



**Universitat Autònoma
de Barcelona**

**TÍTULO: VOICE MARKETING: INFLUENCIA DE LAS VOCES EN LA
PERCEPCIÓN DE LA MARCA**

AUTORES: ANDONI HERNÁNDEZ Y TONI MOREY

MÁSTER UNIVERSITARIO EN MÁRKETING (MUM)

TUTORA: MARIA PILAR LÓPEZ BELBEZE

FECHA: 08/06/2020

Resumen

En este documento se habla sobre el “Voice marketing”, un nuevo sistema de creación de experiencias para el usuario todavía en fase de desarrollo el cual se basa en el uso de una voz artificial como sistema de comunicación con el cliente. Se estudiará el impacto sobre el consumidor de dichas voces en referencia a la marca y se tratará de averiguar los estándares que más gustan a la población. Para ello se recopilarán cuatro diferentes voces creadas mediante páginas web y dos voces presentes en el mercado (Google y Apple). Mediante un cuestionario se analizará la opinión de la gente y con el estudio de diversos análisis (ANOVA, PCA y word clouds) se interpretarán los resultados finales, donde se verá que una voz corporativa puede tener efecto en la visión de los clientes sobre la marca y que una voz que consiga parecer humana, captará con mayor eficacia la atención y la confianza que los clientes depositan en la marca.

ÍNDICE

1. Introducción	3
2. Marco teórico	3
2.1 Identidad de marca	3
2.2 Las experiencias de los consumidores.....	5
2.3 Marketing sensorial	7
2.4 Valor de marca	9
2.5 Características físicas de la voz	11
2.6 Objetivos	13
3. Metodología	14
3.1 Elección de los estímulos: las voces	14
3.2 Elaboración del cuestionario.....	17
4. Resultados.....	21
4.1 Análisis estadísticos realizados.....	21
4.1.2. Análisis bivariantes: ANOVA y Correlaciones	22
4.1.3. Análisis multivariantes: ACP	23
4.2. Rendimiento cognitivo de las voces	24
4.2.1 Pensamiento/asociaciones espontáneas	24
4.2.2 Características de las voces (resultados ANOVA).....	28
4.2.3. Matriz de correlaciones.....	30
4.2.4. Mapa perceptual de las voces.....	31
4.3. Rendimiento emocional de las voces	33
4.3.1. Valor percibido de las voces	33
4.3.2. Grado de confianza de las voces	34
4.4 Rendimiento comportamental de las voces	35
4.4.1 Intención de configuración.....	35
5. Interpretación y conclusiones	37
5.1 Extrapolación al ámbito del marketing.....	40
6. Bibliografía	41
7. Anexo	43
7.1 Anexo 1: Cuestionario	43
7.2 Anexo 2: Valores correlaciones	45
7.3 Anexo 3: Mapa perceptual.....	46

1. Introducción

Internet ha cambiado muchos aspectos en las personas, en los consumidores y estamos ahora ante la aparición e introducción de los asistentes de voz. El modelo de TAM (modelo de aceptación de la tecnología) de Davis et al (1989) nos ayuda a introducir este trabajo. Este modelo trata de explicar cómo los consumidores o el mercado aceptamos las nuevas tecnologías. Según este modelo, hay factores externos como la facilidad de uso de la nueva tecnología o el valor diferencial que aporta una nueva tecnología favorecen la adopción de una nueva tecnología, pero este trabajo se centra más en factores más internos como puede ser la actitud o la predisposición hacia la nueva tecnología por parte del consumidor o del mercado. La aceptación de una nueva tecnología como son los asistentes de voz estará determinada por uno de sus principal o elemento clave, la voz. En este estudio estudiaremos las voces y su efecto en las percepciones del individuo.

Por otro lado, las marcas, como en cualquier contexto, pero en especial en las nuevas oportunidades como puede representar el de los asistentes de voz, están buscando desarrollar marcas con valor. Las marcas utilizan signos, símbolos o una combinación de ellas con las que darse a conocer y deleitar a sus consumidores. La búsqueda de crear experiencias de marca también los ha llevado en algunos casos a buscar en singles o músicas unos signos de identidad con el que atraer a los consumidores. La voz puede ser el siguiente signo identitario de las marcas. En este trabajo también estudiaremos cómo las voces pueden ser diferentes en su notoriedad, nivel de recuerdo o asociaciones y que perfectamente podrían ser trasladadas a una marca.

2. Marco teórico

2.1 Identidad de marca

Como se ha mencionado anteriormente, las marcas deben buscar métodos para diferenciarse y no limitarse a ser una más, estamos hablando de crear una identidad de marca, construir una cultura impulsada por la marca es un compromiso de por vida con una mentalidad y una forma de vida que requiere tiempo, planificación y perseverancia. Esta cultura produce un tipo de intangibles que incluyen mayor satisfacción del cliente,

menor sensibilidad al precio, menos abandono de los clientes, una mayor *share of wallet* de los clientes y un mayor porcentaje de compras repetidas (Knapp, 2000). Los clientes valoran sus relaciones con sus posesiones de marca y con los agentes de marketing e instituciones que poseen y administran la marca (Alexander et al., 2002). Todo esto hace que se cree una identidad de marca la cual debe centrarse en puntos de diferenciación que ofrezcan una ventaja competitiva sostenible para la firma.

Para poder forjar una identidad de marca, la empresa debe basarse en conocer profundamente a sus clientes, competidores y el entorno empresarial en el que se encuentra. La identidad de la marca se debe reflejar en la estrategia comercial y en la voluntad de la empresa de invertir en los programas necesarios para cumplir con las promesas a sus clientes (Aaker y Joachimsthaler, 2000). Las empresas necesitan tener un amplio conocimiento sobre las creencias, comportamientos... de los consumidores y competidores. Las marcas fuertes disfrutan de la lealtad del cliente, el potencial de cobrar precios superiores y la posibilidad de lanzar nuevos productos y servicios de manera más fácil. Es por ello, que todas las marcas buscan crear una identidad propia para poder ser fácilmente reconocidas y puedan convertirse en la preferencia de los clientes frente a otro tipo de marcas.

Las preferencias de marca se han explicado durante mucho tiempo utilizando modelos tradicionales que se centran en el juicio cognitivo de los consumidores sobre los atributos de la marca de manera racional. Sin embargo, el cambio al marketing experiencial expandió el papel de la marca de un conjunto de atributos a experiencias. Esto hizo que también se considerasen los aspectos irracionales del comportamiento del consumidor (Schmitt, 1999, 2009). Además, los avances tecnológicos han aumentado las similitudes entre marcas y los diferentes productos. Por lo tanto, a los consumidores les resulta difícil diferenciar entre marcas solo por sus atributos funcionales (Petruzzellis, 2010). Es por eso que los consumidores buscan que la marca cree una experiencia; que consigan atraerlos de una manera sensorial, emocional y creativa. Tales apelaciones experimentales son importantes componentes de una marca (**Figura 1**), y se utilizan en la diferenciación de marca y mejora de preferencia del consumidor (Zarantonello y Schmitt, 2010). Por lo tanto, las empresas que compiten en dichos mercados enfrentan dificultades, ya que su supervivencia requiere construir una ventaja competitiva al intentar crear experiencias memorables (Schmitt, 1999, 2009).

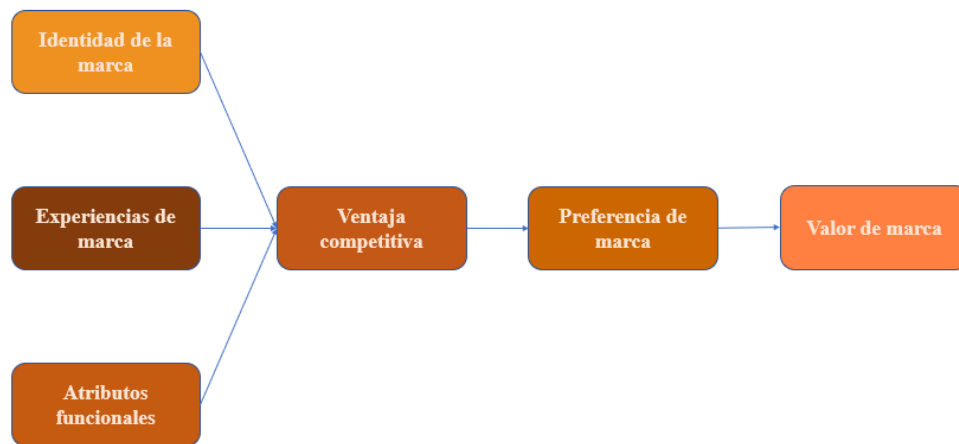


Figura 1 Mapa conceptual sobre el valor de marca

2.2 Las experiencias de los consumidores

El modelo antiguo de publicidad trataba de persuadir al cliente de una manera racional, llamar la atención mediante las características físicas y atributos funcionales de los productos tales como el precio o la calidad (Garcillán López-Rúa, M. ,2015). En la actualidad, ofrecer al mercado una propuesta de valor únicamente centrada en los atributos funcionales del producto es extremadamente complicado debido a la numerosa oferta existente de casi cualquier tipo de producto, por lo que la antigua vía de diferenciación ha quedado, sin dejar de ser relevante, obsoleta.

El concepto de experiencia del consumidor surgió a principios de la década de 1980 (Holbrook & Hirschman, 1982) para superar las limitaciones de las teorías del comportamiento tradicional del consumidor. Esta vista resalta la importancia de las variables descuidadas, considerando así a los consumidores como seres emocionales y pensadores (Addis y Holbrook, 2001). Investiga las respuestas de los consumidores a lo simbólico, estético e imaginativo, elevando el papel de los aspectos de la experiencia multisensorial (Holbrook y Hirschman, 1982). En consecuencia, esta vista expande y complementa el procesamiento de la información perspectiva, enriqueciéndolo con la perspectiva experiencial.

La experiencia de marca se define como el comportamiento subjetivo y conductual interno de los consumidores por las respuestas inducidas a diferentes niveles de interacción, tanto directos como indirectos, con estímulos relacionados con la marca (Meyer y Schwager, 2007). Los resultados de estas interacciones son las respuestas

experimentales almacenadas en la memoria del consumidor (Pine y Gilmore, 1998). La experiencia comienza antes de la compra real, continúa durante la compra y dura como una experiencia memorable (Tynan y McKechnie, 2009). La experiencia de marca captura lo sensorial, emocional, intelectual, conductual, social (Schmitt, 1999), aspectos pragmáticos y de estilo de vida de la marca. Las respuestas psicológicas, interpretaciones cognitivas y expresiones de comportamiento de las interacciones superpuestas mutuas con los insumos constituyen la experiencia del consumidor (Holbrook Y Hirschman, 1982). A través de estas experiencias interactivas, los consumidores construirán sus preferencias y decisiones de compra (Holbrook, 2007).

