AI más destacados son que el sistema tiene un comportamiento similar al hombre y que posee una capacidad de auto-aprender desde el Big Data.

Para la tecnología de voz es imprescindible la AI porque la máquina tiene que aprender a escuchar, entender y responder ante el lenguaje humano. Para ello, hace falta entrenarle con grandes cantidades de datos auditivos (Big Data). En el ámbito de la voz, hay tres técnicas de AI que son fundamentales y cada una actúa sobre una fase de la interacción (Alan,2019):

- Reconocimiento de voz en texto (ASR, *Automatic Speech Recognition*): proceso en el que la máquina traduce los mensajes verbales en textos, parecido a una transcripción.
- Reconocimiento del significado (NER, *Name Entity Recognition*): proceso por el que las máquinas descifran los significados reales de los textos obtenidos. Es el reto más difícil de superar debido a la complicación del lenguaje humano (múltiples formas de expresar una misma cosa, pronombre, doble sentido, connotación, etc.). El objetivo es que la máquina pueda tener en cuenta el contexto.
- Síntesis de voz: proceso por el que la máquina lee las respuestas textuales generadas por el sistema.

2.1.2. Limitaciones de la tecnología de voz: VUI vs GUI

Tradicionalmente los individuos interactúan con la máquina mediante los ojos

y las manos, viendo las pantallas y haciendo clics, a través de GUI (*Graphical User Interface*). La tecnología de voz ofrece una nueva forma de interactuar, mediante la oreja y la boca, conocida como VUI (*Voice User Interface*). No es adecuado decir que una es mejor que otra, sino que una puede ser más conveniente que otra según la ocasión de uso. Entonces, ¿en qué ocasiones la VUI puede aportar diferencias que son valoradas por los usuarios? Para responder a esta pregunta, hay que entender la esencia de ambos tipos de interacción (GrowthHK, 2020):

1. La GUI es más eficiente en los *outputs* de información, más concretamente:
 - Es más rápido captar las informaciones visuales que las auditivas (leer es más rápido que escuchar, una imagen explica más que mil palabras, etc.).
 - La VUI requiere la atención total del usuario, mientras que en el caso de GUI no pasa nada si el usuario se distrae por un momento.
2. La GUI es más conveniente para una interacción que requiere varios pasos (como cuando se reserva un billete de transporte, por ejemplo). Con GUI, el usuario tiene tiempo de pensar, ver las opciones y responder, mientras que VUI obliga respuestas inmediatas tras las preguntas verbales.
3. La VUI es más rápida, pero menos precisa que la GUI en los *inputs*:
 - Hablar es más rápido que teclear.
 - La VUI permite saltar los pasos en las búsquedas de contenidos familiares (por ejemplo, para buscar una canción en la GUI hay que abrir la aplicación, localizar el buscador interno, teclear y buscar la canción dentro de la lista de resultados, en cambio, en la VUI basta con decir el nombre de la canción).
 - La VUI es menos precisa porque puede haber fallos en ASR o NLP (*Natural Language Processing*).
4. La VUI requiere menos esfuerzo de aprendizaje que la GUI. Un ejemplo ilustrativo es cuando se ve a los mayores aprendiendo a utilizar los smartphone (sistema dominado por GUI), suelen tener muchos problemas (no saben a cuál tiene que pulsar, no recuerdan cómo llegar a las opciones, etc.). Hablar es la forma de interactuar más intuitiva y primitiva

para el hombre.

En conclusión, la VUI es menos eficiente en la emisión de informaciones y se debe centrar en aquellas tareas que no contienen demasiadas informaciones. También presenta limitaciones para aquellas tareas en las que el usuario necesita tiempo para decidir. Por contra es más intuitiva de utilizar y es más rápida en introducir los *inputs*. A parte de esto, hay que tener en cuenta la restricción de lugar de uso, que no es posible utilizarlo en los lugares demasiados ruidosos o que requieren silencios (oficina, biblioteca, etc.).

Según todo lo anterior, se obtienen las ocasiones de uso ideales para la tecnología de voz, que se muestran en la Tabla 2.1.

Tareas	Escenarios	Lugares
<ul style="list-style-type: none"> -Tareas familiares para el usuario que pueden ser cumplidas por una sola frase - Tareas simples y rutinarias - Introducir textos - Búsquedas de información 	<p>Escenarios en que NO conviene utilizar la GUI:</p> <ul style="list-style-type: none"> - Multitarea (Ej: conducir, cocinar, etc.) - Cuando no hay luz suficiente - Para públicos específicos (niños, mayores o discapacitados) - En los escenarios de pereza (Ej: pereza de teclear, de moverse para acercarse al dispositivo físicamente cuando uno está en la cama/sofá, etc.) 	<p>Espacio donde no se requiere silencio ni hay demasiados ruidos (Ej: casa, coche, calle, etc.)</p>

Tabla 2.1: Escenarios ideales para la aplicación de la tecnología de voz

2.1.3. Futuro de la tecnología de voz

La combinación de diferentes tipos de interfaces parece ser la tendencia del futuro (GrowthHK, 2020). En el caso de la VUI, dadas sus limitaciones, en algunos casos sería un complemento perfecto de las interfaces visuales (por ejemplo, asistente de voz para los dispositivos de realidad virtual, *Virtual Reality* -VR- y realidad aumentada, *Augmented Reality* -AR-). En otros casos, sería la tecnología protagonista, pero aprovechando la ayuda de las interfaces visuales para ofrecer un servicio más completo (por ejemplo, en los altavoces inteligentes con pantallas).

Por el otro lado, ocupará un peso importante en el campo de Internet de las cosas (IoT, *Internet of Things*), que son los objetos conectables a Wifi. Según la tendencia, al bajar los costes será asequible que los objetos lleven micrófono, altavoz y módulo de Wifi, por lo que la VUI puede tener un lugar importante en los objetos conectables sin pantallas (GrowthHK, 2020).

En el campo del marketing, la tecnología de voz es considerada como una interfaz de usuario y por ello influye directamente en la experiencia de uso. Esa experiencia de uso está relacionada con la satisfacción del usuario de cara al producto (por ejemplo, en una aplicación, página web o producto electrónico que la incorpore). Por tanto, el marketing debe intervenir en el diseño de la interfaz de voz para que sea útil y fácil de utilizar (MartinEdic, 2019).

Por otro lado, la voz ha abierto nuevo mundo en el ambiente digital. La búsqueda por voz ha creado una nueva batalla en el SEO. La forma de buscar por voz es diferente que teclear, cambian las palabras claves. Las marcas que mejor las definen tienen más posibilidad de ocupar una posición ventajosa en los resultados de las búsquedas por voz. Además, las marcas deben considerar la opción de crear su aplicación de voz. Las grandes empresas que poseen los asistentes de voz permiten a los terceros crear aplicaciones de voz (por ejemplo, *skills* de Alexa), lo que implica una nueva posibilidad de conectarse con los consumidores (Wu, 2019).

2.2. Teorías de adopción de una nueva tecnología

El objetivo del trabajo es averiguar las variables que influyen en las intenciones de utilizar la tecnología de voz. Para ello, es necesario revisar los modelos teóricos existentes para explicar la adopción de las tecnologías.

2.2.1. Modelo de aceptación de la tecnología y sus derivados

El modelo TAM (*Technology Acceptance Model*) se basa en dos factores para explicar la decisión de utilizar una nueva tecnología: la utilidad percibida -el usuario cree que sirve de algo- y la facilidad de uso percibida -el usuario cree que le cuesta poco para poder manejarla- (Davis, 1989). El modelo es usado en muchas investigaciones y ha sido comprobado como una herramienta confiable para conocer la aceptación de las tecnologías (Marangunić y Granić, 2014). Pese

a su uso frecuente, el modelo es criticado por dar un peso excesivo a la utilidad percibida en el proceso de adopción, ignorando otros aspectos importantes como las influencias sociales, costes, etc. (Lunceford, 2009). Una crítica más reciente contrastó que el modelo TAM solo valdría para explicar un 40% del uso de un sistema tecnológico (Legris, Ingham y Colletette, 2003).

Más Adelante surgió el modelo UTAUT (*Unified Theory of Acceptance and Use of Technology*) que combina el modelo TAM, sus extensiones y otros modelos de adopción existentes hasta ese momento. Está basado en cuatro bloques de variables: 1) expectativa de rendimiento, 2) expectativa de esfuerzo, 3) influencia social, y 4) condiciones facilitadoras. La influencia de los bloques está moderada por otras cuatro variables: género, edad, experiencia y voluntariedad del uso. En un estudio posterior se demostró que el modelo explicó aproximadamente un 50% del uso real (Venkatesh, Morris, Davis y Davis, 2003). Pese a ello, el modelo UTAUT también recibe críticas. Bagozzi acusa que el modelo utiliza demasiadas variables para pronosticar intenciones y comportamientos (casi 50 variables), lo que lleva al estudio de adopción a un estado de caos (Bagozzi, 2007).

2.2.2. Uses and Gratification theory U>

La teoría U> (*Uses and Gratification Theory*) fue desarrollada para entender por qué un individuo utiliza uno u otro medio de comunicación. Partiendo de los propios usuarios, se postula que el individuo utiliza un medio porque cree que satisface mejor sus necesidades. Entonces, la necesidad del usuario se convierte en el motivo principal en la decisión de uso (Vinney, 2019).

El mismo enfoque es aplicable a la adopción de las nuevas tecnologías. En el caso de la tecnología de voz, podemos pensar que un individuo la utiliza porque cree que de alguna manera satisface sus necesidades.

Ante la presencia de críticas a los modelos de aceptación basados en el TAM, se cree que la perspectiva de necesidad es un punto de vista interesante. En consecuencia, el presente estudio se basa en la teoría de U> incorporando algunos conceptos derivados del modelo TAM.

En una investigación reciente realizada por Rauschnabel, He y Ro (2018), los autores agrupan las necesidades en tres grandes categorías: beneficios utilitarios, beneficios hedónicos y beneficios simbólicos, que serán las variables

que analizaremos en nuestro estudio. Por otra parte, siguiendo la perspectiva de U>, incluimos otra necesidad que el asistente de voz puede cumplir, la necesidad social, en tanto que beneficio social.

2.3. Propuesta de un modelo explicativo y desarrollo de los conceptos

El estudio trata de explicar las intenciones y las previsiones de cambio en el uso de los cuatro niveles de la tecnología de voz. Se han concebido una serie de variables explicativas que influyen en ellas y que se presentan a continuación:

2.3.1. Beneficio utilitario

La tecnología es útil si aporta mejores soluciones a las necesidades del usuario. La tecnología de la voz ofrece soluciones más convenientes en controlar los dispositivos electrónicos y realizar búsquedas de información en los escenarios de multi-tarea o de pereza. Por tanto, en principio la utilidad percibida debe facilitar su adopción. Pero la tecnología todavía presenta defectos (como errores en reconocimiento o de entendimiento) por lo que quizás para mucha gente aún no supera al nivel mínimo de aceptación, y no es percibido como útil. En consecuencia, se propone que:

H1: A mayor beneficio utilitario percibido, mayor intención de uso de la tecnología de voz.

2.3.2. Beneficio hedónico

La interacción con el asistente de voz puede generar sensaciones de disfrute, lo cual facilita la adopción de la tecnología. El hecho de poder hablar con la máquina ya es divertido porque es algo novedoso. Además, las conversaciones con el asistente de voz son entretenidas porque a veces responden con frases humorísticas. Por otro lado, los errores que producen también pueden ser graciosos para los usuarios.

H2: A mayor beneficio percibido, mayor intención de uso de la tecnología de voz.

2.3.3. Beneficio simbólico

El individuo puede utilizar una tecnología con el motivo de crear imagen personal, y más allá de eso ganar estatus social. Así, dado que el reconocimiento

de voz es visto como algo novedoso y avanzado, el uso de la tecnología puede otorgar al usuario ciertos valores simbólicos (ser innovador, creativo, exclusivo, etc.). Por ello,

H3: A mayor beneficio simbólico percibido, mayor intención de uso de la tecnología de voz.

2.3.4. Beneficio social

El asistente de voz tiene una característica humana que le diferencia del resto, la voz. Es la forma de interactuar más primitiva para el hombre, por tanto, el asistente puede dar sensación de compañerismo. Es por ello que se piensa que los asistentes de voz pueden cubrir cierta necesidad social de los usuarios. A diferencia de los amigos o familias, la máquina está disponible en cualquier momento para conversar. Además, a través de la AI puede llegar a entender al usuario incluso mejor que sus amigos y ofrecer respuestas que mejor encajen a su estado del momento.

Entonces, se trata de ver cómo piensan la gente sobre una relación social entre la máquina y el hombre, más allá de un simple asistente, ¿es posible convertirse en un amigo, tutor o amantes? En concreto se propone que:

H4: A mayor beneficio social percibido, mayor intención de uso de la tecnología de voz.

2.3.5. Riesgo percibido sobre la privacidad

El riesgo percibido sobre la privacidad produce efectos negativos en la adopción y uso de los asistentes de voz (Hoy, 2018). Las empresas recolectan los audios con el objetivo de mejorar la tecnología y ofrecer servicios más personalizados. Pero es posible que al usuario no le guste esa sensación de que la empresa le está escuchando y que sepa todo de él. Además, puede preocuparse sobre la seguridad de esos datos almacenados. En consecuencia,

H5: A mayor riesgo percibido sobre la privacidad, menor intención de uso de la tecnología de voz.

2.3.6. Resistencia al cambio

Para utilizar la tecnología hay que hablar con la máquina, y hablar con la

máquina es una forma de interactuar totalmente novedosa. En teoría la voz es más intuitiva y fácil de utilizar para los usuarios, pero las personas se pueden resistir en adoptarla porque no están acostumbradas a conversar con la máquina y pueden sentirse incómodas. Cada innovación requiere adaptación por parte del usuario, la cuestión es ver si la gente está dispuesta a incorporar un nuevo hábito en su vida cotidiana. Así es esperable que:

H6: A mayor resistencia al cambio, menor intención de uso de la tecnología de voz.

2.3.7. Compatibilidad con el estilo de vida

Teniendo en cuenta el análisis de VUI, existen determinados escenarios y tareas ideales para la aplicación de la tecnología de voz. Por ello queremos estudiar la frecuencia con la que los individuos se encuentran en estos escenarios o la frecuencia con la que ejecutan estas tareas. Si la frecuencia es alta, se deduce que la tecnología de la voz será una solución atractiva para ellos y por tanto hará subir más las intenciones de utilizarla.

H7: A mayor compatibilidad con el estilo de vida (mayor frecuencia de experimentar determinados escenarios), mayor intención de uso de la tecnología de voz.

El modelo propuesto aparece representado en el Gráfico 2.1.

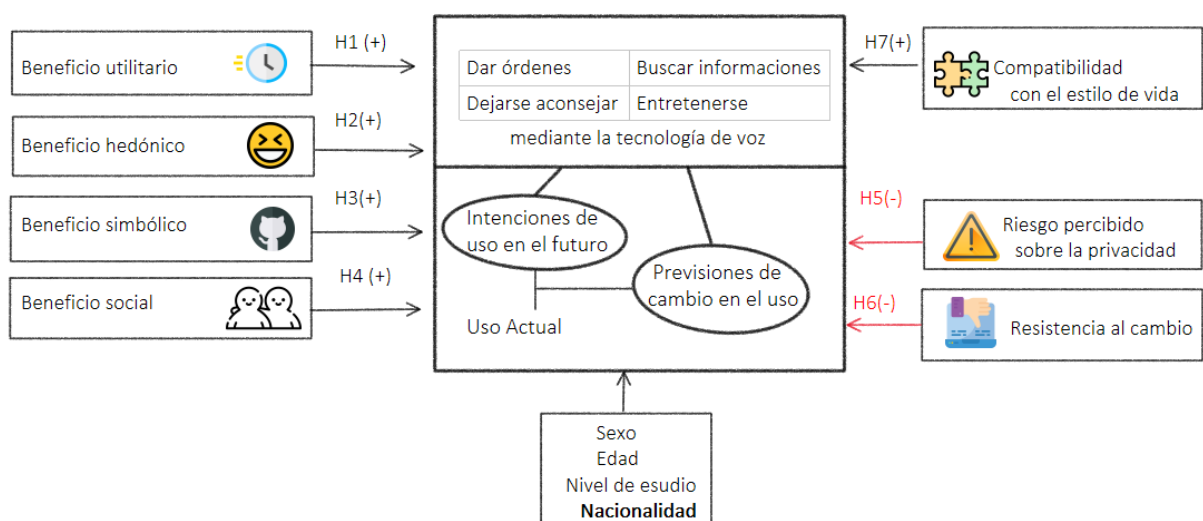


Gráfico 2.1: Modelo propuesto

3. METODOLOGÍA

3.1. Plan de muestreo y selección de la muestra

Para alcanzar los objetivos propuestos, se ha recogido información a través de una encuesta. Debido a la naturaleza del trabajo y a la falta de recursos por parte del investigador, la muestra se ha seleccionado por conveniencia. El cuestionario fue elaborado con el formulario de Google y difundido a través de las aplicaciones de mensajería (Whatsapp, Line). Con el fin de ampliar el tamaño de la muestra y poder hacer comparaciones entre países, se han elaborado dos versiones del cuestionario, una en español y otra en chino. Los cuestionarios completos pueden verse en el Anexo I.

3.2. Medida de las variables

Las variables y sus indicadores de medida se muestran en la Tabla 3.1 junto con los estadísticos descriptivos para las dos muestras: encuestados españoles y encuestados taiwaneses.

En cuanto a las variables dependientes, las intenciones de uso futuro de la tecnología de voz están medidas en una escala de siete valores en términos de grado de seguridad (ver Tabla 3.1), para cada uno de los cuatro niveles de uso. Y para obtener el paso de uso actual a las intenciones de uso en el futuro, es necesario obtener datos de uso actual. Así pues, el uso actual también se mide en una escala de siete valores, en términos de frecuencia. Para evitar confusiones a la hora de responder, al principio del cuestionario se ha incorporado una descripción breve y algunos ejemplos para cada uno de los cuatro niveles de uso.

Respecto a las variables explicativas, las variables de beneficios, el riesgo percibido y la resistencia al cambio se miden con escalas Likert de siete valores (donde 1= totalmente en desacuerdo y 7= totalmente de acuerdo). La compatibilidad con el estilo de vida se mide con una escala de siete valores en términos de frecuencias (donde 1= nunca y 7= me pasa muchas veces). La mayoría de los indicadores que se emplean para medir las variables se han adaptado de investigaciones previas (McLean y Osei-Frimpong, 2019). No obstante, la medida de la compatibilidad se ha desarrollado para este trabajo.

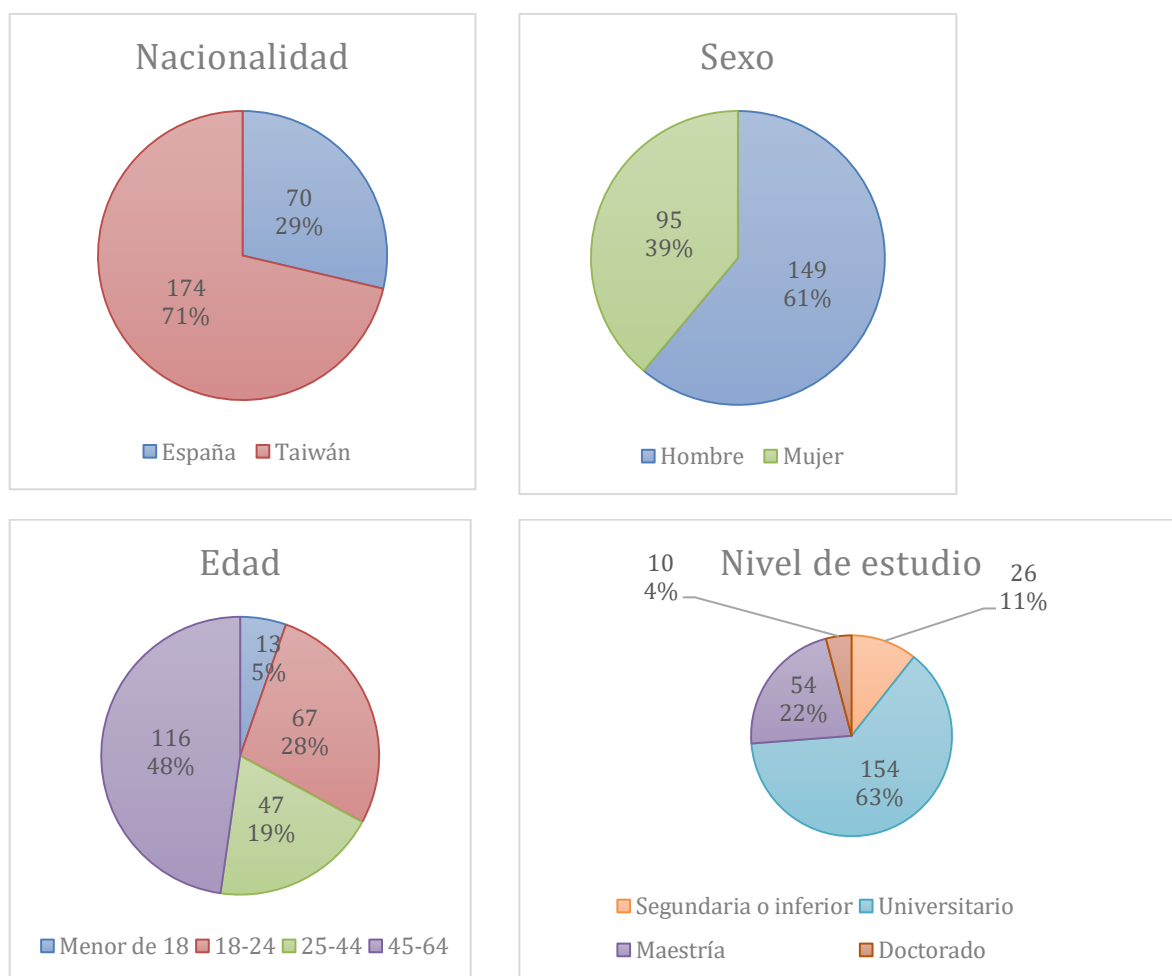
Variables y estadísticos descriptivos				
Variables e indicadores	Media ESP.	D.T. ESP.	Media TAIWAN	D.T. TAIWAN
Beneficio utilitario				
Creo que muchas veces la tecnología de voz falla.	4,77	1,39	4,57	1,41
Creo que la inteligencia del asistente de voz está por debajo de mi expectativa.	4,14	1,62	4,76	1,52
La tecnología de voz es conveniente para cumplir muchas tareas diarias.	5,08	1,45	5,64	1,27
La tecnología de voz es fácil de utilizar.	5,62	1,25	5,33	1,44
Creo la tecnología de voz es útil.	5,93	1,18	6,02	1,1
Beneficio hedónico				
Para mí la tecnología de voz es "guay".	5	1,66	5,76	1,30
Es entretenido pasar tiempo con mi asistente de voz.	3,9	1,74	5,21	1,4
Utilizar la tecnología de voz es divertido.	4,34	1,65	5,32	1,26
Beneficio simbólico				
Utilizar la tecnología de voz encaja con mi imagen personal.	4,08	1,8	4,16	1,57
Utilizar la tecnología de voz me hace o me haría destacar en mi grupo de amigos o conocidos.	2,83	1,67	3,71	1,64
Beneficio social *Así se refiere a un asistente casi perfecto				
Creo que un asistente de voz así* podría llegar a ser amigo mio.	3,3	2,07	4,71	1,69
Creo que podría compartir mis asuntos personales o sentimientos con una máquina así*.	2,69	1,93	3,7	1,78
Cuando tuviera que hablar con el asistente de voz, lo haría como si fuera una persona real.	3,58	2,23	4,68	1,63
Siempre tengo en cuenta que el asistente de voz es una máquina, una herramienta. Será difícil que desarrolle una relación social con ella.	5,56	1,78	5,17	1,54
Riesgo percibido sobre la privacidad				
Me preocupa que el asistente de voz guarde demasiadas informaciones mías.	5,63	1,62	5,16	1,57
Tengo dudas sobre la confidencialidad de mi interacción con el asistente de voz.	5,9	1,33	5,39	1,58
Creo que es muy inseguro comprar a través del asistente de voz.	5,41	1,43	5,55	1,45
Resistencia al cambio				
Creo que es raro hablar con la máquina.	3,72	1,85	3,75	1,7
Pienso que es incómodo hablar con la máquina.	3,38	1,76	3,68	1,68
Me sorprende ver a otros hablando con la máquina.	3,20	1,85	3,14	1,63
Me molesta si veo a otros hablando con la máquina.	2,14	1,62	2,9	1,59
Me da vergüenza si otros me vean hablando con la máquina.	3,15	1,75	3,51	1,74