Las empresas buscan nuevas estrategias de diferenciación y, según la práctica empresarial actual, parece que aspectos relacionados con el contexto de compra que antes eran considerados secundarios e irrelevantes, están marcando la diferencia (Ballester, M. E. D., & Sabiote, E. F. (2011)). Las empresas captan la atención de los consumidores mediante “experiencias” (Carrillo, M. V., & Castillo, A. (2005)): el concepto de experiencia reside en la capacidad de aprendizaje del consumidor, que se adquiere con el uso y la práctica, que es la base de la fidelización. Con la creación de experiencias interactivas y que satisfacen las expectativas del usuario, se consigue establecer en la mente del consumidor un vínculo entre la marca y uno o varios refuerzos positivos derivados de la experiencia de la compra. Por lo que, cuando el consumidor se disponga a realizar el proceso de compra otra vez, existirá cierta probabilidad que desee repetir la experiencia anteriormente ofrecida procediendo a la compra del producto o servicio. Conseguir impactar en las emociones del consumidor de manera positiva es la base de los métodos de fidelización actuales.

Existen cinco tipos de experiencias que las empresas pueden establecer en sus procesos de compra para llegar a la satisfacción del consumidor (Schmitt, B. H. (2010)):

- Experiencias sensoriales: se basa en la percepción voluntaria e involuntaria del cliente, este método usa los 5 sentidos para acompañar la experiencia del consumidor, como ejemplo de dicha estrategia podemos encontrar: los colores de la marca, el logotipo, el perfume y la música de ambientación en las tiendas físicas entre otros ejemplos.
- Experiencias de sentimientos: Se trata de establecer un vínculo sentimental con el cliente, llevar a cabo acciones y experiencias durante el transcurso de la

relación, que aviven los sentimientos de los consumidores con el fin que lleguen a sentir sentimientos positivos hacia la marca.

- Experiencias del pensamiento: Se refiere a crear un desafío intelectual para el cliente, retar su manera de pensar, atraer su atención son temas relacionados con eventos o problemas actuales (contaminación, ecología, educación...), incluso llegar a plantar un nuevo modelo de consumo para generar interés en los consumidores.
- Experiencias de actuaciones: Se trata de crear un estilo de vida entorno a la marca, enriquecer la vida de los clientes con experiencias diferentes, nueva música, estilos de vida, decoración.
- Experiencias de relaciones: La finalidad de estas experiencias es conseguir que el cliente se sienta parte de un movimiento o tendencia, se tratan de experiencias colectivas fuertes que hacen sentir cierto orgullo al consumidor al formar parte de la marca.

Información extraída de: Casado, J. C. A., & Merino, M. J. (2011)

Como el principal objetivo de este estudio es analizar el impacto de la voz sobre la percepción de la marca de los clientes, se priorizarán las experiencias sensoriales por encima de las otras debido a que, el uso de una voz corporativa pretende estimular al cliente mediante el sentido auditivo.

2.3 Marketing sensorial

En el momento en el que se usan experiencias sensoriales para crear signos de identidad con los que atraer a los consumidores, pasamos a hablar de “marketing sensorial”.

El marketing sensorial usa herramientas de persuasión mediante los cinco sentidos (oído, gusto, tacto, olfato y vista), la idea principal del marketing sensorial es poder distinguirse de la competencia ante la saturación publicitaria a la que estamos sometidos, y llegar a conseguir que el cliente tenga una buena experiencia de compra para que en un futuro recuerde la marca de manera positiva (Garcillán López-Rúa, M. 2015). Para ello se debe conseguir transmitir emoción al cliente, si la emoción es positiva, funcionará como refuerzo positivo en el momento de la relación para que dicho cliente quiera repetir la experiencia de compra.

La manera en la que el marketing sensorial influye en el proceso de compra dista en gran parte del método tradicional de atracción de clientes, la principal diferencia es el uso del subconsciente. Cuando se usan técnicas que involucran el olfato, el oído o la vista, se busca provocar una reacción inconsciente sobre el cliente, las personas no respiramos de manera intencionada en busca de un olor determinado o no oímos solo lo que nos interesa, estos actos se llevan a cabo de manera inconsciente, no funciona así con el tacto o el gusto, que son dos sentidos que requieren de un deseo de participación del cliente para poder influir sobre ellos. El hecho de que sean inconscientes es lo que los hace más atractivos para las empresas, ¿Por qué una marca no genera más interés que otra solo con el logo? ¿Por qué el exterior de una tienda me llama más la atención que otra? ¿Por qué recuerdo la melodía de anuncios de hace meses?

La respuesta a estas preguntas consigue la diferenciación entre empresas semblantes en oferta, el hecho de estar en la mente del consumidor ya supone cierta ventaja respecto a las otras empresas. La manera de conseguir dicho efecto es mediante el uso de herramientas sensoriales:

Los consumidores no son tan racionales como nos ha enseñado la teoría económica, el proceso de toma de decisiones conlleva mucha incertidumbre y la decisión final suele ser tomada en un periodo de tiempo bastante limitado, según demuestran los estudios, un 70% de las decisiones se toma en el mismo punto de venta (Ramos-Serrano, M. (2007)). Este hecho es explicado por Goleman, D. (2001), el psicólogo argumenta que la mente humana está compuesta por dos mentes diferentes: la emocional y la racional.

La mente emocional es más rápida que la mente racional y nos permite tomar decisiones en los momentos críticos, en su contraparte la mente racional es más lenta pero las decisiones tomadas mediante dicha mente suelen ser más acertadas debido a que el número de variables que entra en juego a la hora de tomar una decisión es mayor. Hay que añadir que, aunque existan dos tipos diferentes de mentes, no implican que sean excluyentes entre sí, en un mismo proceso de compra, pueden llegar a interactuar las dos mentes, aunque la medida en la que actúan aun es un incierto y depende de cada situación.

Gracias a la interacción del marketing con la neurociencia, sabemos que, en algunas situaciones la elección entre marcas se produce por la influencia de cuestiones emocionales y culturales, Ramos-Serrano, M. (2007) ponen como ejemplo un estudio llevado a cabo por científicos de la Facultad de Medicina de Baylor en EEUU, donde a

los participantes que conocían las marcas que bebían se inclinaban por Coca-Cola y se les activaban partes del cerebro relacionados con cambios de conducta provocados por la emoción, en cambio, los que bebían sin conocer la marca, se les activaba el córtex prefrontal ventromedial, relacionado con los deseos básicos en los procesos de recompensa, que en este caso hace referencia al sabor, en este caso, los participantes se decantaban por la marca Pepsi.

Así, podemos definir 5 objetivos con los que el marketing sensorial puede ser relevante en la búsqueda de valor de la marca:

- Atraer y fidelizar clientes: tanto en los puntos de venta presenciales como online, se pueden aplicar técnicas para que el consumidor se sienta atraído por la marca y proporcionar una experiencia que quiera repetir.
- Crear una identidad de marca: permite a las marcas diferenciarse del resto consiguiendo que el consumidor la escoja o la prefiera por encima de las demás.
- Humanizar la marca: permite crear una experiencia más cercana, este objetivo tomará una posición relevante en el este trabajo ya que una de las finalidades del análisis de la voz es que el consumidor tenga la impresión de estar interactuando con un humano, y no con una inteligencia artificial.
- Poner valor a los productos: el hecho de asociar los sentimientos percibidos por una buena estrategia de marketing sensorial provocará que los consumidores identifiquen los productos de una manera positiva y le otorguen un mayor valor.
- Ofrecer una experiencia única: permite a la empresa ofrecer al consumidor una experiencia de compra agradable, con lo que el consumidor tendrá más probabilidades de repetir la compra.

información extraída de: <https://es.semrush.com/blog/marketing-sensorial/>

2.4 Valor de marca

Como ejemplo de la eficacia del marketing sensorial podemos analizar dos casos donde una buena estrategia ha conseguido que la empresa se diferencie de la competencia:

En primer lugar tenemos el caso de BMW, explicado por Vierna Carrasco, L. (2014), con el objetivo de reposicionar la empresa para recuperar el terreno que estaban

perdiendo contra Audi en el mercado de los coches de lujo, para ello llevan una campaña de remodelación del mensaje de la marca y empiezan a diferenciarse mediante “el placer de conducir”, mediante un cambio del slogan por una pregunta que, literalmente, incitaba al consumidor a responderla y mediante una serie de anuncios televisivos que abandonaban el modelo antiguo de publicidad y apostaban por unos anuncios más simples pero dirigiéndose al marketing visual y auditivo, con una combinación de paisajes bonitos y una música relajante acompañada con el viento característico de un viaje en coche (<https://youtu.be/bmYYpHI6clY>). Gracias a estas campañas se habla de una revolución del mundo de los anuncios televisivos, donde los estímulos no apelan a la capacidad racional del consumidor para decidir, sino a las propias emociones.

Otro ejemplo más claro sobre el marketing sensorial en el punto de venta sería el de Starbucks que, a pesar de que no suele invertir en publicidad, ha conseguido obtener éxito internacional haciendo que la marca se vea como un lujo accesible y consiguiendo un ambiente sociable en sus establecimientos (un lugar donde pensar, hablar y hasta trabajar), el precio de los servicios de Starbucks (más elevado que el su competencia), se justifica mediante esta promesa de valor experiencial donde no solo se compra un café sino que se está comprando un sentimiento (Casado, J. C. A., & Merino, M. J. 2011).

El marketing sensorial se convierte en una estrategia dominante en el ámbito publicitario actual, como hemos observado en los ejemplos, el uso de diferentes estímulos para atraer y estimular las emociones de los clientes puede ser la clave para la diferenciación de la marca, de hecho, “los seres humanos podemos recordar el 1% de lo que tocamos, el 2% de lo que escuchamos, el 5% de lo que vemos, el 15% de lo que degustamos y el 35% de lo que olemos” (Schmitt, B. 1999). Por eso no es de extrañar que gran parte de las tiendas tengan un perfume que impacte en el consumir justo al entrar, diferentes muestras de productos olfativos o gustativos a la disposición del consumidor, colores vivos que llaman la atención entre otras muchas estrategias similares.

Pero si nos fijamos en el contexto de compra actual, observamos que el volumen de negocio del e-commerce no ha parado de aumentar, y la previsión hacia el futuro es que lo siga haciendo. Un nuevo mundo de posibilidades de experiencias para los clientes se abre ante las marcas, sentidos como el olfato, el tacto y el gusto pasan a un segundo

plano debido a que su estimulación mediante servicios on-line es imposible con la tecnología actual, por lo tanto, los dos sentidos restantes (vista y oído) pasan a ganar severa importancia a la hora de definir experiencias para los consumidores.