VARIABLES E INDICADORES	Media ESP.	Des. ESP.	Media TAIWAN	Des. TAIWAN
Compatibilidad con el estilo de vida				
Frecuencia de experimentar los siguientes escenarios es: (1= Nunca me ha pasado, 7= Me pasa muchas veces)				
Tengo que poner el GPS cuando estoy conduciendo.	4,43	2,29	4,76	2,51
Hago llamadas o envío mensajes cuando estoy conduciendo/andando,	4,07	2,17	4,68	2,19
Me gusta escuchar algo cuando estoy haciendo tareas de casa,	4,49	2,24	5,36	1,92
Utilizo el móvil cuando estoy cocinando,	4,45	2,1	4,34	2,34
Me da pereza levantarse del sofá/cama para apagar la luz,	3,21	2,34	4,05	2,27
Me canso de escribir/teclear,	3,56	2,16	4,05	2,2
Cuando necesito sentirme acompañado e interactuar, no sé con quién puedo hablar,	2,06	1,59	4,05	2,2
Me cuesta hacer decisiones simples,	2,55	1,87	2,71	1,75
Frecuencia de uso actual de los 4 niveles de usos (1= Nunca, 7=De forma habitual)				
Utilizo la tecnología de voz para dar órdenes,	3,5	1,9	3,72	1,67
Utilizo la tecnología de voz para buscar informaciones,	4,46	1,99	3,91	1,82
Utilizo la tecnología de voz para dejarme aconsejar,	2,47	1,52	2,93	1,75
Utilizo la tecnología de voz para entretenerme,	2,87	1,89	3,4	1,9
Intención de utilizar los 4 niveles de usos				
Me gustaría seguir utilizando la tecnología de voz o empezar a utilizarla para:...				
(7=seguramente que sí y 1= seguramente que no)				
Dar órdenes	3,89	2,28	4,83	1,69
Buscar informaciones	4,67	2,17	4,83	1,70
Dejarme aconsejar	3	1,76	4,11	1,76
Entretenerme	3,41	1,95	4,34	1,75

Tabla 3.1: Variables y estadísticos descriptivos

3.3. Composición de la muestra

La muestra final está compuesta por 244 individuos, 70 de ellos son españoles y 174 son taiwaneses. En el Gráfico 3.1 puede verse que el porcentaje de españoles es solo del 29%. Para evitar el problema de que las opiniones de los taiwaneses tengan un mayor peso en los resultados, se analizarán los efectos de interacción para detectar diferencias significativas entre los dos países.



*En el gráfico de Edad, hay 1 individuo que pertenece al grupo de mayor de 65 años

Gráfico 3.1: Composición de la muestra (I)

También cabe mencionar que hay más hombres que mujeres, el rango de edad que cuenta con más individuos es el de 45-64 años y gran parte de la muestra tiene un nivel de estudios universitarios. No es una muestra muy equilibrada porque el cuestionario ha sido distribuido por conveniencia. De nuevo,

para evitar que estos desequilibrios confundan los resultados, se van a introducir estas variables como variables de control en el análisis.

Además, como puede verse en el Gráfico 3.2. la mayoría de la muestra, un 88% de los encuestados, tiene experiencia con la tecnología de voz porque la ha utilizado al menos en una ocasión. Por tanto, las opiniones obtenidas serán más fiables porque están basadas en experiencias reales.

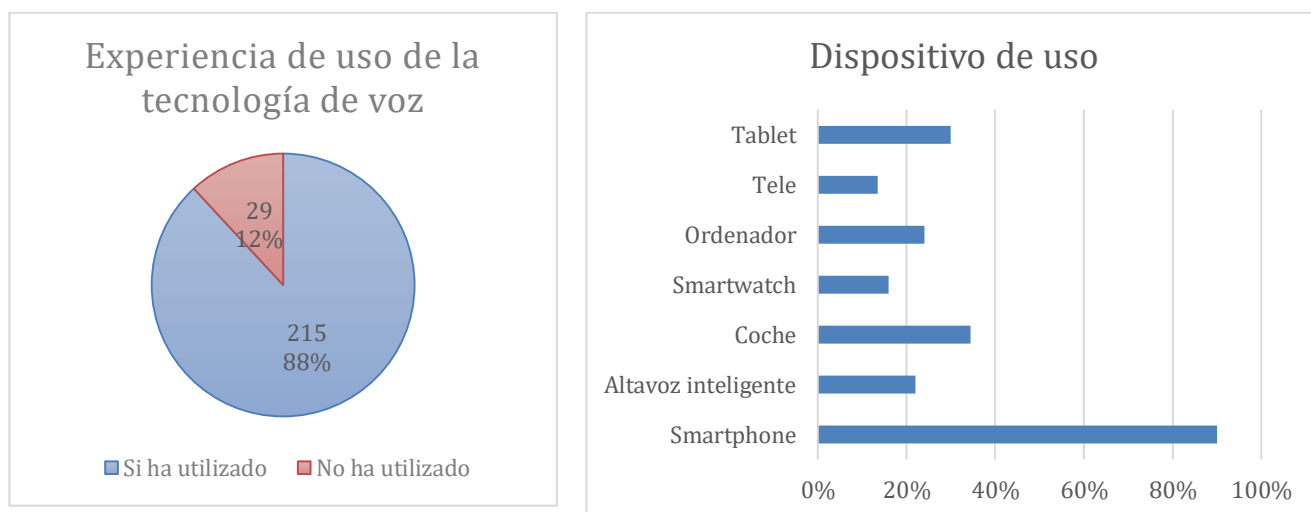


Gráfico 3.2: Composición de la muestra (II)

Por último, el dispositivo más utilizado por la muestra para acceder a la tecnología es el smartphone, seguido por el sistema inteligente de coche y por la tablet. El dispositivo estrella de la tecnología, el altavoz inteligente, aún no se ha extendido de forma llamativa en ninguno de los dos países analizados.

4. ANÁLISIS DE RESULTADOS

SPSS es el programa informático utilizado para llevar a cabo los análisis de los datos.

4.1. Análisis factorial para construir las medidas de las variables

Para poder someter a contraste empírico el modelo propuesto se han de construir las medidas de las variables que participan en él. En primer lugar, se trata de comprobar si los indicadores de medida empleados para cada variable se agrupan empíricamente en sus correspondientes variables. Para ello, se ha

decidido aplicar la técnica de análisis factorial. Al principio había preocupación acerca de si los individuos de los dos países entendían de distinta forma las preguntas del cuestionario, por cuestión de idioma y/o cultura, lo que podría acarrear problemas de medición. Pero tras los análisis preliminares se pudo comprobar que el proceso y el resultado de generar los factores son muy similares en las muestras de ambos países. Por tanto, se pueden crear las medidas de las variables con los factores para la muestra completa.

Para evitar el problema de multicolinealidad, se ha aplicado el análisis factorial para tres grandes bloques de indicadores:

- Bloque 1: Todos los indicadores de las variables de Beneficios (utilitario, hedónico, simbólico y social).
- Bloque 2: Todos los indicadores de las variables de freno (Riesgo y Resistencia).
- Bloque 3: Todos los indicadores de los escenarios (Compatibilidad con el estilo de vida).

Los resultados obtenidos muestran que el Beneficio utilitario se desprende en 2 factores: Utilidad y Fallo técnico, Compatibilidad con el estilo de vida también se desprende en 2 factores: escenarios de multitarea y escenarios de pereza. Y cada una de las demás variables tiene su factor correspondiente. En el Anexo II se proporcionan los resultados de estos análisis factoriales.

Es decir, el output de los análisis factoriales son nueve factores que serán utilizados como variables explicativas en el posterior análisis de regresión:

- | | |
|------------------------|---|
| 1. Utilidad | 6. Riesgo percibido sobre la privacidad |
| 2. Fallo técnico | 7. Resistencia al cambio |
| 3. Beneficio hedónico | 8. Escenarios de multitarea |
| 4. Beneficio simbólico | 9. Escenarios de pereza |
| 5. Beneficio social | |

4.2. Regresión lineal: Análisis de los determinantes de la intención de utilizar la tecnología de voz

Las intenciones de uso de la tecnología en el futuro tienen valores medios que oscilan entorno al 4, que es el punto medio de la escala de medida, y

desviaciones típicas elevadas. Estos valores ponen de manifiesto que los encuestados no tienen mucha seguridad en sí utilizarán el asistente de voz en el futuro, aunque las opiniones de los encuestados están bastante dispersas.

Según los valores medios de las intenciones para toda la muestra, los usos futuros de la tecnología más aceptados son utilizar el asistente para buscar informaciones y dar órdenes. El nivel de uso con la media de intención más baja es dejarse aconsejar. Los dos primeros tipos o niveles de uso son más instrumentales. Quizás los individuos perciben más utilidad real en emplear la tecnología para dar órdenes y buscar información. Además, son los usos más difundidos y familiares en el mercado actual y a los que se recurre en la publicidad de los asistentes de voz. Por el contrario, utilizar la tecnología para dejarse aconsejar y entretenerse son usos que no tienen una finalidad tan práctica, son menos conocidos e implican un mayor grado de adopción de la tecnología que dar órdenes al asistente o mandarle buscar información.