A la hora de estimular la vista, aparecen términos como usabilidad y diseños web, que mejoran la experiencia del usuario con el simple retoque de la interfaz para que sea más agradable a la vista y fácil de usar, consiguiendo una mejora de la experiencia de compra del usuario al interactuar con la página web, no obstante, el sentido del oído es obviado, son pocas las marcas que usan una voz o sintonía para acompañar al consumidor en sus paseos por la página web consiguiendo así un trato más humano, esta tecnología ya está aplicada a los asistentes por voz de marcas como Apple, Google o Amazon, por tanto, ¿por qué no aplicar esta tecnología a más ámbitos web?

2.5 Características físicas de la voz

En este documento se pretende estudiar un estímulo auditivo como son las voces, el hecho de que una marca puede tener una voz corporativa y usarla en sus diferentes productos, analizar el efecto que puede llegar a tener sobre la visión de la marca por parte del consumidor. Para ello debemos conocer el instrumento humano que estamos tratando, así como definir ciertas características tanto intrínsecas como extrínsecas.

La voz se puede definir como el sonido producido por la vibración de las cuerdas vocales. Teniendo esto en cuenta, cada vez que se habla sobre las voces, en realidad se está hablando sobre diferentes sonidos.

El sonido se puede ver como el movimiento de ondas a través del aire o de otro medio elástico (en este caso el sonido es un estímulo) o también se puede ver como la excitación del mecanismo auditivo que hace que percibamos las ondas (en este caso el sonido es una sensación).

Hay dos tipos de ondas: las transversales y las longitudinales. En el caso de las ondas transversales, el desplazamiento de las partículas es perpendicular a la dirección de la propagación mientras que en el caso de las ondas longitudinales es paralela. El sonido está formado por ondas longitudinales que además son periódicas formando así varios ciclos consecutivos.

Una de las principales características del sonido, que luego servirá para poder definir o relacionar otras propiedades, es la velocidad. A través del aire y en unos valores temperatura y presión normales (20°C y 1atm) la velocidad del sonido es de alrededor de 1240km/h pero esto puede variar si cambian las condiciones; por ejemplo, un aumento en la temperatura, supone un aumento en la velocidad del sonido.

Otras de las características principales del sonido, y por tanto de la voz, son la longitud de onda y la frecuencia, la intensidad, el timbre y la duración (**Figura 2**).

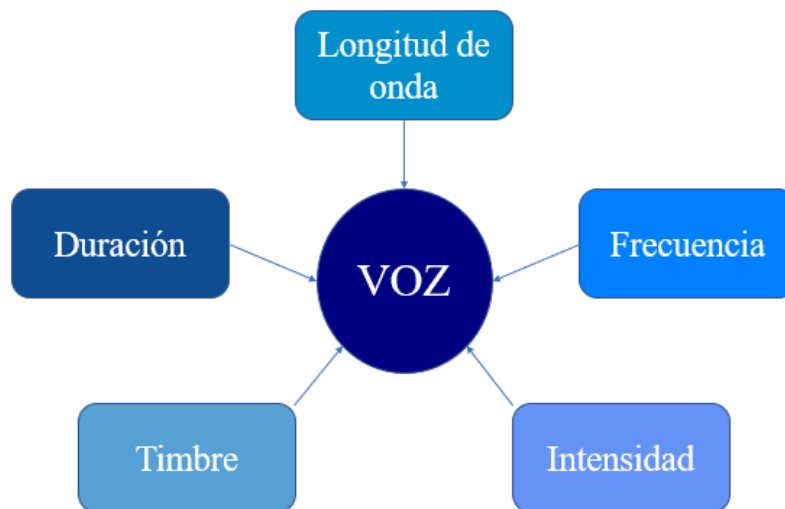


Figura 2 Características físicas de la voz

Por un lado, la longitud de onda es la distancia que recorre una onda hasta completar su ciclo y se puede medir teniendo en cuenta los picos de las ondas o midiendo dos puntos correspondientes en diferentes periodos. Por otro lado, la frecuencia hace referencia al número de ciclos que realizan las ondas por segundo y se mide en hercios (Hz). Para que un sonido pueda ser detectado por los humanos, debe tener una frecuencia comprendida entre los 20Hz y los 20.000Hz; por debajo de esa frecuencia se consideran infrasonidos y por encima, ultrasonidos. Para el caso de la voz, los hombres acostumbran a tener voces con frecuencias entre 100-200Hz mientras que el rango para las mujeres se encuentra en torno a 150-300Hz. Ambos parámetros (la longitud de onda y la frecuencia) están relacionados tal y como se puede ver en la **Ecuación 1**.

$$Frecuencia = \frac{Velocidad\ del\ sonido}{Longitud\ de\ onda} \quad (1)$$

Otra de las características fundamentales del sonido es la intensidad, parámetro medido habitualmente en decibelios (dB) y que depende de la amplitud de onda y de la fuerza con la que se ejecute el cuerpo sonoro. La intensidad de un sonido se puede medir utilizando un sonómetro y se definen como sonidos auditivos los que se encuentran entre los 0dB y los 130dB.

Los otros dos parámetros restantes son el timbre y la duración del sonido. Como su nombre bien indica, la duración del sonido hace referencia al tiempo que se mantienen las vibraciones producidas por el sonido. El timbre, depende de la complejidad de las ondas que forman el sonido; los sonidos que normalmente escuchamos son sonidos complejos formados por más de una onda simultánea (armónicos) aunque nosotros solo seamos capaces de detectar una. Por tanto, dependiendo de la cantidad de armónicos que compongan el sonido y la intensidad de cada uno de ellos, el timbre será uno u otro.

La voz es un elemento básico en la comunicación entre seres humanos, pero además de comunicar el mensaje, la voz puede comunicar otro tipo de datos. Por ejemplo, las variaciones en la voz permiten atribuir un estado emocional al hablante y, junto con eso, las emociones influyen en la sonoridad de la voz y los mensajes emitidos (Bänzinger y Scherer, 2005).

2.6 Objetivos

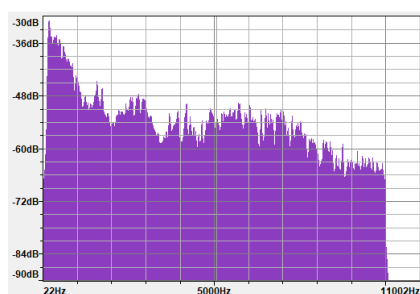
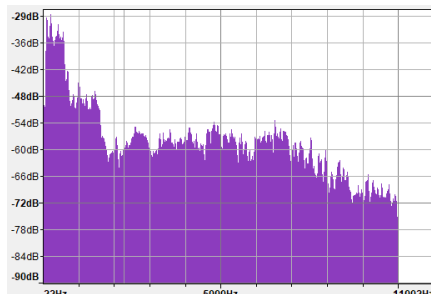
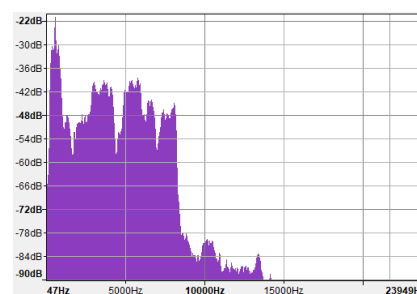
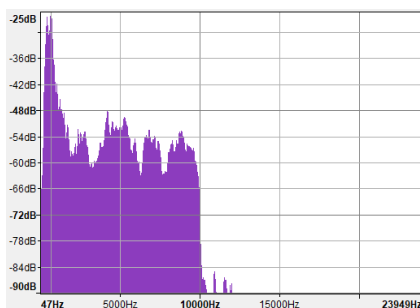
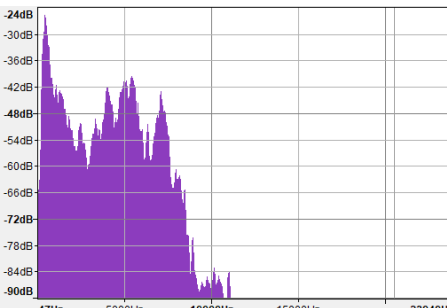
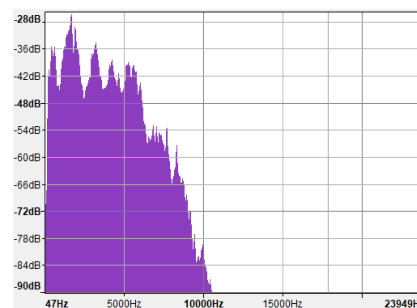
En este trabajo se pretende, mediante el análisis de seis voces diferentes, descubrir si existe relación entre los sentimientos generados por una voz corporativa y la visión que tiene el cliente de la marca. También se aprovechará para analizar cuales son las características que más gustan a la población en una voz; principalmente se tratará de descubrir si el género de la voz es relevante y en que medida puede llegar a afectar que una voz suene humana o robótica.

3. Metodología

Para dar respuesta a las preguntas de investigación planteadas se ha realizado una encuesta con la que se ha podido medir las reacciones de los participantes ante seis voces de asistentes virtuales. Gracias a esto, se ha realizado una valoración sobre la eficacia de las voces en términos de su recuerdo o actitud hacia las mismas. De las seis voces elegidas, dos de ellas corresponden a las voces de asistentes virtuales que se encuentran actualmente en el mercado: Siri y OK Google. Las cuatro voces restantes han sido creadas específicamente para esta investigación, siendo dos de ellas representativas de la voz de un hombre y las dos restantes de mujer. Para la recogida de datos se ha utilizado la plataforma CoolTool, una plataforma automática de investigación de mercados que permite una recolección de datos vía internet. A continuación, se describirá el diseño del estudio realizado donde se detallarán las voces seleccionadas, el diseño de la encuesta realizada y la descripción de la mi

3.1 Elección de los estímulos: las voces

Como ya se ha mencionado anteriormente se han utilizado seis voces (divididas en dos cuestionarios diferentes) para realizar este estudio. Todas las voces relatan la misma oración: **“Hola, soy tu asistente virtual, ¿en qué puedo ayudarle?”** Gracias a esto, se ha podido analizar cada una de las voces para obtener sus diferentes características mediante el programa de análisis de voces “Audacity”. Además, el hecho de utilizar una misma oración con todas las voces implica que el encuestado no se sienta influenciado por el contenido del mensaje, sino que simplemente pueda centrar toda la atención en la voz que está escuchando. En la **Figura 3** se representan los espectros de las diferentes voces analizadas en el cuestionario son su pertinente interpretación más adelante.

Voz 1**Voz 2****Voz 3****Voz 4****Voz 5****Voz 6****Figura 3** Espectros de frecuencia de las diferentes voces utilizadas

En el caso de la **voz 1**, en primer lugar, cabe destacar que se trata de una voz femenina. En cuanto a los aspectos puramente físicos, se puede ver como se trata de una voz bastante grave que ya los picos que más altura obtienen a lo largo del espectro se encuentran en las frecuencias más bajas. Además, escuchando el contenido, se puede apreciar un nivel de entonación en la voz. Por último, se ha de mencionar que esta voz ha sido obtenida en la página web www.animaker.es.

En referencia a la **voz 2**, se ha de mencionar que se trata de una voz masculina y con una entonación muy neutra. En lo que a aspectos físicos se refiere, se ve como al igual que en el caso anterior, se trata de una voz grave ya que en el espectro se acentúan mucho más las zonas de menos frecuencia. Como en el caso de la voz 1, esta voz también se ha obtenido gracias a www.animaker.es.

Refiriéndonos a la **voz 3**, cabe destacar que es una voz femenina y con una entonación muy monótona. En cuanto a sus características físicas, se observa como estamos frente a

una voz más aguda que las anteriores ya que, aunque se puede apreciar un pico pronunciado en las frecuencias bajas, los picos obtenidos en frecuencias más altas son destacables. Finalmente, se ha mencionado que esta voz 3 hace referencia a una de las voces corporativas de la marca Google

En cuanto a la **voz 4**, se trata de una voz femenina sin casi entonación. Si se observan sus cualidades físicas, se puede ver que esta voz es una voz más grave que la voz anterior y similar a la voz 1. Cabe destacar que la voz se ha obtenido en la web www.readspeaker.com.

En el caso de la **voz 5**, estamos ante una voz masculina con cierta entonación. Analizando el espectro de su frecuencia se puede observar que la voz es bastante grave ya que en las frecuencias más altas (las agudas) no destaca ningún pico ya que pierde completamente la altura. Al igual que en caso anterior, esta voz se ha obtenido en la web www.readspeaker.com.

Refiriéndonos a la **voz 6**, cabe destacar que es una voz femenina y con cierto grado de entonación. En cuanto a sus características físicas. En el espectro, se ve como estamos frente a una voz muy similar a la voz 3 mencionada anteriormente. Por último, se ha mencionado que esta voz 6 hace referencia a una de las voces corporativas de la marca Apple.

De cada una de las páginas web se ha obtenido una voz masculina y una voz femenina, generando así cuatro voces en total. Dada la posibilidad de poder escoger la frase formulada por cada voz, se ha escogido una misma frase para todas ellas:

“Hola soy su asistente virtual, ¿en qué puedo ayudarle?”

El hecho de elegir esta frase se debe a su neutralidad de contenido impidiendo así que el usuario sienta algún sentimiento derivado de ella más allá del generado por la propia voz. Por ejemplo, si la frase fuese “hola soy Siri (...)” el oyente podría parecerle una voz más humana solo por el hecho de tener un nombre y así ser más o menos preferida solo por la naturaleza del mensaje. Además, la frase incorpora una pregunta, este hecho permite poder diferenciar mejor a los oyentes cual voz tiene una entonación más humana o cual es más robótica puesto que la entonación en una pregunta es diferente a la entonación realizada en una afirmación corriente.

En la **Figura 4** se presenta un árbol conceptual con las características principales de cada voz



Figura 4 Características principales de cada voz

3.2 Elaboración del cuestionario

Para llevar a cabo la recogida de datos, se ha usado la plataforma Cooltool dada su funcionalidad al poder incluir un sonido entre las preguntas, de esta manera, los encuestados han podido contestar las encuestas sin necesidad de recurrir a terceros para oír las voces.

Con la intención de hacer los cuestionarios más amenos y para facilitar el trabajo de los encuestados, las seis voces se han dividido en dos cuestionarios diferentes (**Anexo 1**). Por un lado, en el primero de los cuestionarios se ha introducido la voz 1, voz 2 y voz 3 mientras que en el segundo cuestionario se han insertado la voz 4, voz 5 y voz 6. Para dividir las voces en los cuestionarios, se ha tomado en cuenta el criterio de incluir los mismos tipos de voces en ambos cuestionarios; es decir, ambos cuestionarios cuentan con una voz actualmente en el mercado, una voz femenina desconocida y una voz masculina desconocida (**Figura 5**).

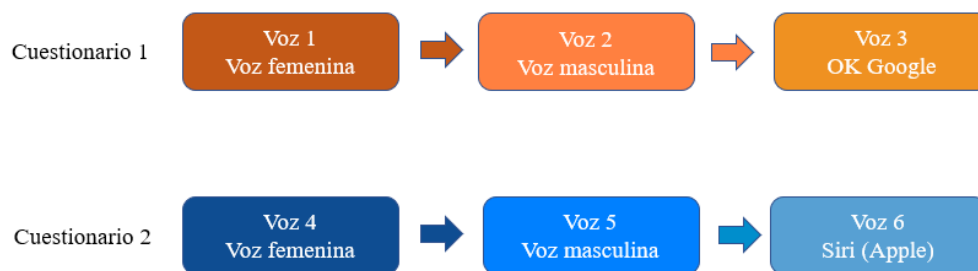


Figura 5 Disposición de las voces en los cuestionarios

La estructura de los dos cuestionarios es idéntica: empiezan con una pequeña introducción donde le explica al encuestado que procederá a escuchar tres voces diferentes y debe contestar las preguntas pertinentes, también se aclara que los datos obtenidos van a ser usados única y exclusivamente con carácter académico.

La primera pregunta se trata de una pregunta de carácter abierto donde se le proporciona una voz al encuestado. Después de escucharla debe responder con una palabra o expresión el sentimiento transmitido por dicha voz, gracias a esto se puede conocer el pensamiento o sentimiento principal que transmite la voz en el encuestado, también se le ofrece la posibilidad de añadir una respuesta adicional de manera opcional. Con esta pregunta se pretende descubrir la opinión del usuario sin tener un precedente de valoración de dicha voz, es decir, el usuario al responder esta pregunta aún no conoce los adjetivos definitorios que se aparecen en el cuestionario más adelante. De esta manera la respuesta no se ve condicionada por los adjetivos proporcionados por los investigadores y se obtiene una muestra más fiable.

En el próximo paso del cuestionario se exponen ocho características diferentes las cuales han de ser valoradas por los encuestados mediante una escala Likert 1-5. Los rasgos utilizados para este apartado están basados en el estudio de Soto Sanfiel, M. T. (2008) y son los siguientes:

- Nivel de agresividad
- Nivel de amabilidad
- Nivel de confianza
- Nivel de agrado
- Nivel de naturalidad

- Nivel de familiaridad
- Grado de credibilidad
- Nivel de artificialidad

Estos rasgos han sido usados con anterioridad para el análisis de las voces y en el cuestionario se pide a los participantes su percepción en las voces escuchadas. Gracias a esto, podemos ver el nivel o el valor otorgado por cada encuestado a cada una de las voces escuchadas y así poder comprender que rasgos caracterizan más a cada voz (**Tabla 1**).

Tabla 1 Medias y desviaciones estándares de las variables representativas de las características.

	Agresividad	Amabilidad	Confianza	Agrado	Naturalidad	Familiaridad	Credibilidad	Artificialidad
Media	1,731	2,721	2,469	2,395	2,019	1,974	2,298	3,688
Desv. Est.	1,005	1,236	1,236	1,25	1,164	1,151	1,226	1,401

A continuación, se realizan dos preguntas específicas para poder obtener datos concretos a cerca de la intención de los encuestados sobre una supuesta marca que pueda llegar a usar estas voces. Las preguntas realizadas son las siguientes:

- ¿Confiaría en una marca que use esta voz?
- ¿Configuraría esta voz como su asistente virtual?

Para la primera de las preguntas las respuestas posibles se limitaban a “Sí” y “No” mientras que, en el caso de la segunda de las preguntas, se amplía el rango de respuestas siendo estas “Seguro que sí”, “Probablemente sí”, “No lo sé”, “Probablemente no” y “Seguro que no”. Gracias a estas dos preguntas, se puede obtener información acerca del nivel de confianza transmitido por cada una de las voces y sobre la intención de configuración de cada uno de los encuestados hacia cada una de las voces

La estructura anteriormente mencionada sobre la pregunta abierta, los rasgos a valorar y las dos preguntas cerradas se repite para cada una de las tres voces y una vez que se haya respondido a estas preguntas para cada voz, se realizan varias preguntas de carácter selectivo, donde el usuario debe escoger la voz predominante en la

característica establecida. Así, se consigue obtener una comparación entre las diferentes voces, para poder analizarlas de manera conjunta, debido a que la otra parte del cuestionario las trataba de manera individual:

- ¿Qué voz te ha gustado más?
- ¿En cuál de las voces confiarías más?
- ¿Cuál de las voces te parece más artificial?

Con estas preguntas se pretende poder diferenciar las características que hacen que una voz sea preferida por delante de las otras en caso de que, anteriormente, hubieran obtenido puntuación similar en los adjetivos presentados.

Finalmente, y con el fin de conocer las características sociodemográficas de los encuestados, se pregunta acerca de su género, franja de edad y país de procedencia.

3.3 Obtención de la muestra

Se usaron diferentes vías de difusión de los cuestionarios, todas de forma on-line. El enlace del colector se distribuyó a través de redes sociales como Whatsapp, Facebook y Twitter. El periodo de recolección se llevó a cabo entre el 6 y el 14 de abril en los que se llegó a obtener un total de 74 respuestas en el cuestionario 1 y un total de 66 respuestas en el cuestionario 2, sumando así una muestra total de 140 respuestas (**Tabla 2**).

Tabla 2 Descripción de la muestra

		Cuestionario 1			Cuestionario 2		
Género	Femenino	69,33%			67,16%		
	Masculino	30,66%			32,83%		
	No contesta	0%			0%		
Edad	Menor de 18	2,66%			4,47%		
	Entre 18 y 34	77,33%			71,64%		
	Mayor de 34	20%			23,88%		
		Voz 1	Voz 2	Voz 3	Voz 4	Voz 5	Voz 6
¿Confiaría en una marca que use esta voz?	Si	42,66%	78,66%	28%	37,32%	58,20%	59,70%
	No	57,33%	21,33%	72%	62,68%	41,80%	40,30%
¿configurarías esta voz como tu asistente virtual?	Seguramente si	2,41%	4,90%	0,86%	2,85%	2,42%	20,45%
	Probablemente si	22,45%	56,25%	6,38%	19,07%	27,72%	32,59%
	No lo sé	27,78%	23,89%	20,18%	17,51%	27,84%	14,86%
	Probablemente no	29,87%	9,12%	39,35%	28,25%	22,27%	14,16%
	Seguramente no	17,48%	5,84%	33,23%	32,33%	19,75%	17,94%
Total de muestra		74	74	74	66	66	66

Según se desprende la **Tabla 2** es mayor la representatividad del género femenino en ambos proyectos. También entorno al 70% de los participantes a este estudio tienen una edad comprendida entre los 18 y los 34 años. Según esta muestra la Voz 2 es la que más confiarían y la configurarían un 56%.