Según los valores medios de las intenciones por país (ver Tabla 3.1), se puede decir que los encuestados de Taiwán tienen más seguridad de que utilizarán la tecnología en el futuro, para todos los niveles de uso.

Tras conseguir las variables explicativas a través del análisis factorial, se procede a realizar la regresión para explicar la intención de utilizar los cuatro niveles de la tecnología de voz, lo que sería el primer objetivo del presente estudio. Además, para ver las diferencias entre los dos países estudiados, tercer objetivo del trabajo, se han creado las variables de interacción multiplicando cada variable explicativa por la nacionalidad. La regresión lineal se hace por bloques de variables que entran en la ecuación en pasos sucesivos. Cada bloque que hemos formado agrupa variables de distinta naturaleza:

- Bloque 1. Variables de control: Sexo, Edad, Nivel de estudio y Nacionalidad.
- Bloque 2. Variables explicativas: Utilidad, Fallo Técnico (FTécnico), Beneficio Hedónico (BHedónico), Beneficio simbólico (BSimbólico), Beneficio social (BSocial), Riesgo percibido sobre la privacidad (Riesgo), Resistencia al cambio (Resistencia), Escenarios de multitarea (Multitarea) y Escenarios de pereza (Pereza).

- Bloque 3. Variables de interacción: Utilidad x Nacionalidad, Fallo Técnico x Nacionalidad, ... lo mismo para todas las variables explicativas, En este bloque se ha aplicado el método por pasos, con F de entrada de 0,1 y F de eliminación de 0,15, de esta forma solo entran las interacciones significativas con este criterio en el modelo.

Los resultados del análisis de regresión se muestran en la Tabla 4.1. Como se puede observar, para cada variable independiente se proporcionan los resultados de los tres modelos (M1, M2 y M3) a los que da lugar el procedimiento de introducción de las variables agrupadas por bloques en la ecuación de regresión.

En cuanto a la bondad de ajuste de las regresiones, el modelo propuesto explica cerca del 40% de las intenciones de dar órdenes y buscar informaciones, y el 33% de las intenciones de dejarse aconsejar y entretenerse. Los porcentajes de explicación no son altos, pero en el contexto del comportamiento del consumidor pueden ser aceptables. En todas las intenciones, el cambio del R^2 ajustado más significativo se produce al meter las variables explicativas principales (beneficios, frenos y compatibilidad).

	Resultados de la regresión											
	Intención de dar órdenes			Intención de buscar informaciones			Intención de dejarse aconsejar			Intención de entretenerse		
	M1	M2	M3	M1	M2	M3	M1	M2	M3	M1	M2	M3
(Constante)	3,175	4,019	3,779	3,408	4,059	4,104	2,146	2,753	2,577	2,862	3,683	3,653
Sexo	0,535	0,261	0,239	0,439	0,342	0,361	0,399	0,277	0,299	0,348	0,220	0,276
Edad	-0,121	-0,097	-0,097	0,072	0,098	0,086	0,017	0,008	0,010	-0,124	-0,136	-0,162
NivelDeEstudios	0,335	0,069	0,134	0,381	0,106	0,086	0,265	0,118	0,167	0,323	0,167	0,150
Nacionalidad	0,979	0,740	0,813	0,072	-0,025	-0,004	1,068	0,818	0,839	0,952	0,442	0,533
Utilidad		0,405	0,385		0,571	0,580		0,268	0,240		0,254	0,262
FTécnico		0,092	0,089		-0,030	-0,015		-0,117	-0,125		0,069	0,103
BHedonico		0,287	0,312		0,212	0,240		0,214	0,224		0,555	0,615
BSimbólico		0,082	0,105		-0,012	-0,358		0,271	0,311		0,255	-0,303
Bsocial		0,015	0,011		-0,202	-0,224		-0,009	0,000		-0,083	-0,120
Riesgo		-0,280	-0,264		-0,206	-0,183		-0,041	-0,488		0,006	0,059
Resistencia		-0,384	-0,771		-0,230	-0,208		-0,173	-0,146		-0,007	0,052
Multitarea		0,602	0,608		0,568	0,558		0,491	0,471		0,461	0,459
Pereza		0,081	-0,167		0,389	0,403		0,310	-0,205		0,295	0,032
In_BSimbólico					0,465							0,759
In_Pereza			0,384						0,742			0,403
In_Resistencia			0,548									
In_Riesgo									0,612			
R ²	0,084	0,401	0,427	0,35	0,398	0,409	0,096	0,324	0,37	0,075	0,323	0,37
R ² Ajustado	0,069	0,367	0,389	0,018	0,364	0,373	0,8	0,286	0,328	0,06	0,285	0,328
F	5,468	11,859	11,822	2,140	11,713	11,333	6,312	8,479	8,921	4,875	8,443	8,919
Cambio R ²	0,084	0,317	0,025	0,35	0,3364	0,011	0,096	0,228	0,046	0,075	0,248	0,046
Cambio F (sig.)	0	0	0,088	0,077	0	0,04	0	0	0,015	0,001	0	0,085
Estadísticos descriptivos												
Media		4,56			4,79			3,79			4,08	
D.T.		1,92			1,85			1,82			1,85	

*El coeficiente β aparece en negrita si la variable correspondiente es significativa. El nivel de confianza es de 95% para las variables de control y las explicativas, y 90% para las variables de interacción.

Tabla 4.1: Resultados de la regresión lineal para la intención de uso

Variables de control

- La variable sexo solo es significativa para la intención de dar órdenes y muestra que los hombres tienen mayor intención de dar órdenes por voz

que las mujeres. El nivel de estudios solo es significativo para la intención de buscar informaciones, de forma que a mayor nivel de estudios mayor es la intención, Y la edad no es significativa para ninguna de las intenciones.

- La nacionalidad no afecta a la intención de buscar informaciones, pero si es significativa en las otras tres. En general, los encuestados de Taiwán muestran más intenciones de probar o seguir utilizando la tecnología de voz que los de España.

Variables explicativas

Variables de beneficios

- La utilidad percibida es significativa en todas las intenciones y su efecto positivo es de los más fuertes, sobre todo en dar órdenes y buscar informaciones utilizando el asistente de voz. Quizás es porque estas dos funciones tienen una finalidad más utilitaria. Es decir, los usuarios quieren llegar a determinados resultados a través de estos dos usos.
- El beneficio hedónico es significativo en todas las intenciones salvo en la de dejarse aconsejar. Su influencia positiva es especialmente relevante en la intención de entretenerse utilizando el asistente de voz, lo cual es algo lógico.
- El beneficio simbólico tiene un efecto positivo y significativo en las intenciones de dejarse aconsejar y entretenerse utilizando el asistente de voz. Quizás estas dos funciones son relativamente más novedosas, por tanto, sirven más para que los usuarios expresen su imagen tecnológica o destacarse dentro de sus grupos de iguales.
- El fallo técnico y el beneficio social no poseen efectos significativos en ninguna de las intenciones.

Variables de frenos

Las variables de frenos propuestos son el riesgo percibido sobre la privacidad y la resistencia al cambio. Tal y como esperábamos, ambos tienen efectos negativos significativos sobre las intenciones de dar órdenes y buscar informaciones por voz. ¿Por qué solo en estas dos funciones y no en otras?

Respecto al riesgo percibido, quizás estas dos funciones tengan un carácter más utilitario, por tanto, las informaciones guardadas de esas actividades son

más serias para los usuarios (por ejemplo, datos de compra, datos personales de dirección, datos personales de llamadas, etc.), mientras que la función de dejarse aconsejar y entretenerse no suele derivar datos serios de los usuarios.

Respecto a la resistencia al cambio, en concreto la resistencia a la acción de hablar a la máquina, quizás la gente se resiste más a hablar con tono de mandatos, que es el caso de estas dos funciones. En cambio, cuando la gente se deja aconsejar o se entretiene con el asistente de voz, se utilizaría un tono más de conversación, con lo cual habría más aceptación porque se actúa como si el receptor del mensaje fuese otra persona.

Variables de compatibilidad con el estilo de vida (escenarios)

- La variable escenarios de multitarea muestra un efecto positivo y significativo en todas las intenciones. Es más, su efecto es uno de los más importantes en todas las variables dependientes. Es decir, cuanto más experimentan los encuestados estos escenarios, más intenciones tendrán de utilizar la tecnología de voz.
- La variable escenarios de pereza tiene, también, una influencia positiva y significativa sobre todas las intenciones de utilizar el asistente de voz salvo para la intención de dar órdenes. Sus efectos tienen un grado de importancia intermedio en las intenciones. Resulta extraño que la frecuencia con la que se presentan escenarios de pereza no afecte a la intención de dar órdenes. Por ejemplo, los momentos de apagar la luz cuando uno no quiere levantarse o los casos en los que hay que escribir texto cuando uno tiene pereza de teclear, deberían ser un impulso al uso de esta función. Parece, sin embargo, que esta “pereza” no es tan crítica como para que los usuarios se planteen utilizar el asistente de voz para dar órdenes.

Variables de Interacción

Algunas de las variables explicativas tienen efectos muy diferentes por país.

- El beneficio simbólico tiene el efecto de impulsar las intenciones de entretenerse y buscar informaciones, mucho mayor para el público taiwanés que para el español. En el caso de España, no tiene efecto significativo sobre la intención de entretenerse y lo tiene negativo sobre la intención de

buscar informaciones.

- Los escenarios de pereza motivan mucho más a los taiwaneses a utilizar la tecnología de voz, excepto en la función de buscar informaciones por voz.
- La resistencia tiene un mayor efecto negativo para los españoles en su intención de dar órdenes. Y el efecto negativo derivado del riesgo también es mayor para los españoles en la intención de dejarse aconsejar.

Conclusiones parciales

- La compatibilidad con el estilo de vida (reflejada en las variables de escenarios) es importante para explicar la intención de utilizar la tecnología de voz. Podemos observar que incluso es más importante que la utilidad percibida, ya que el individuo puede pensar que la tecnología de voz es útil, pero si no encaja con su estilo de vida, no tendría tantas intenciones de utilizarla.
- Los taiwaneses muestran mayores intenciones de utilizar la tecnología de voz. Además, en el caso de Taiwán, se intensifican los efectos de las variables impulsoras y se atenúan los efectos de las variables de freno en algunos de los cuatro niveles de uso analizados.
- No sirven para explicar las intenciones ni el fallo técnico ni el beneficio social, En otras palabras, el hecho de que la tecnología produzca muchos fallos o que no cumpla las expectativas no influye sobre la intención de utilizarla. Tampoco influye el grado de aceptación por parte del individuo de la existencia de una “relación” entre el asistente de voz y el hombre.
- Es posible dividir las intenciones en dos bloques. Dar órdenes y buscar informaciones como el primer bloque, dejarse aconsejar y entretenerse como el segundo. El primer bloque tiene una naturaleza más instrumental y utilitaria, contiene las funciones más extendidas hasta el momento y el usuario tiende a interactuar con la máquina a través de un tono de mandatos. El segundo bloque tiene una naturaleza más informal y menos instrumental, contiene las funciones consideradas como más novedosas en las que el usuario tiende a interactuar con los dispositivos con tonos de conversación. Quizás por ello, se pueden observar dos patrones de los efectos de las variables explicativas. El efecto de la utilidad es mucho mayor en el primer

bloque de intenciones, el riesgo y la resistencia solo tienen efectos en el primer bloque y el beneficio simbólico solo tiene efecto en el segundo bloque.