4. Resultados.

De cara a dar respuesta a nuestras preguntas de investigación se ha dividido los resultados en dos partes. Por un lado, se realizará un análisis de los estadísticos realizada para dar respaldo teórico a la metodología usada. En la segunda parte se analizarán aquellos sentimientos surgidos de forma espontánea además de analizar la predisposición de los encuestados hacia las diferentes voces. Para ello, se utilizarán wordclouds para conocer esas asociaciones para cada una de las voces analizadas y se analizarán otras preguntas realizadas en el cuestionario para poder disponer de unos resultados más completos. También se analizarán los resultados de las características predefinidas por los investigadores mediante test de medias (ANOVA), posibles correlaciones entre los diferentes rasgos y un mapa perceptual elaborado en base a un análisis de componentes principales.

4.1 Análisis estadísticos realizados

4.1.1. Wordclouds

Esta sección va referida a las preguntas que dejan cierto margen de libertad al usuario para definir los sentimientos creados por las voces. Como se ha mencionado antes, la respuesta a esta pregunta no está condicionada por los adjetivos proporcionados por el mismo cuestionario, por lo que la respuesta ofrecida por el encuestado será una aproximación más fiable a la realidad que quiere reflejar con la palabra/expresión.

Para su interpretación se usarán nubes de palabras o wordclouds para ofrecer una mejor idea de la utilidad de las nubes de palabras, se podrían definir como *“una representación visual de palabras”*. *Cuanto más frecuente aparece la palabra dentro del texto que se analiza, más grande se vuelve la palabra. "En esencia, una nube de*

palabras "traza" la frecuencia de palabras por el tamaño de la palabra" (Ramsden, A & Bate, A 2008). Se ha elegido este método porque es una forma simple y visual de poder analizar la frecuencia de aparición de ciertas palabras debido a que se observó que las respuestas obtenidas bastante homogéneas en cuanto a opinión se refieren.

De los mismos autores de los que se obtuvo la definición, Ramsden, A & Bate, A (2008), se ha recolectado información sobre cómo obtener unas nubes de palabras fiables y evitar ciertos problemas que se pudieran ocasionar provocando una representación gráfica no representativa de la realidad.

- En una primera instancia, se debe asegurar que la muestra es significativa, que cuente con un cierto número de palabras para no dar una impresión que no refleja la realidad.
- Los datos obtenidos deben adaptarse para el software usado para la obtención de las nubes, si no se cumple esta norma, el gráfico puede obviar ciertas palabras interesantes para la interpretación.
- Por último, la muestra debe tener un lenguaje estandarizado y se deben realizar los cambios ortográficos pertinentes para que el software los detecte de forma correcta.
- También hay que tener en cuenta que las nubes de palabras no representan el contexto de la palabra, en este caso no es un factor relevante puesto que la muestra se construyó a base de palabras sueltas, no de frases.

Cumpliendo estos cuatro puntos, se modificó la muestra obtenida corrigiendo y estandarizando algunas palabras con el fin de obtener unas nubes lo más fiables posibles que serán expuestas e interpretadas más adelante.

4.1.2. Análisis bivariantes: ANOVA y Correlaciones

Para poder interpretar los resultados obtenidos de las valoraciones de los adjetivos facilitados mediante las escalas Likert, se ha realizado un análisis de la varianza (ANOVA (Fisher, 1992)).

En este caso se crearon ocho variables explicativas cada una representativa de la media de los adjetivos escogidos (agresividad, amabilidad, confianza, agrado, naturalidad,

familiaridad, credibilidad y artificialidad) obtenida mediante las valoraciones de los usuarios en una escala Likert del 1-5.

Luego se procederá a la creación de una variable índice (categórica) la cual represente de los valores 1-6 las distintas voces que forman la muestra.

Una vez obtenidas las variables necesarias para el análisis, se podrá observar la dispersión de las medias de los grupos creados. Para ello generaremos dos hipótesis:

- H_0 = Las medias de los grupos son iguales
- H_1 : Al menos una de las voces tiene una media diferente del resto.

Para ello será necesario comprobar el P-valor obtenido en cada variable saber si rechazamos o aceptamos la hipótesis nula, en este caso consideraremos una significatividad del 95% ($\alpha=0.05$), por lo que si se obtienen valores superiores a dicho número se entenderá que no existen pruebas suficientes para rechazar la hipótesis nula; por el contrario, si obtenemos un P-valor inferior a 0.5 se rechaza la hipótesis nula y se interpretará que, al menos un grupo, tiene una media que difiere de las demás.

Además, gracias a las salidas que nos ofrece el programa, se podrá observar las medias obtenidas en cada elemento de la variable índice, con ello se podrá determinar cuáles son los adjetivos que definen mejor las voces según las opiniones de los encuestados.

También se tendrá en cuenta la desviación estándar a modo de guía para comprobar la disparidad de las opiniones (a cuanta menor desviación estándar, mayor semejanza de opiniones).

4.1.3. Análisis multivariantes: ACP

El análisis de componentes principales (ACP) se puede definir como *“una técnica estadística que se basa en transformar linealmente una base de datos con numerosas variables cuantitativas en una base de datos con un menor número de variables cuantitativas, pero intentando representar el mayor número de información posible”* (George H. Dunteman 1989).

El objetivo principal es el reducir la dimensionalidad de la base de datos ya que es más fácil de entender una base de datos con pocas variables y no-correlacionadas que una base de datos con muchas variables y correlacionadas entre sí. Otro de los objetivos o ventajas que ofrece el análisis de componentes principales es la capacidad de

representación de los datos de una manera simplificada para poder sacar así información clara de forma sencilla. Las componentes principales se caracterizan por ser una combinación lineal de las variables originales además de no estar correlacionadas entre sí.

Para saber el número de componentes a elegir, es importante tener en cuenta que se busca representar los datos intentando explicar la máxima información posible. Para esto, se pueden encontrar dos métodos diferentes. Por un lado, existe el método de la varianza explicada, la cual se basa en elegir el número de componentes en base al porcentaje de variabilidad que se explica con cada uno de ellos (normalmente se busca alcanzar un 75% como mínimo). Por otro lado, está el método de los valores propios para el cual se observan los valores propios de los componentes y se elige tantas componentes como valores propios haya superiores a 1.

Una vez se ha realizado el ACP, los datos se pueden representar mediante un mapa perceptual en el cual se muestran tanto los datos como las variables que forman las componentes.

La idea de este análisis fue concebida por Pearson (1901) aunque después fue desarrollada independientemente por Hotelling (1933) (George H. Dunteman 1989).

4.2. Rendimiento cognitivo de las voces

4.2.1 Pensamiento/asociaciones espontáneas

En una primera instancia, se pidió a los encuestados que respondieran, con una sola palabra, el primer sentimiento o pensamiento que les venía a la cabeza al escuchar las voces con esto se pretende investigar y descubrir la primera reacción involuntaria a la voz. En esta pregunta, no entra en juego ningún adjetivo ni ejemplo por parte de los autores, por eso es la más genuina, parte directamente de la mente del consumidor sin ningún tipo de condicionamiento. Para poder analizar mejor los datos obtenidos, se han realizado diferentes *wordclouds*, uno para cada una de las voces, con ello se pretende ofrecer de manera visual un resumen donde se mostrará la frecuencia de aparición de las respuestas de los encuestados.

En la **Figura 6** se pueden observar los *wordclouds* de las voces 1, 2 y 3. Haciendo referencia a la Voz 1, se puede observar dos tipos de opiniones, en primer lugar, la opinión más generalizada es que la voz suena muy artificial, las palabras más mencionadas por los encuestados han sido “Robot”, “Robótica”, “Mecanizada” ... todas demostrando una misma sensación o connotación hacia la Voz 1. Además, aparte de las palabras mencionadas también destacan palabras como “Artificial”, “Amable”, “Seria”, “Correcta” o “Seca”; adjetivos que pueden ayudar a describir mejor la voz.

En cuanto a la nube de palabras formada con las opiniones realizadas sobre la Voz 2, se observa el predominio de palabras con connotación positiva como es el caso de “Natural”, “Amable”, “Confianza”, “Agradable” o “Cálida” aunque también aparecen palabras que describen un sentimiento contrario como pueden ser “Fría”, “Falso”, “Desganada” o “Agresiva”, aunque hay que añadir que la frecuencia de estos adjetivos es mucho más baja que los primeramente mencionados.

En el caso de la Voz 3, se observa un claro predominio de las palabras “Robot”, “Robótica”, “Artificial” o “Máquina” las cuales muestran claramente el sentimiento generado por esta voz. La voz número tres es de carácter especial puesto, junto con la de Apple, es una de las que ha sido sacada de una empresa existente, en este caso Google. Se observa que algunos encuestados han oído la voz anteriormente y lo expresan con palabras como “familiar” o “conocida”, algunos encuestados hasta han adivinado su procedencia.

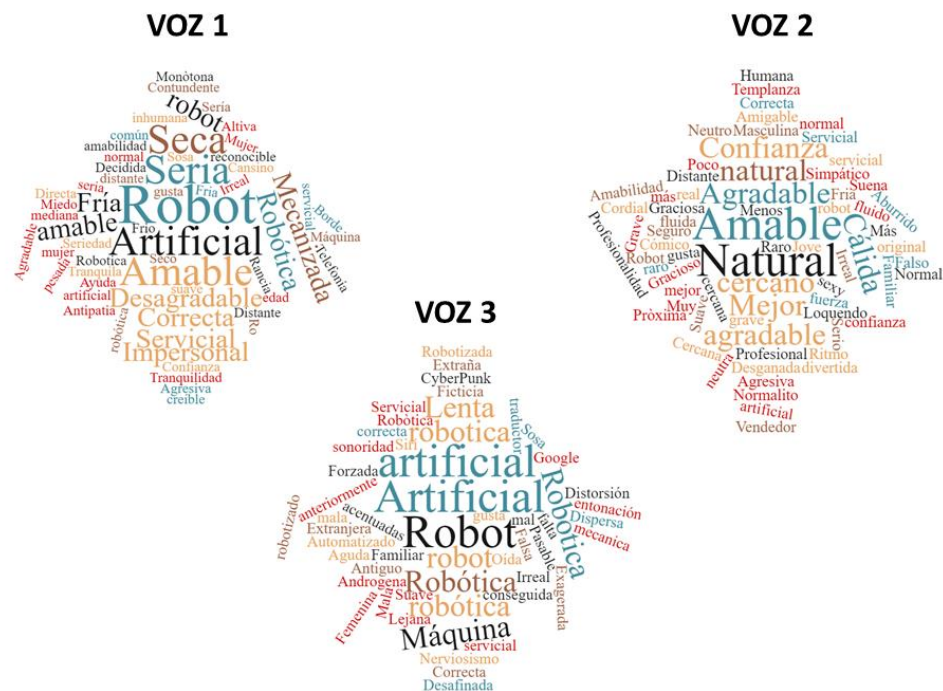


Figura 6 Wordclouds referentes a la Voz 1, Voz 2 y Voz 3 en los cuales se puede ver las primeras palabras con las que han reaccionado los encuestados al escuchar cada una de las voces.

Para analizar las sensaciones generadas por las voces 4, 5 y 6 también se han realizado diferentes *wordclouds* los cuales están representado en la **Figura 7**.

Como ya ha sucedido con otras voces, las palabras más destacadas acerca de la Voz 4 son “Robot”, “Robótica” o “Artificial” aunque frente a estas palabras también podemos observar palabras con connotaciones positivas como “Agradable”, “Confianza” o “Amable”, que aparecen con mayor frecuencia que palabras con connotaciones negativas.

La voz 5 es la segunda voz masculina analizada, esta, a diferencia de la primera, ha sido también definida como robótica, aunque con una menor frecuencia que en otras voces, los adjetivos en este caso son más diversos existen opiniones bastante dispares, algunos definen la voz como “amable”, “cálida”, “cordial” o “cercana” (siendo el sentimiento positivo la mayoría), en cambio también se observa una parte con un sentimiento

negativo hacia ella, definiéndola como “fría”, “desconfiada” y “falsa”. Aun así, el sentimiento positivo es el predominante.

La voz 6 en este caso es una voz supuestamente conocida para los encuestados puesto que es la que usa Apple en sus dispositivos móviles. La voz, más conocida como Siri consigue muy buenos resultados en cuanto a sentimiento, observamos que las palabras con mayor número de apariciones son adjetivos como: “Alegre”, “Dulce”, “Natural”, “Agradable”, “Simpática”. Esta voz, a diferencia de las otras femeninas presenta muy pocos adjetivos calificándola de robótica o artificial.

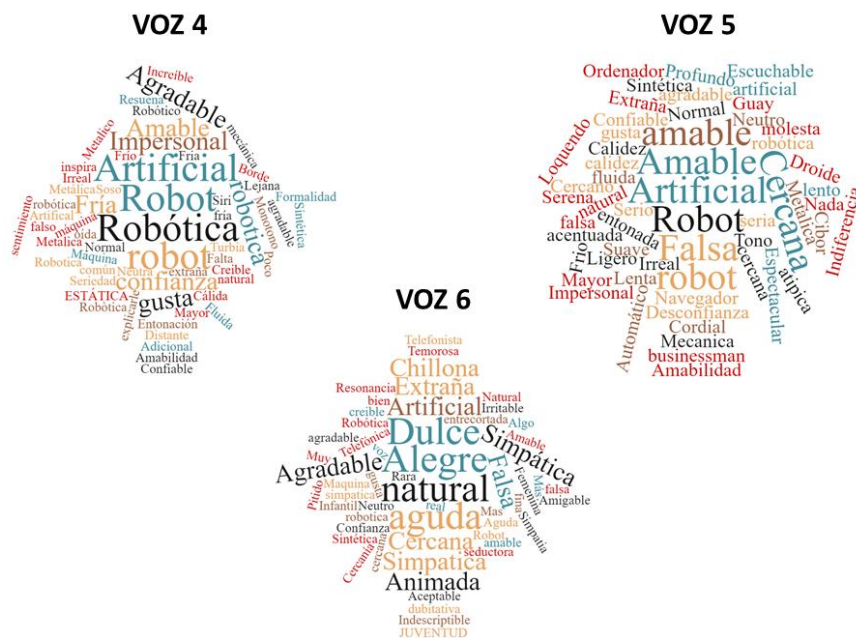


Figura 7 Wordclouds referentes a la Voz 4, Voz 5 y Voz 6 en los cuales se puede ver las primeras palabras con las que han reaccionado los encuestados al escuchar cada una de las voces.

Hay que notar también que, a diferencia de la voz de Google, la de Siri ha tenido poco recuerdo por parte de los encuestados, ningún adjetivo da pie a entender que el encuestado ha reconocido la voz, al contrario que pasa con Google.

4.2.2 Características de las voces (resultados ANOVA)

Con el fin de analizar las diferencias entre las medias de los grupos, se ha realizado un análisis ANOVA de cada una de las diferentes características extrínsecas con el fin de observar variaciones en ellas y poder demostrar diferencias entre las voces. Para ello se muestra en la **Tabla 3** un resumen de los resultados obtenidos. Se observa que los P-Value de todas las características son menores a 0,05, por lo que se interpreta que, al menos en cada adjetivo, hay una voz que difiere de las demás. (Tener en cuenta que la media está comprendida entre el 1-5, que son los valores otorgados al cuestionario). Se procede a analizar las medias de las diferentes voces respecto a los adjetivos especificados, para ello se adjuntará la **Tabla 3** donde se expondrán los resultados obtenidos mediante el análisis ANOVA.

Tabla 3 Resultados obtenidos mediante ANOVA

Agresividad	media	Desv. Est.	Test F (sign)	Familiaridad	media	Desv. Est.	Test F (sign)
voz 1	2,054	1,145	5.20 (0,00012)	voz 1	1,838	1,159	9,36 (1,80E-08)
voz 2	1,486	0,954		voz 2	2,432	1,251	
voz 3	1,892	1,200		voz 3	1,473	0,831	
voz 4	1,985	1,157		voz 4	1,682	1,055	
voz 5	1,515	0,808		voz 5	1,985	1,030	
voz 6	1,424	0,805		voz 6	2,455	1,205	
Promedio	1,731			Promedio	1,974		
Amabilidad	media	Desv. Est.	Test F (sign)	Credibilidad	media	Desv. Est.	Test F (sign)
Voz 1	2,581	1,170	22.09 (1.22E-09)	Voz 1	2,081	1,070	20,76 (1,59E-18)
voz 2	3,595	0,905		voz 2	3,149	1,143	
voz 3	1,986	0,972		voz 3	1,500	0,815	
voz 4	2,121	1,247		voz 4	1,939	1,094	
voz 5	2,924	1,181		voz 5	2,394	1,188	
voz 6	3,121	1,144		voz 6	2,742	1,281	
Promedio	2,721			Promedio	2,298		
Confianza	media	Desv. Est.	Test F (sign)	Artificialidad	media	Desv. Est.	Test F (sign)
Voz 1	2,324	1,172	15.76 (3,21E-04)	Voz 1	4,081	1,132	37,03 (2,31E-31)
Voz 2	3,203	1,072		Voz 2	2,338	1,197	
Voz 3	1,757	0,919		Voz 3	4,676	0,829	
Voz 4	2,076	1,256		Voz 4	4,227	1,187	
Voz 5	2,561	1,191		Voz 5	3,606	1,239	
Voz 6	2,909	1,224		Voz 6	3,197	1,406	
Promedio	2,136			Promedio	3,688		
Agrado	media	Desv. Est.	Test F (sign)	Naturalidad	media	Desv. Est.	Test F (sign)
Voz 1	2,189	1,155	21,89 (1,78E-09)	Voz 1	1,716	0,973	34,16 (3,22E-19)
voz 2	3,324	1,087		voz 2	3,027	1,146	
voz 3	1,622	0,753		voz 3	1,284	0,483	
voz 4	1,970	1,136		voz 4	1,439	0,844	
voz 5	2,424	1,229		voz 5	2,061	1,051	
voz 6	2,848	1,304		voz 6	2,591	1,252	
Promedio	2,395			Promedio	2,019		

En cuanto al sentimiento de **amabilidad** transmitido por las voces, se observa una clara diferencia entre la Voz 2 y las demás, teniendo esta una media de 3.59 con una desviación estándar bastante baja, por el contrario, la Voz 3 es la cualificada como menos amable por los encuestados presentando una media muy por debajo de las medias de los otros. Respecto a las otras voces, presentan valores bastante neutros tirando al alta en la voz 6 y 5, y a la baja en las voces 4 y 1.

La **confianza** transmitida sigue unos parámetros similares a la amabilidad donde, destaca la Voz 2 por encima de las demás de una manera positiva, mencionar también la Voz 6, que obtiene resultados por encima de la media del 2'5 pero no tan positivo como la Voz 2. Si nos fijamos en la voz que menos confianza transmite, la Voz 3 es la más destacada con una media de 1.75, muy baja en comparación con las otras voces.

Las voces que más **agradan** a los encuestados vuelven a ser las voces 2 y 6, destacando la Voz 2 con una media de 3.32, bastante más alta que el resto de las voces. Las que menos agradan en este caso son las voces 3 y 4 con medias por debajo del 2, destacando la Voz 3 con una media de 1.62.

En este caso se analizará cuál de las voces suena más **natural** al oído de los consumidores, es la voz número 2 la que obtiene la puntuación más, aunque tampoco sobresale demasiado, el resto de las voces obtienen puntuaciones bastante bajas (a excepción de la voz 6 que obtiene una puntuación neutra) destacando la Voz 3, 4 y 1 como las peor puntuadas.

En cuanto al grado de **familiaridad**, observamos que ninguna voz destaca por obtener buenas puntuaciones, solo las voces 2 y 6 han conseguido llegar al 2.4 de media, las otras voces tienen poco grado de familiaridad.

Respecto a la **credibilidad** las voces presentan valores neutros destacando la voz 2, 6 y 5, (3.14, 2.74, 2.39 respectivamente) que obtienen mejores puntuaciones que las otras voces, la Voz 3 vuelve a obtener una puntuación muy baja en esta característica (1.5)

En el caso de la **artificialidad** de las voces, todas, en este caso presentan valores elevados, algunas hasta rozando el 5, como la voz 3 (4.67) la Voz 1 (4.08) y la voz 4 (4.22), por otro lado, la Voz número 2 obtiene la media más baja de todas las voces, aunque sigue siendo elevada según la media de las medias anteriormente analizadas.

4.2.3. Matriz de correlaciones

Para poder realizar un buen análisis sobre los resultados obtenidos en el apartado de los adjetivos facilitados es importante observar en primer lugar el grado de correlación que se encuentra entre cada uno de los adjetivos utilizados. Para ello, en la **Figura 8** se puede observar un cuadro en el cual mediante diferentes colores se puede apreciar la correlación entre diferentes adjetivos. (En el **Anexo 2** se podrán consultar los valores exactos correspondientes a todas las correlaciones).

Para el caso del adjetivo “agradable”, esta característica muestra un grado de correlación positivo con un gran mayoría de las otras características a excepción de con “agresiva” y “artificial”. Con éstas se observa una correlación negativa, siendo éste bastante acenturado para el caso de “artificial”.

Analizando la característica “agresiva” se ve como tan solo muestra una correlación postiva con el adjetivo “artificial” mientras que la correlación hacia las demás características es negativa. Cabe destacar, que tanto en el caso de la correlación positiva como en el caso de la correlación negativa los valores obtenidos no son muy elevados.