4.3. Regresión logística: análisis de los determinantes de la previsión de cambio en el uso

La variable “previsión de cambio en el uso” se ha construido a partir de la diferencia entre la intención de uso en el futuro y la frecuencia de uso actual (Intención – Uso actual). De esta forma podemos conocer el paso desde el uso actual a intención de uso del futuro. El uso actual y la intención de uso futuro están medidas en distintas escalas, el primero en una escala de frecuencia y la segunda en una escala de seguridad, a pesar de ello, es posible interpretar sus diferencias.

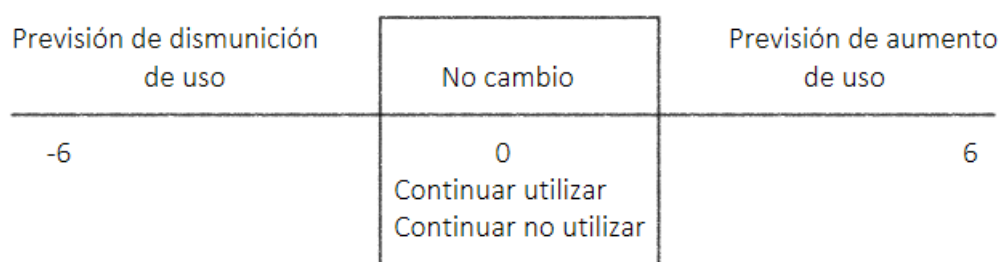


Gráfico 4.1: Ilustración de la interpretación de la previsión de cambio del uso

Si la diferencia es 0, significa que el individuo seguirá haciendo lo que está haciendo ahora. Si está utilizando la tecnología de voz de forma habitual, está seguro de que la va a seguir utilizando. Si la utilizando poco ahora, no está tan seguro sobre su uso continuado.

Si la diferencia no es 0, significa que el individuo cree que hará algo muy diferente a lo que está haciendo ahora. En este caso hay dos posibilidades:

- Diferencia negativa: el individuo prevé que en el futuro utilizará menos que ahora.
- Diferencia positiva: el individuo prevé que en el futuro utilizará más que ahora.

El análisis se centra solamente en las dos previsiones de cambio, porque se cree que es más interesante hallar las razones por las que los individuos prevén

que van a cambiar sus comportamientos de uso de la tecnología de voz. Para facilitar el análisis, se clasifican los individuos en tres grupos, los que no van a cambiar, los que van a utilizar menos que ahora y los que van a utilizar más que ahora. La técnica de análisis que se aplica es la regresión logística, donde se atribuye el valor 0 al grupo que va a utilizar menos y 1 al grupo que va a utilizar más. Los individuos que no prevén cambio en su comportamiento serán excluidos en el análisis.

Las variables explicativas consideradas para el análisis son las mismas que en la regresión lineal. También aplicaremos el procedimiento de introducir las variables por pasos sucesivos en la ecuación a estimar, agrupadas en los mismos tres bloques que ya explicamos en el apartado anterior. Primero entrarán en la ecuación las variables de control, luego las variables explicativas y por último las de interacción. En el último bloque se aplica el método hacia adelante de Wald con el criterio valor de entrada 0,1 y de eliminación 0,15. Los resultados del análisis de regresión logística se muestran en la Tabla 4.2.

Según el criterio de Nagelkerke, el modelo aporta mayor explicación a la previsión de cambio en dejarse aconsejar por voz (50%), seguida por dar órdenes y buscar informaciones (42%). La menor explicación le corresponde a la previsión en entretenerse (35%). Los porcentajes de explicación se consideran aceptables al ser un modelo de comportamiento del consumidor. Por el otro lado, los porcentajes de casos bien clasificados superan el 80% en todas las previsiones, por tanto, el modelo está bastante bien ajustado.

Resultados de la regresión logística										
	Previsión cambio de dar órdenes			Previsión cambio de buscar informaciones		Previsión cambio de dejarse aconsejar			Previsión cambio de entretenerse	
	M1	M2	M3	M1	M2-M3	M1	M2	M3	M1	M2-M3
(Constante)	-0,307	0,267	-0,107	0,085	0,483	0,665	1,020	0,32	-0,271	0,020
Sexo	-0,113	-0,304	-0,222	-0,296	-0,102	-0,526	-0,526	-0,145	-0,082	0,162
Edad	-0,165	-0,107	-0,132	-0,131	0,043	-0,409	-0,394	-0,642	-0,237	-0,190
NivelDeEstudios	0,639	0,398	0,465	0,522	0,266	0,782	0,617	1,049	0,825	0,654
Nacionalidad	1,101)	1,603	1,928	0,737	0,723	1,186	1,822	2,312	0,974	1,152
Utilidad		0,318	0,287		0,092		0,027	-0,218		-0,049
FTécnico		-0,048	-0,066		-0,074		0,016	0,100		0,458
BHedónico		0,318	-0,160		0,351		0,186	-1,104		0,282
BSimbólico		-0,726	-0,786		-0,734		-0,774	-0,850		-0,340
BSocial		-0,758	-0,831		-0,627		-0,566	-0,756		-0,813
Riesgo		-0,264	-0,347		-0,345		-0,260	-0,310		-0,196
Resistencia		-0,085	-0,105		0,175		-0,141	-0,040		0,016
Multitarea		0,781	0,780		1,109		0,936	1,028		0,787
Pereza		0,267	0,250		0,584		0,073	-0,788		0,191
In_BHedónico			0,967					2,354		
In_Pereza								1,270		
R ² Cox y Snell	0,076	0,26	0,277	0,047	0,285	0,083	0,233	0,31	0,069	0,223
R ² Nagelkerke	0,116	0,394	0,42	0,07	0,421	0,134	0,377	0,502	0,109	0,349
Sig, Prueba Omnibus	0,005	0	0,038	0,064	0	0,003	0	0,025	0,012	0
% bien clasificados	75,9%	80,7%	84%	71,6%	80,3%	78,1%	81,4%	85,8%	76,4%	84,3%
Casos analizados	187 (76,7%)			183 (75%)		183(75%)			178 (73%)	
Casos perdidos	57 (23,4%)			61(25%)		61(25%)			66 (27%)	

*El coeficiente β aparece en negrita si la variable correspondiente es significativa. El nivel de confianza es de 95% para las variables de control y las explicativas, y 90% para las variables de interacción.

Tabla 4.2: Resultados de la regresión logística para la previsión de cambio en el uso

Variables de control:

- Las variables sexo y edad no son significativas para explicar las previsiones del cambio en el uso.
- La variable nivel de estudios es significativa en todas las previsiones salvo la de buscar informaciones. Su efecto es positivo, de forma que un mayor

nivel de estudio implica mayor probabilidad de darse un aumento en el uso de dar órdenes, dejarse aconsejar y entretenerse utilizando el asistente de voz.

- La nacionalidad posee un efecto significativo en todas las previsiones. En general, los taiwaneses tienen más probabilidad de aumentar su uso de la tecnología de voz, frente a los españoles.

Variables explicativas

- No tienen efectos significativos en ninguna previsión: utilidad, fallo técnico, beneficio hedónico, riesgo y resistencia.
- El beneficio simbólico afecta de forma significativa a todas las previsiones salvo a la de entretenerse. En contra de lo esperado, un mayor beneficio simbólico supone una mayor probabilidad de que el individuo anticipe una disminución en su uso. Quizás a los individuos no les guste que la tecnología les haga destacar demasiado dentro de su grupo de iguales y, por ello, el aumento de beneficio simbólico provoque la previsión a la baja. O bien que el beneficio simbólico se pierda a medida que pasa el tiempo porque la tecnología deje de ser novedosa.

Este resultado entra en contradicción con el resultado obtenido en la regresión lineal. Según el resultado de la regresión lineal sobre las intenciones, el beneficio simbólico contribuye a aumentar las intenciones de utilizar las dos funciones más novedosas, dejarse aconsejar y entretenerse. Pero ahora hemos observado que ese beneficio provoca una previsión a la baja en la función de dejarse aconsejar. Creemos que puede deberse a que en este análisis hemos dejado fuera una parte de la muestra (los que no cambian sus previsiones), y quizás dentro de esa parte eliminada existen individuos para los que el beneficio simbólico sí que sirve como acicate a su intención de dejarse aconsejar utilizando el asistente de voz.

- El beneficio social influye significativamente en todas las previsiones. De nuevo en contra de lo esperado, mayor beneficio social implica una mayor probabilidad de encontrarse a individuos con previsiones a la baja. Quizás es porque a medida que suba el beneficio social, el individuo otorgaría mayor identidad personal al asistente de voz, lo que puede cambiar el

contenido de la interacción. Es posible que el hombre deje de ver al asistente de voz como una herramienta y espere que desempeñe otros papeles más allá del ámbito de las cuatro funciones actuales.

- Respecto a las variables de compatibilidad, la influencia de escenarios de multitarea resulta significativa para todas las previsiones, mientras que Escenarios de pereza solo es significativa para la previsión de buscar informaciones. Una mayor frecuencia de experimentar dichos escenarios conduce a mayor probabilidad de que el individuo anticipe un aumento de su uso de la tecnología de voz.

Variables de interacción

- El beneficio hedónico influye sobre las previsiones de los taiwaneses en dar órdenes y dejarse aconsejar, pero no tiene efecto sobre las de los españoles. En concreto, hace aumentar la probabilidad de que los taiwaneses anticipen un aumento de uso en estas dos funciones.
- La variable escenarios de pereza presenta efectos significativamente diferentes para los dos países sobre la previsión de cambio en dejarse aconsejar por el asistente de voz. Su efecto sobre la probabilidad de prever aumentos de uso es negativo para los españoles, y positivo para los taiwaneses.

Conclusiones parciales

- En general, son más numerosos en la muestra los individuos que prevén que utilizarán más a la tecnología de voz en el futuro que en el momento actual.
- Los taiwaneses tienen más probabilidad de pertenecer al grupo que anticipe un aumento de uso de la tecnología de voz. Además, esa probabilidad es fomentada por el beneficio hedónico en la función de dar órdenes y dejarse aconsejar.
- El beneficio simbólico y social provocan las previsiones a la baja. El primero afecta a las funciones de dar órdenes y buscar informaciones porque los individuos no quieren destacar frente a otros por el uso de estas dos funciones. El segundo afecta a todas las funciones, quizás cuando se

percibe el beneficio social, el individuo tiene menos ganas de usar estas funciones superficiales y tratará de buscar nuevas posibilidades en la aplicación de la tecnología de voz.

- Escenarios de multitarea es la única variable que fomenta las previsiones de aumento en el uso. Entonces, podemos reafirmar que la compatibilidad con el estilo de vida, más concretamente la frecuencia con la que se presenta la necesidad de realizar varias tareas a la vez es relevante en la adopción de la tecnología de voz.

5. CONCLUSIONES

El presente estudio persigue tres objetivos, buscar y analizar los determinantes de la intención de utilizar la tecnología de voz en el futuro distinguiendo por tipo o nivel de uso, buscar y analizar los determinantes de las previsiones en el cambio de uso de la tecnología y averiguar si existen diferencias entre los dos países de análisis.

Para ello se ha propuesto un modelo basado en las teorías y modelos de adopción de tecnología y usos y gratificaciones y se ha contrastado empíricamente.

En general, los resultados muestran que los efectos de las variables explicativas (beneficios, frenos y compatibilidad) sobre las intenciones de uso de la tecnología cumplen con las expectativas del modelo planteado, aunque esos efectos difieren según el nivel o tipo de uso (dar órdenes, buscar informaciones, dejarse aconsejar y entretenerse). También es cierto que algunas hipótesis no se cumplen, porque hay variables en el modelo cuyos efectos no son significativos.

Respecto a la intención de uso de la tecnología de voz, se ha obtenido que la utilidad y la compatibilidad con el estilo de vida son importantes para todas las funciones o niveles de uso de la tecnología considerados. La compatibilidad parece tener un peso más importante que la utilidad. Para los individuos es importante percibir que la tecnología es útil y resuelve problemas, pero lo es más que encaje en su estilo de vida y tenga ocasiones de “brillar”, jugando su papel en algunos de sus momentos diarios. También cabe mencionar que el beneficio

hedónico es otra variable motivadora de la intención de adopción, sobre todo para la función de entretenerse.

Partiendo del comportamiento de los coeficientes, es posible dividir las funciones en dos grandes bloques, el primero formado por dar órdenes y buscar informaciones y el segundo formado por dejarse aconsejar y entretenerse. El primero tiene una naturaleza más utilitaria, formal e instrumental y contiene las funciones más conocidas y familiares hasta ahora. Por tanto, tiene sentido que las variables de freno (resistencia y riesgo) solo sean significativas en este bloque. Los individuos se resisten a dar órdenes y mandatos de búsqueda de información quizás porque es más difícil hablar a la máquina con un tono de mandatos. La importancia del riesgo quizás se deba a que las actividades de este primer bloque derivan más datos serios del usuario. También, el efecto de la utilidad es mayor en este primer bloque.

El segundo bloque, en cambio, es más informal y contiene las funciones más novedosas y divertidas. El beneficio simbólico percibido del uso de la tecnología solo tiene efectos significativos en las intenciones futuras de este segundo grupo. Los individuos creen que pedir consejo al asistente o entretenerse con él encaja con su imagen personal y les ayuda a destacar entre amigos y conocidos.

Respecto a la previsión en el cambio de uso, se ve que no hay tantas diferencias entre los cuatro niveles de uso, salvo para la variable beneficio simbólico. Además, se puede observar que los individuos que prevén un mayor uso en el futuro son mayoría.

Entrando en el propio análisis, vemos que un mayor nivel de estudio implica más probabilidad de anticipar un incremento de uso, pues es probable que los que tienen más nivel de estudios tengan más contactos con informaciones sobre las tendencias de futuro, ámbito en el que la tecnología de voz está reclamando continuamente su importancia. También cabe mencionar que el beneficio social y el simbólico tienen un efecto contrario al propuesto, porque, según los resultados, mayor beneficio social y mayor beneficio simbólico conducen a mayor probabilidad de anticipar una disminución en el uso. Puede que alcanzado un determinado nivel de beneficio social los usuarios dejen de ver a los asistentes de voz como máquinas y traten de buscar otras aplicaciones más allá de estas cuatro funciones superficiales. De la misma forma, alcanzar un

determinado nivel de beneficio simbólico puede que haga que los individuos sientan que destacan demasiado por el uso del asistente. O bien, que piensen que la tecnología ha dejado de ser novedosa como para ofrecerles un beneficio simbólico. Por último, otra vez destaca escenarios de multitarea como el único factor que fomenta el aumento de uso en el futuro.

En cuanto a la comparación entre los dos países, los taiwaneses muestran más tendencia a la adopción de la tecnología de voz. Tienen mayor intención de probar las cuatro funciones, salvo la de buscar informaciones, en la que destacan los españoles. Además, el hecho de ser taiwanés intensifica los efectos positivos de las variables impulsoras y atenúa los efectos negativos de las variables de freno en algunas de las funciones. En la Tabla 5.1 se presenta un resumen de las diferencias entre ambos países.

Efecto impulsor por ser un taiwanés				
	Dar órdenes por voz	Buscar informaciones	Dejarse aconsejar	Entretenerse
Intenciones	--- Resistencia +++ Pereza	+++BSimbólico	--- Riesgo +++ Pereza	+++Simbólico +++Pereza
Previsiones de cambio en el uso	+++BHedónico		+++BHedónico +++Pereza	

*--- : Atenúa significativamente el efecto de dichas variables.

*+++ : Fomenta significativamente el efecto de dichas variables.

Tabla 5.1: Resumen de las diferencias entre España y Taiwan

Escenarios de pereza parece ser una variable que provoca una diferencia importante entre los dos países, pues quizás a los taiwaneses les importa mucho más la comodidad, esto se evidencia con la alta densidad geográfica de las tiendas de conveniencias en Taiwán. Por tanto, si hay una solución tecnológica que hace que sus vidas sean más cómodas, el grado de atracción es mucho mayor.

Otros resultados de interés.

- Las funciones del primer bloque son más populares (más usuarios habituales y mayores intenciones de uso) posiblemente porque su uso está más extendido en este momento. Entre ellas, la función de buscar informaciones es la más utilizada en ambos países (más encuestados situados en la parte de alta frecuencia de uso).

- Las funciones del segundo bloque son menos utilizadas que las del primero en ambos países, aunque una minoría afirma que mantiene su uso habitual de estas funciones. Entre ellas, la intención de dejarse aconsejar por voz es menor en ambos países.

6. SUGERENCIAS PARA INVESTIGACIONES FUTURAS

Existen muchos más asuntos interesantes en la adopción de la tecnología de voz que el presente estudio no ha explorado y que pueden ser el objetivo de futuras investigaciones.

- Cuestiones relacionadas con la pantalla. Quizás los individuos creen que la tecnología de voz está bien, pero no tienen mucha intención de probarla porque están acostumbrados a utilizar la interfaz visual-táctil de siempre y por tanto no creen necesaria otra interfaz de voz. Sería interesante ver cómo esa actitud puede influir en la intención de uso de la tecnología de voz.
- Disponibilidad de los dispositivos complementarios. En el presente trabajo no se ha entrado mucho en la cuestión de la disponibilidad de dispositivos complementarios, por ejemplo, para apagar la luz a través de la voz, el individuo debe tener una bombilla conectable.
- Tono para hablar con la máquina. Parece que el usuario está más dispuesto a hablar con la máquina como si fuera una persona real, es por eso que en las funciones más utilitarias el efecto de la resistencia al cambio es mayor ¿Y cómo se puede reducir esos efectos negativos? Se observa que en la vida normal los individuos también dan órdenes o piden informaciones a otras personas. Las empresas quizás pueden intentar hacer un diseño de interacción más humano para estas dos funciones y ver los efectos sobre el uso (por ejemplo, permitir personalización del nombre o voz del asistente).
- Marcas de asistente de voz. Cuestiones tales como si los usuarios de Siri tienen mayor intención de uso en algunas funciones que los usuarios de Google asistente, la fidelidad y la satisfacción a las distintas marcas de asistente de voz.
- Cuestiones relacionadas con el entorno. En este estudio no se ha considerado la influencia del entorno sobre los individuos. Por ejemplo: ¿Cómo perciben la generalización de la tecnología de voz?, ¿Alguien de sus

círculos cercanos utiliza la tecnología de voz? o ¿Ve mucha publicidad o películas que contienen la tecnología de voz?...

7. BIBLIOGRAFÍA

AIMC (2019): "AIMC elabora un estudio sobre el uso de los asistentes virtuales de voz". Disponible en <http://electro-imagen.com/articulo/aimc-elabora-un-estudio-sobre-el-uso-de-los-asistentes-virtuales-de-voz>, [consulta: 18/08/2020],

AIMC (2020): "22º Encuesta a usuarios de internet", pp. 25-26.

Alan (2019): "What is a Voice User Interface?". Disponible en <https://alan.app/blog/voiceuserinterface/>, [consulta: 22/07/2020].

Bagozzi, R. P. (2007): "The legacy of the technology acceptance model and a proposal for a paradigm shift". *Journal of the Association for Information Systems*, 8(4), pp. 244–254.

Carreras, R. (2019): "Datos y estadísticas sobre el crecimiento del mercado de asistentes de voz". Disponible en <https://robertocarreras.es/datos-y-estadisticas-sobre-el-crecimiento-del-mercado-de-asistentes-de-voz/>, [consulta: 18/08/2020].

Davis, F. D., Bagozzi, R. P. y Warshaw, P. R. (1989): "User acceptance of computer technology: a comparison of two theoretical models". *Management Science*, 35(8), pp. 982-1003.

Davis, F. D. (1989): "Perceived usefulness, perceived ease of use, and user acceptance of information technology". *MIS Quarterly*, 13(3), pp. 319-340.

Galeano, S. (2019): "Así usamos los asistentes de voz: una tecnología con mucho margen de mejora". *MARKETING ECOMMERCE*. Disponible en <https://marketing4ecommerce.net/asi-usamos-los-asistentes-de-voz-una-tecnologia-con-mucho-margen-de-mejora/>, [consulta: 18/08/2020].

GrowthHK (2020): "Comparación entre VUI y GUI". Disponible en https://www.growthhk.cn/guan/28737.html?utm_source=tuicool&utm_medium=referral, [consulta: 20/07/2020].

Hoy, M. B. (2018): "Alexa, Siri, Cortana, and more: An introduction to voice

assistants". *Medical Reference Services Quarterly*, 37(1), pp. 81–88.

Lee, K. F. y Wang, Y. G. (2017): "Que viene la inteligencia artificial". Editorial Commonwealth Publishing Co., Ltd., Taipei, pp. 40-55.

Legris, P., Ingham, J. y Colletette, P. (2003): "Why do people use information technology? A critical review of the technology acceptance model". *Information & Management*, 40 (3), pp. 191-204.

Lunceford, B. (2009): "Reconsidering Technology Adoption and Resistance: Observations of a Semi-Luddite". *Explorations in Media Ecology*, 8 (1), pp. 29-48.

Marangunić, N. y Granić, A. (2014): "Technology acceptance model: a literature review from 1986 to 2013". *Universal Access in the Information Society*, 14 (1), pp. 81-95.

MartinEdic (2019): "Why Marketers Need a Basic Understanding of UI and UX", Medium. Disponible en medium.com/better-marketing/the-importance-of-ui-ux-for-marketers-and-business-writers-6b2459fa31fa, [consulta: 12/08/2020].

McLean, G. y Osei-Frimpong, K. (2019): "Hey Alexa...examine the variables influencing the use of artificial intelligent in-home voice assistants". *Computers in Human Behavior*, 99, pp. 32.

MoneyDJ (2020): "Las ventas de altavoz inteligente crecen un 40% respecto al año al mismo trimestre del año anterior, Amazon, Google ocupan los 2 primeros puestos". TechNews. Disponible en <https://ccc.technews.tw/2020/02/14/smart-speaker-shipments-2019-q4/>, [consulta: 20/07/2020].

Rauschnabel, P. A., He, J., y Ro, Y. K. (2018): "Antecedents to the adoption of augmented reality smart glasses: A closer look at privacy risks". *Journal of Business Research*, 92, pp. 374–384,

ReasonWhy (2019): "Las cifras detrás de la revolución de asistentes de voz y altavoces inteligentes". Disponible en. <https://www.reasonwhy.es/actualidad/numeros-asistentes-voz-altavoces-inteligentes>, [consulta: 18/08/2020].

Sharoon Emmanuel (2019): "¿Por qué la voz es el futuro del marketing digital?". Medium. Disponible en <https://planetachatbot.com/por-que-voz-futuro-marketing-digital-2b5b5336a8ea>, [consulta: 20/07/2020].