En cuanto al adjetivo “amable”, se puede observar una correlación postiva y más elevada con los rasgos “agradable” y “confiable” aunque la correlación también es positiva para todas las demás variables a excepción de “agresiva” y “artificial” con las que mantiene una correlación negativa (y elevada para el caso concreto de “artificial”).

El rasgo “artificial” muestra una correlación negativa respecto a la mayoría de adjetivos a excepción de “agresiva” con la que muestra una correlación positiva aunque de un valor no muy elevado.

La característica “confianza” es una de las características que muestra correlaciones más elevadas positivamente con muchos de los adjetivos menos para “agresiva” y “artificial”.

Al igual que en el caso anterior, para el rasgo “creible” también se aprecian correlaciones positivas con todas las variables a excepción de “agresiva” y “artificial”.

El adjetivo “familiar” no muestra correlaciones elevadas con ninguno de los adjetivos anteriores aunque, tal y como se comprueba en otros adjetivos, la mayoría de correlaciones son positivas excepto “agresiva” y “artificial”.

Por último, para el rasgo “natural” se pueden ver correlaciones positivas y de valor elevado en casi todos los casos menos en los ya mencionados “agresiva” y “artificial” cuyas correlaciones con “natural” son negativas.

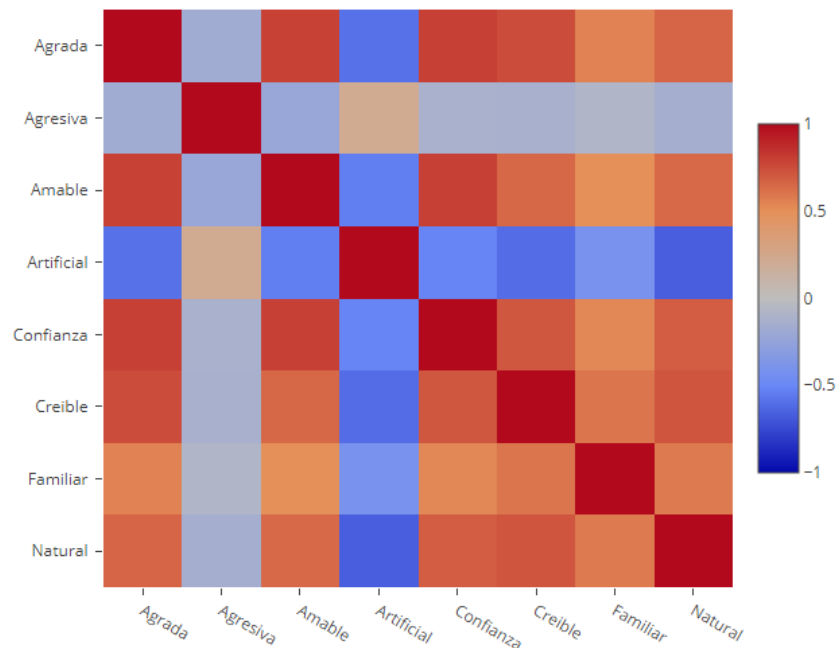


Figura 8 Cuadro de correlaciones entre los adjetivos utilizados para el estudio

4.2.4. Mapa perceptual de las voces

Finalmente, una vez observadas las medias y las correlaciones entre los distintos adjetivos se realiza un análisis de componentes principales.

En este caso, se utilizan dos componentes de los cuales el primero explica un 82,36% de la varianza mientras que el segundo componente explica el 12,97%. También cabe destacar que los adjetivos que más aportan a explicar el primero componente son “agradable”, “amable”, “artificial”, “creible” mientras que el que mejor define el segundo componente es “confianza” de una manera bastante clara respecto a los demás.

Observando el mapa perceptual (**Figura 9**) se puede ver como se ubica cada una de las voces dentro de los dos componentes utilizados (en el **Anexo 3** se puede encontrar un

mapa perceptual con los datos de cada una de las respuestas obtenidas en los cuestionarios).

En el caso de la voz 1 se puede ver como destaca por su agresividad y por su artificialidad. Para la voz 2 las características más destacadas son la amabilidad, la credibilidad y el agrado mientras que la voz 3 demuestra unas características totalmente contrarias destacando sobre todo por su artificialidad.

En el caso de la voz 4, al igual que en la voz 1 y la voz 3, destacan la artificialidad y la agresividad mientras que la voz 5 muestra unos rasgos similares a la voz 2 aunque de forma mucho más discreta y estando muy cerca del promedio. Para acabar, la voz 6 destaca principalmente por su familiaridad y su naturalidad dejando muy de lado atributos como la agresividad.

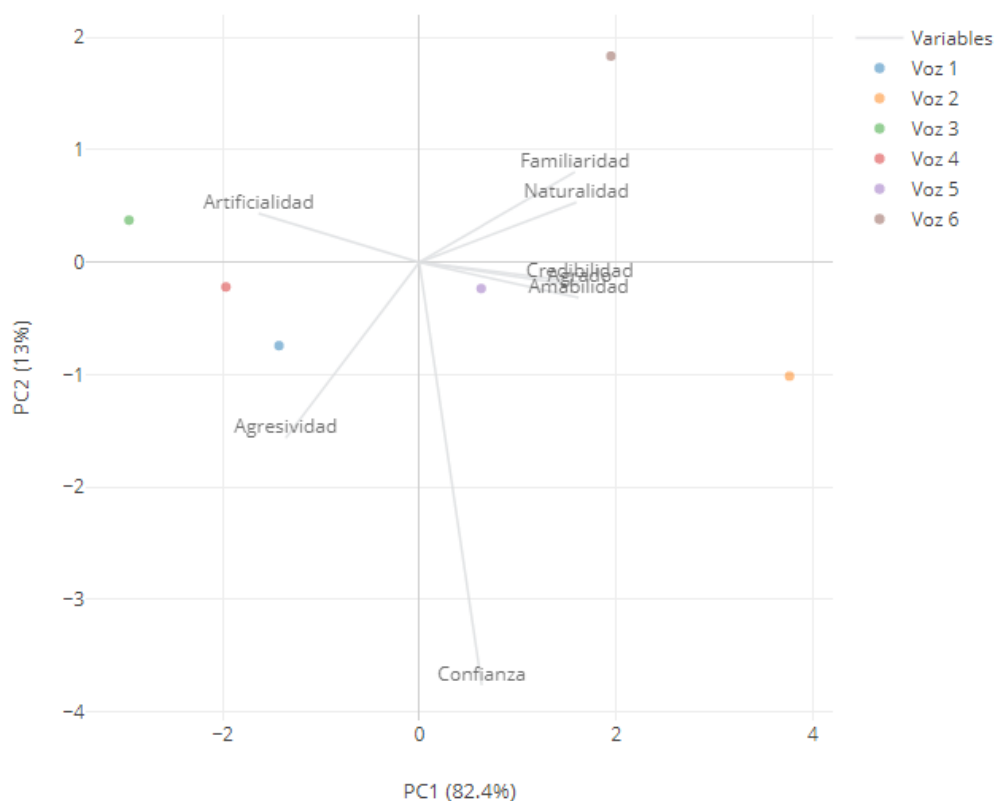


Figura 9 Mapa perceptual del análisis de componentes principales

4.3. Rendimiento emocional de las voces

4.3.1. Valor percibido de las voces

Una de las cuestiones preguntadas a los encuestados es la preferencia o la predilección hacia una de las voces. Para ello, se preguntaba acerca de cuál es la voz que más gustaba, la voz en la más podría confiar y la voz más artificial. En la **Tabla 3** se pueden apreciar los distintos datos obtenidos como resultados a las preguntas mencionadas.

Tabla 3 Porcentaje de respuesta sobre las preguntas realizadas acerca del valor añadido de las voces

	Gusto	Confianza	Artificial		Gusto	Confianza	Artificial
Voz 1	19%	16%	20%	Voz 4	18%	18%	59%
Voz 2	76%	81%	1%	Voz 5	27%	26%	20%
Voz 3	5%	3%	79%	Voz 6	55%	56%	20%
Tamaño de la muestra	74	74	74	Tamaño de la muestra	67	67	67

En lo que se refiere a cuál de las voces es la que más gusta entre los encuestados, se observa que en el primer cuestionario (Voz 1, 2 y 3) y con un gran 76% la voz preferida es la Voz 2. A continuación, aunque con un porcentaje muy inferior, se encuentra la Voz 1 con un 19% dejando a las Voz 3 en último lugar con un 5%. Observando los resultados obtenidos para esta misma pregunta en el segundo cuestionario se ve que, aunque no hay la misma superioridad que en el cuestionario anterior, la voz que más gusta es la Voz 6 con un 55% seguida de la Voz 5 con un 27%. Finalmente se encuentra la Voz 4 con un 18% de las respuestas.

Si centramos la atención en la segunda de las preguntas, la cual se centra en la confianza generada por las voces, se pueden observar los siguientes resultados. En el primer cuestionario la voz en la que más confían los encuestados es la Voz 2 con un 81% dejando muy por debajo a las demás (16% para la Voz 1 y 3% para la Voz 3), En cuanto al segundo cuestionario, la voz en la que más confiarían los participantes de la encuesta es la Voz 6 con un 56% seguida de la Voz 5 con un 26% y la Voz 4 con un 18%.

Por último, observando cuál de las voces consideran más artificial los encuestados, se aprecia que en el primer cuestionario la voz más artificial es la Voz 3 con un 79% de las respuestas seguida de la Voz 1 con un 20% y dejando en tan solo un 1% a la Voz 2. En cuanto al segundo cuestionario, la voz claramente detectada como más artificial es la

Voz 4 con un 60% de las respuestas mientras que para terminar tanto la Voz 5 como la Voz 6 reciben un 20% de las respuestas sobre la artificialidad.

4.3.2. Grado de confianza de las voces

Otro de los temas a analizar es la percepción de los consumidores de la marca dependiendo del uso de las diferentes voces analizadas en el trabajo. Para ello se mostrarán seis gráficos circulares (**Figura 10, Figura 11**) representativos del porcentaje de personas que, escuchada la voz pertinente, confiarían en la marca.

Para el primer set de voces (**Figura 10**) se observa que la voz 2 obtiene unos resultados positivos con mayor diferencia a las otras voces analizadas, se observa que la voz 1 y 3 obtienen calificaciones negativas en lo que a confianza en la marca respecta.