Venkatesh, V., Morris, M. G., Davis, G. B. y Davis, F. D. (2003): "User Acceptance of Information Technology: Toward a Unified View". *MIS Quarterly*, 27 (3), pp. 425-478.

Vinney, C. (2019): "What is Uses and Gratifications Theory? Definition and Examples". ThoughtCo. Disponible en <https://www.thoughtco.com/uses-and-gratifications-theory-4628333>, [consulta: 13/08/2020].

Wu, Y. (2019): "Oportunidad comercial de la tecnología de voz, 3 tendencias que impactan a los marketers". Business Next. Disponible en <https://www.bnext.com.tw/article/54611/voice-control>, [consulta: 12/08/2020].

8. ANEXOS

Anexo I. Cuestionario completo Cuestionario en español

Tecnología de voz (Introducción)

El presente cuestionario será utilizado para completar un trabajo fin de grado. Los datos obtenidos se destinarán exclusivamente para fines académicos y se garantizará la protección a la privacidad de los datos personales de los encuestados.

La tecnología de voz es aquella que permite al usuario interactuar con la máquina a través de la voz. Ej: Siri, dictado por voz, búsqueda por voz etc. El funcionamiento de la máquina está basado en la Inteligencia Artificial, ya que la máquina tiene que aprender a identificar y entender el lenguaje humano y luego ofrecer una respuesta adecuada. Para ello, la máquina necesita entrenarse con muchos datos que serían etiquetados por el hombre.

Se dicen que los controles por voz están ganando cada vez más peso en nuestra vida, ¿pero realmente es así? Por lo menos el propio investigador no ha percibido el aumento de uso de este tipo de tecnología, y se quiere acercar a la realidad de su uso y explicarla.

Tecnología de voz (Introducción)

*Obligatorio

Uso de la tecnología de voz

⚠ Importante leer esta parte para poder contestar ⚠

Se agrupan las acciones que pueden realizarse a través de la tecnología de voz en 4 tipos de uso:

- 🔊 1. Dar órdenes: Realizar el control del dispositivo (sustituyendo los "clicks" por la voz)
Ej: Encender/apagar la luz, control de electrodoméstico, poner alarma/música, hacer llamadas, apuntar eventos en el calendario, escribir textos a través de la voz...
Ojo: La tecnología de voz no incluye mandar audios en el whatsapp, pero sí incluye los mensajes escritos a través del dictado pues la tecnología debe convertir el audio en texto.
- 🔊 2. Buscar informaciones: Acceder a informaciones
Ej: Buscar información de tiempo/Clima/localización/ resultados de los partidos, búsqueda de canción/película, búsquedas generales sobre cualquier tema...
- 🔊 3. Dejarse aconsejar: Dejar al asistente de voz que me recomiende algo, incluso darle el poder de decisión
Ej: Pedirle que me recomiende un restaurante, pedirle que elija una película, pedirle que elija la ropa que tengo que poner...
- 🔊 4. Entretenimiento: Pasar un buen rato mediante la interacción con el asistente de voz
Ej: Conversación con el asistente, Juegos de voz, Pedirle que imita la voz de personajes famosos, juguetes...

1. Utilizar la tecnología de voz para... *

	1 Muy Mala idea	2	3	4 Ni bien ni mal	5	6	7 Muy buena idea
Dar órdenes me parece...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Buscar informaciones me parece...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dejarme aconsejar me parece...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entretenerme me parece...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. ¿Has utilizado en alguna ocasión la tecnología de voz para los usos descritos? *

- Sí
 No

3. ¿Y para otros usos no descritos? *

- Sí
 No

4. Si lo has utilizado para otros usos ¿Cuáles son?

Tu respuesta _____

5. ¿Cuándo utilizaste por primera vez la tecnología de voz?

- En los últimos 6 meses
 Entre 6 meses y 12 meses
 Entre 1 año y 2 años
 Más de 2 años

6. Utilizo la tecnología de voz para dar órdenes *

1 2 3 4 5 6 7

Nunca De forma muy habitual

7. Utilizo la tecnología de voz para buscar informaciones *

1 2 3 4 5 6 7

Nunca De forma muy habitual

8. Utilizo la tecnología de voz para dejarme aconsejar *

1 2 3 4 5 6 7

Nunca De forma muy habitual

9. Utilizo la tecnología de voz para entretenerme *

1 2 3 4 5 6 7

Nunca De forma muy habitual

10. Me gustaría seguir utilizando la tecnología de voz o empezar a utilizarla para: *

	1 Seguramente que no	2	3	4 No lo tengo claro	5	6	7 Seguramente que sí
Dar órdenes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Buscar informaciones	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dejarme aconsejar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Entretenerme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Indica en qué dispositivos has utilizado la tecnología de voz (Permite múltiples respuestas. Si no la utiliza, salte esta pregunta)

- Sistema inteligente de coche
- Smartphone
- Ordenador
- Tele
- Altavoz inteligente
- Smartwatch o dispositivos llevables
- Tablet
- Electrodomésticos inteligentes
- Otro: _____

Características de la tecnología de voz

Se pide al encuestado que responda según su percepción (si no es usuario) o según su experiencia (si es usuario) sobre los posibles beneficios de la tecnología de la voz en varios aspectos.

12. Creo que muchas veces la tecnología de voz falla (La máquina no reconoce bien lo que digo, no entiende lo que digo, o entiende pero no tiene respuestas para lo que pregunto, entre otros posibles fallos) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

13. Creo que la inteligencia del asistente de voz está por debajo de mi expectativa *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

14. No me parece necesario el asistente de voz *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

15. La tecnología de voz es conveniente para cumplir muchas tareas diarias (Ahorra tiempo, Permite hacer varias cosas a la vez, realiza las tareas cuando no quiero moverme...) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

16. La tecnología de voz es fácil de utilizar *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

17. Creo que la tecnología de voz es útil *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

18. Para mí la tecnología de voz es "guay" *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

19. Es entretenido pasar tiempo con mi asistente de voz (Son graciosos los errores que producen y las respuestas humorísticas que inventan) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

20. Utilizar la tecnología de voz es divertido *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

21. Utilizar la tecnología de voz encaja con mi imagen personal *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

22. Utilizar la tecnología de voz me hace o me haría destacar en mi grupo de amigos y conocidos *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

Advertencia para las preguntas 23-26

Para evitar la influencia de que la tecnología todavía es imperfecta, se pide al encuestado imaginar un asistente de voz casi perfecto (como Jarvis en la película de Ironman) y responder cuestiones sobre "una posible relación entre máquina y hombre".

23. Creo que un asistente de voz así podría llegar a ser amigo mío *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

24. Creo que podría compartir mis asuntos personales o sentimientos con una máquina así *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

25. Cuando tuviera que hablar con el asistente de voz, lo haría como si fuera una persona real *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

26. Siempre tengo en cuenta que el asistente de voz es una máquina, una herramienta. Será difícil que desarrolle una relación social con ella. *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

Riesgo percibido sobre la privacidad

Se pide al encuestado que responda según su percepción (si no es usuario) o según su experiencia (si es usuario) sobre los posibles riesgos de la tecnología de la voz en varios aspectos.

27. Me preocupa que el asistente de voz guarde demasiadas informaciones mías *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

28. Tengo dudas sobre la confidencialidad de mi interacción con el asistente de voz *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

29. Creo que es muy inseguro comprar a través del asistente de voz (Requiere ofrecer datos personales, datos financieros...) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

Resistencia al cambio

En este bloque se pregunta por tu actitud hacia el hecho de hablar con la máquina: si no te importa, si te parece absurdo o si no tienes una opinión clara pero prefieres utilizar la interfaz de siempre (Como pantalla táctil y/o a través de teclado mecánico).

30. Creo que es raro hablar con la máquina *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

31. Pienso que es incómodo hablar con la máquina *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

32. Me sorprende ver a otros hablando con la máquina (Empezando con palabras clave como Hey, Siri o OK Google, por ejemplo) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

33. Me molesta si veo a otros hablando con la máquina *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

34. Un altavoz inteligente con pantalla es más atractivo que uno sin pantalla *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

35. Utilizar el asistente de voz no es mejor que emplear una pantalla táctil, ratón y/o botones (Dan los mismos resultados) *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

36. Me da vergüenza si otros me vean hablando con la máquina *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

Compatibilidad con el estilo de vida

En las siguientes preguntas se van a exponer ciertos escenarios considerados como ideales para la aplicación de la tecnología de voz. Se preguntará por la frecuencia que el encuestado siente o experimenta dichas situaciones.

37. ¿Con qué frecuencia experimentas las siguientes circunstancias? *

	1	2	3	4	5	6	7	
	Nunca me ha pasado						Muchas veces me ha pasado	No me ha pasado porque no conduzco, no cocino, no tengo móvil....
Tengo que poner el GPS cuando estoy conduciendo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hago llamadas o envío mensajes cuando estoy conduciendo y/o andando	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Me gusta escuchar algo cuando estoy haciendo tareas de casa

Utilizo el móvil cuando estoy cocinando (Ver receta, temporizador...)

Me da pereza levantarme del sofá/cama para apagar la luz

Me canso de escribir/teclear

Cuando necesito sentirme acompañado e interactuar, no sé con quién puedo hablar

Me cuesta hacer decisiones simples (Qué comer, qué película ver, qué marca comprar...)

38. Marca las costumbres que tengas de las que están a continuación (Permite seleccionar múltiples opciones) *

- Escuchar audiolibro
- Escuchar podcast o programas de radio
- Mandar mensaje de audio en whatsapp

39. Me gustaría reducir el tiempo que veo a las pantallas digitales *

	1	2	3	4	5	6	7	
Totalmente en desacuerdo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Totalmente de acuerdo

Características del encuestado

En este bloque se trata de conocer un poco las características personales de los encuestados con el fin de detectar sus diferencias y hacer posibles agrupaciones.

40. Me interesan las nuevas tecnologías y sigo informaciones sobre ellas *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

41. Por lo general, intento recabar tanta información como sea posible antes de comprar un nuevo producto tecnológico *

1 2 3 4 5 6 7

Totalmente en desacuerdo Totalmente de acuerdo

42. Sexo *

- Hombre
 Mujer

43. Edad *

- Menor de los 18 años
 De 18 a 24 años
 Entre 25-44 años
 Entre 45-64 años
 Mayor que los 65 años

44. Nivel de estudios *

- Secundaria o inferior
 Universitario
 Maestría
 Doctorado

Muchas gracias por su colaboración !!!

Questionario en Chino

簡介

此問卷是為了完成本人的大學畢業論文，收集到的資料只會用於研究而非其他用途，並且將會確保任何有關作者的資料不會遭到洩漏。

語音科技指的是任何可以讓人用聲音跟機器來互動的技術。像是 Siri, 語音輸入文字, 語音搜尋... 等等。此一技術建立在人工智慧的基礎之上，因為需要用大量語音數據來訓練機器，才能讓它學習如何分辨和理解人類的語言並做出適當的回應。

資料顯示，有越來越多的人開始使用語音科技，但真的是如此嗎？至少問卷的發行者並沒有查覺到這個現象。所以此問卷希望能夠認識中文語音科技真正的使用情況，並了解人們對它的態度。

語音科技的使用

▲為了讓答題能夠更加順暢，請務必閱讀這個部分▲

大致上，能夠用語音科技完成的事可以分成4大類：

◆1. 聲控：用聲音來控制裝置或應用程式（用聲音取代點擊滑鼠，觸控螢幕或是鍵盤）

Ex: 開/關燈，控制智慧家電，設定鬧鐘，播放音樂，打電話，紀錄活動到行事曆上，備忘錄，用聲音來打字...
注意：語音科技並不包含通訊軟體上中的語音訊息(Line, 微信)，因為這只算是錄音，機器並不需要理解使用者的話語。反之，使用語音輸入來打字回訊息的話就包含在語音科技的範圍裡。

◆2. 搜尋資訊：用聲音來獲取資訊

Ex: 查詢時間/天氣/地址/比賽結果/新聞...，搜尋歌曲/電影，語音搜尋任何事...

◆3. 尋求建議：讓語音助理來推薦，甚至幫使用者做決定

Ex: 讓語音助理推薦一家餐廳，叫它選擇一部電影，叫它幫我決定今天要穿甚麼...

◆4. 娛樂：從和語音助理的互動之中獲得樂趣

Ex: 跟它聊天，語音遊戲，叫它模仿動物的聲音...

1. 我覺得用語音科技來... (總共有7個選項) *

	1 非常不好的主意	2	3	4 不好不壞	5	6	7 非常好的主意
聲控 是...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
搜尋資訊 是...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
尋求建議 是...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
娛樂 是...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. 您有使用過語音科技嗎？ *

- 有
 沒有

3. 您有在其他領域使用過語音科技嗎？ (不包含開頭描述的4大類型) *

- 有
 沒有

4. 如果有在其他領域使用過，請問您用它來做甚麼事呢？

您的回答 _____

5. 您第一次使用語音科技是在多久以前呢？

- 最近6個月之內
 最近6-12個月之內
 最近1-2年之內
 2年或更久之前

6.我會使用語音科技來聲控 *

	1	2	3	4	5	6	7	
從不	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	常常

7.我會使用語音科技來搜尋資訊 *

	1	2	3	4	5	6	7	
從不	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	常常

8.我會使用語音科技來尋求建議 *

	1	2	3	4	5	6	7	
從不	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	常常

9.我會使用語音科技來娛樂 *

	1	2	3	4	5	6	7	
從不	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	常常

10.我會想要開始使用或是繼續使用語音科技來... (總共有7個選項) *

	1 絕對不會	2	3	4 不確定	5	6	7 絕對會
聲控	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
搜尋資訊	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
尋求建議	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
娛樂	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11.請問您在以下哪些裝置上使用過語音科技呢? (可複選, 如果並沒有使用過, 可跳過這個問題)

- 車上的智慧系統
- 智慧型手機
- 電腦
- 電視
- 智慧音箱
- 智慧手錶 或 可配戴式裝置(Ex: 手環)
- 平板電腦
- 其他: _____

語音科技的性質

請作答者依據自身的經驗(如果已經有在使用)或是 自身的印象(如果並沒有在使用)來回答

12.我覺得語音科技常常會出錯(Ex:它分辨不出我說的是甚麼,不懂我說的是甚麼意思,或是聽得懂但找不到相關的答案)*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

13.我覺得現有的語音技術並沒有達到我所預期的標準*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

14.我覺得根本不需要語音助理*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

15.語音科技可以讓人方便的完成許多日常生活上的任務(節省時間,讓人能同時處理多個任務,當人不想移動時完成一些簡單的小事 ex: 關燈)*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

16.使用語音科技非常的簡單*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

17.我覺得語音科技是有用的*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

18.我覺得語音科技很酷*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

19.跟語音助理互動的時候很有趣 (好笑的回答 或是 好笑的答非所問) *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

20.使用語音科技很好玩 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

21.使用語音科技符合我的個人形象 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

22.使用語音科技符合會使我在我的社交圈中顯得突出 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

提示: 當作答 23-26 題時

我們接著會詢問您對於: 人和語音助理之間的關係的看法, 為了避免科技不成熟的影響, 當作答 23 至 26 題時, 需要您設想一個幾乎完美的語音助理。(像是網織人之中的賈維斯)

23.我覺得一個像這樣的語音助理可能可以成為我的朋友 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

24.我覺得我有可能會向一個這樣的語音助理分享自己的私事或情緒 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

25.當我需要跟語音助理說話時, 我會用平時的講話方式跟他對話 (像是我在跟其他真人對話) *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

26.不管怎樣, 語音助理對我來說都會是個機器和工具。所以我想自己應該很難跟它發展出感情 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

對於隱私的擔憂

請作答者依據自身的經驗(如果已經有在使用)或是自身的印象(如果並沒有在使用)來回答

27.我會擔心我的語音助理儲存太多有關於我的資訊*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊同

28.我會懷疑跟語音助理對話的保密性*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊同

29.我覺得用語音助理來買東西很不安全(當需要提交個人或帳戶的資料時)*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊同

對語音科技的抵抗性

這一部份將會詢問您對於'對機器講話'這個行為的看法。你會在意嗎? 你會覺得愚蠢嗎? 或是你並沒有特別的看法但你現在寧願使用更熟悉的介面(看的到的介面, 像是觸控式螢幕或是電腦螢幕+鍵盤+滑鼠)

30.我覺得跟機器講話很奇怪*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

31.跟機器講話會讓我覺得不自在*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

32.看到其他人跟機器講話的話我會覺得驚訝(當有人用 Hey, Siri 或是 Ok, Google 來做開頭)*

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

33.當看到其他人跟機器講話時我會覺得困擾 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

34.一個有螢幕的智慧音箱會比沒有螢幕的更吸引我 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

35.使用語音科技並不會比使用 觸控螢幕/滑鼠/按鈕 要好多少 (因為兩者都能輸出相同的結果 Ex: 開啟/關閉 wi-fi) *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

36.假如我被別人看到我在跟機器說話, 我會覺得尷尬 *

	1	2	3	4	5	6	7	
完全反對	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	完全贊成

跟生活風格的契合度

我們將會描述一些公認較適合使用語音的場景, 並詢問您是否常常經歷這些場景。

您有常常經歷以下這些場景嗎? (總共有8個選項) *

	1 從來不會	2	3	4	5	6	7 常常會	從來不會, 因為我不開車, 或不煮飯, 或是我沒有手機...
我會在開車時使用GPS導航	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
我會在開車或走路時打電話或傳訊息	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
我會邊聽東西(音樂, 廣播...)邊做家事	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
我會邊用手機邊作菜(看食譜, 計時...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

我會懶得
從床上/
沙發上起
來開燈

我會懶得
寫字/打
字

當我覺得
需要陪伴
和交流
時, 不知
道可以找
誰說話

我會很難
決定一些
小事情
(要吃什
麼, 要看
什麼電
影, 要買
哪一個品
牌...)

38. 您有以下這些習慣嗎? (假如都沒有的話, 可以不答)

- 聽有聲書
- 聽podcast 或 廣播節目
- 用line或微信發送語音訊息

39. 我會想用減少看電子螢幕的時間 *

1 2 3 4 5 6 7

完全反對 完全贊成

作答者的個性

在這一環節我們想要認識您的一些基本資料, 目的是要檢視不同群體是否會對語音科技有不同的看法

40. 我對新的科技感興趣並會關注這一方面的消息 *

1 2 3 4 5 6 7

完全反對 完全贊成

41. 一般來說, 當我想要買一個科技產品, 我會盡量的收集有關的資訊(產品描述, 其他人的評價...)*

1 2 3 4 5 6 7

完全反對 完全贊成

42. 性別 *

- 男
- 女

42.性別 *

男

女

43.年齡 *

18歲以下

18-24歲

25-44歲

45-64歲

65歲以上

44.學歷 *

高中或以下

大學

碩士

博士

非常謝謝您的參予!!!

Anexo II. Análisis factorial

Factorial 1 para el bloque 1: Incluyen todos los indicadores de las variables de beneficios.

KMO=0,816, con cinco factores que explican el 73% de la información contenida en los indicadores de beneficios. En realidad, solo hay cuatro factores que tienen un autovalor mayor que uno, pero al sacar el quinto factor, se ve más claro el significado de los factores.

Componente	Varianza total explicada		
	Total	Autovalores iniciales	
		% de varianza	% acumulado
1	4,870	34,782	34,782
2	2,154	15,387	50,169
3	1,431	10,218	60,387
4	1,007	7,193	67,581
5	,888	6,343	73,924
6	,668	4,771	78,696
...			

Matriz de componente rotado^a

	Componente				
	1	2	3	4	5
La tecnología sirve para pasar tiempo	,842				
Es divertido utilizar la tecnología	,804				
La tecnología es guay	,696				
Es posible compartir con un asistente de voz		,851			
Asistente de voz puede ser un amigo		,813			
Se hablaría con el asistente de voz igual que cuando habla con una persona		,780			
Asistente como herramienta, difícil de desarrollar relaciones		-,576			
La tecnología es fácil de utilizar			,846		
La tecnología es conveniente			,728		
La tecnología es útil			,718		
La tecnología no cumple la expectativa				,868	
La tecnología tiene muchos fallos				,854	
Uso de la tecnología le haría destacar					,809
Uso de la tecnología encaja con su imagen					,656

Se ha pedido al programa que solo muestre las cargas de puntuación mayor que 0,5.

El nombre de los factores sería:

- Factor 1: Beneficio Hedónico
- Factor 2: Beneficio Social
- Factor 3: Utilidad
- Factor 4: Fallo Técnico
- Factor 5: Beneficio simbólico

Factorial 2 para el bloque 2: Incluye todos los indicadores de las variables de freno.

KMO=0,786, con dos factores, que tienen autovalor mayor que uno se explica un 70%.

Componente	Varianza Total Explicada		
	Total	Autovalores iniciales % de varianza	% acumulado
1	3,510	43,874	43,874
2	2,064	25,795	69,669
3	,597	7,460	77,129
...			

Matriz de componente rotado^a

	Componente	
	1	2
Hablar a la máquina es incómodo	,875	
Hablar a la máquina es raro	,868	
Viendo otros hablando con la máquina me sorprende	,815	
Viendo otros hablando con la máquina me molesta	,786	
Hablar con la máquina es vergonzoso si me vean otros	,750	
Riesgo de que el asistente guarde informaciones tuyas		,894
Riesgo de confidencialidad de las interacciones		,892
Riesgo de compra a través del asistente		,762

El nombre de los factores sería:

- Factor 1: Resistencia al cambio
- Factor 2: Riesgo percibido sobre la privacidad

Factorial 3 para el bloque 3: Incluyen los indicadores de la variable de compatibilidad con el estilo de vida.

KMO=0,811, con dos factores, que superan el autovalor de uno, se explica un 60%.

Componente	Varianza total explicada		
	Total	Autovalores iniciales	
		% de varianza	% acumulado
1	3,587	44,834	44,834
2	1,257	15,709	60,543
3	,836	10,455	70,998
'''			

Matriz de componente rotado^a

	Componente	
	1	2
Frecuencia de contactar con otros por el móvil cuando se está conduciendo/andando	,778	
Frecuencia de poner GPS cuando se está conduciendo	,768	
Frecuencia de escuchar cosas cuando se está haciendo tareas de casas	,733	
Frecuencia de utilizar móvil cuando se está cocinando	,686	
Frecuencia de costarse tomar decisiones simples		,815
Frecuencia de necesitar compañerismo		,779
Frecuencia de sentir pereza de teclear		,691
Frecuencia de sentir pereza de levantarse		,590

El nombre de los factores sería:

- Factor 1: Escenarios de multitarea
- Factor 2: Escenarios de Pereza