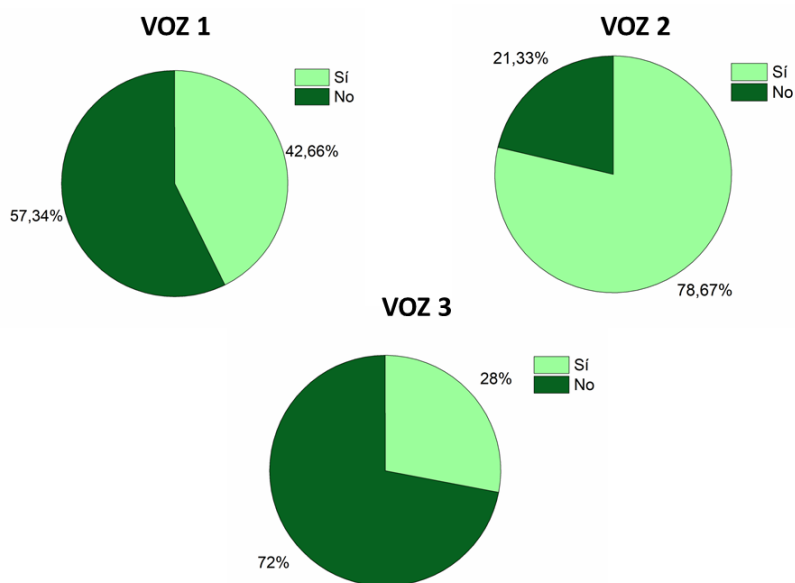


Figura 10 Gráficos circulares sobre la confianza otorgada por los encuestados a cada una de las voces representando el “sí” un nivel de confianza positivo y el “no” un nivel de confianza negativo.

El segundo set de voces (**Figura 11**) es un poco mas ambiguo, destaca la Voz 6, que consigue obtener buenos resultados en generar confianza sobre la marca, mientras que

las voces 4 y 5 consiguen una intencion de confianza bastante neutra, sin decantarse excesivamente hacia el sí o el no.

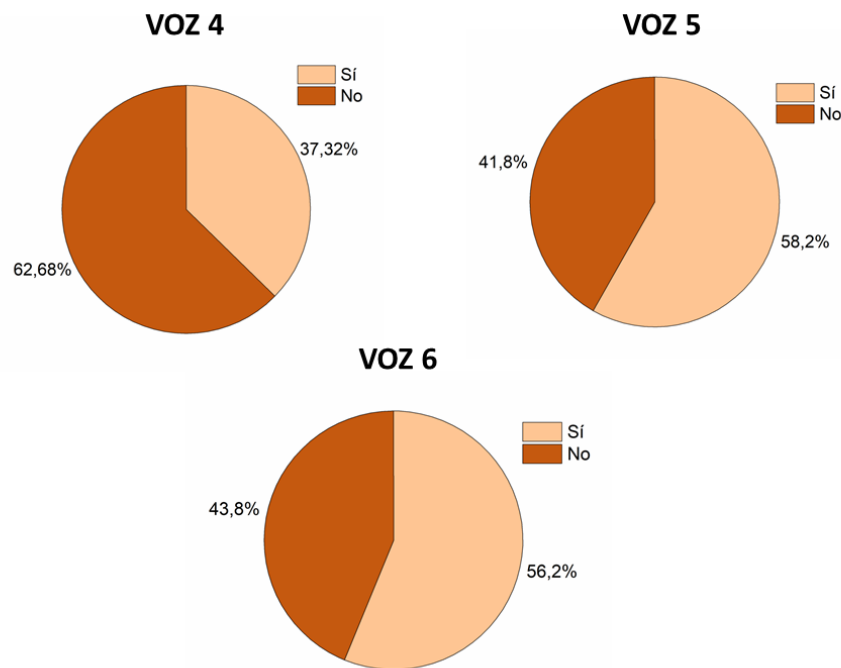


Figura 11 Gráficos circulares sobre la confianza otorgada por los encuestados a cada una de las voces representando el “sí” un nivel de confianza positivo y el “no” un nivel de confianza negativo.

4.4 Rendimiento comportamental de las voces

4.4.1 Intención de configuración

Además de estudiar la confianza que produce cada una de las voces, también se ha hecho una valoración sobre la intención de cada uno de los encuestados a configurar las voces como sus asistentes personales.

En la **Figura 12** se pueden ver los gráficos donde se indica la intención de configuración para la voz 1, voz 2 y voz 3. En referencia a la voz 1 se puede apreciar como una cantidad mínima de la gente indica estas completamente segura de configurar la voz mientras que las demás opiniones adquieren un número de respuestas similares.

En cuanto a la voz 2, la tendencia es claramente positiva y la gran mayoría de la muestra se decanta por que probablemente sí que llegaría a configurar esta voz como su propio asistente personal. Siguiendo con la voz 3, al contrario que en la voz anterior, la

tendencia es negativa ya que se aprecia un claro predominio de las respuestas “seguro que no” y “probablemente no”.

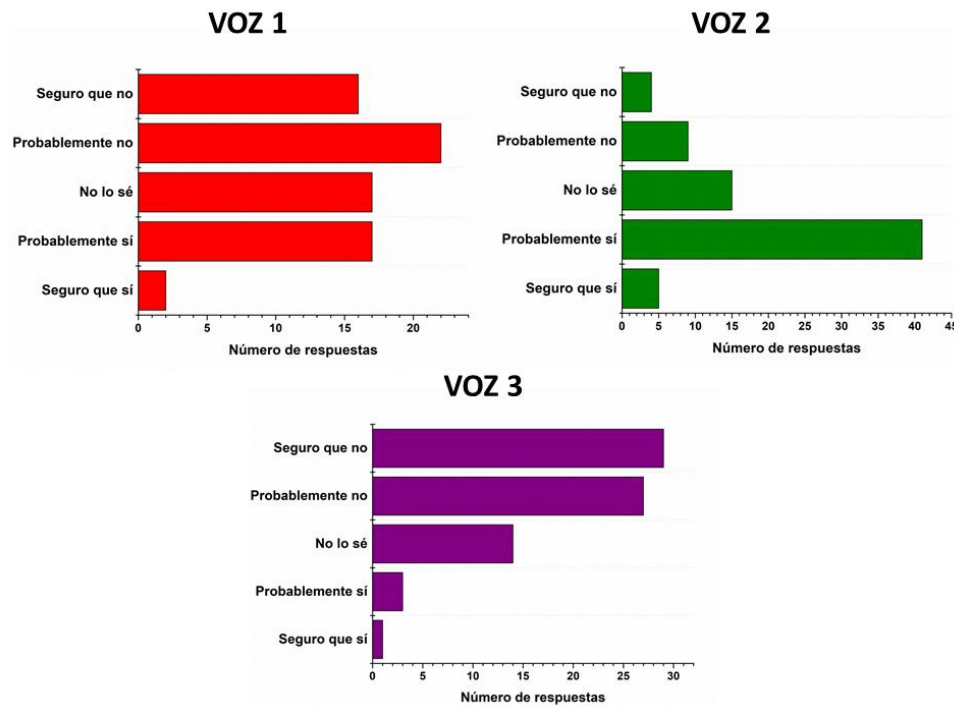


Figura 12 Gráficos sobre la intención de configuración de las personas encuestadas sobre la voz 1, voz 2 y voz 3.

En la **Figura 13** se puede observar la intención para la voz 4, voz 5 y voz 6. En el caso de la voz 4 se puede ver como una minoría de los encuestados se decanta por configurar esta voz como su asistente mientras que un gran número de ellos decidiría no configurar esta voz. En cuanto a la voz 5 se refiere, la tendencia se repite en cuanto las respuestas “seguro que sí” mientras que en esta ocasión la mayoría se decanta tanto por el “no lo sé” como por el “probablemente sí”.

Por último, observando las tendencias de la voz 6 se puede observar que hay una intención muy positiva de configurar esa voz como asistente personal por parte de los encuestados ya que las respuestas más numerosas son el “probablemente sí” y el “seguro que sí”.

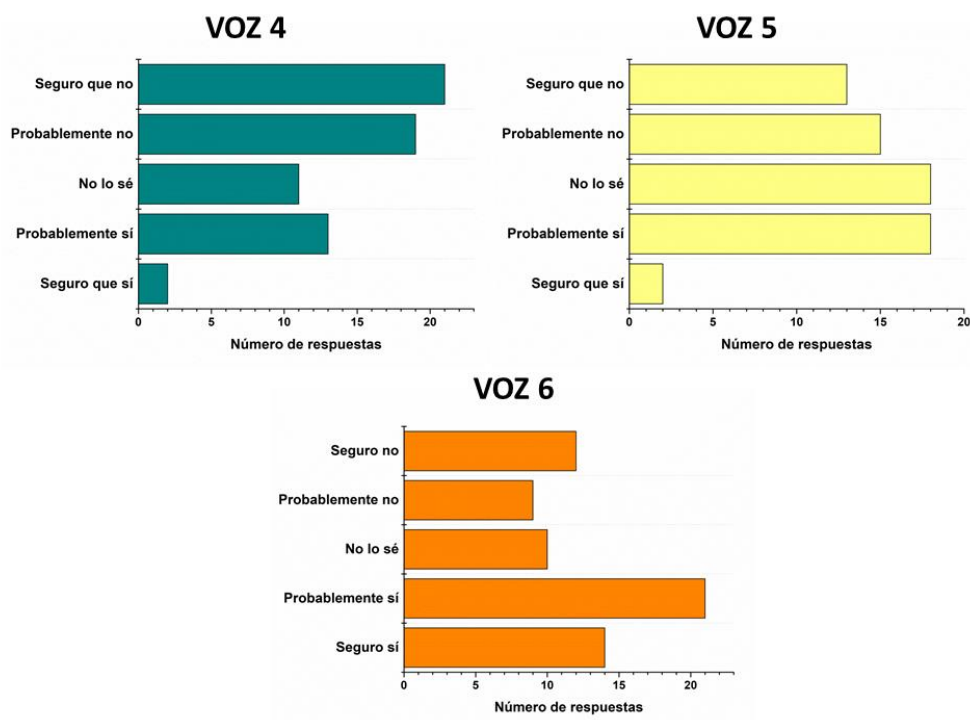


Figura 13 Gráficos sobre la intención de configuración de las personas encuestadas sobre la voz 4, voz 5 y voz 6.

5. Interpretación y conclusiones

Tras analizar los resultados obtenidos de los dos cuestionarios, se han realizado diferentes interpretaciones sobre los mismos.

En cuanto al cuestionario número uno (Voz, 1, 2, 3) cabe destacar la voz 2 en cuanto a los resultados obtenidos. En este caso, nos encontramos ante una voz que genera un claro sentimiento positivo además de transmitir un nivel de confianza elevado entre los encuestados. Sumando todo esto, podemos además ver como los participantes han demostrado una predisposición positiva hacia la configuración de dicha voz, aunque no se aprecia una clara seguridad en este aspecto ya que no destacan respuestas como “seguro que sí”.

Al hilo del primer cuestionario, la voz 3 sobresale por razones completamente opuestas. Observamos como genera un pensamiento preliminar ciertamente negativo entre la muestra y la confianza generada por esta voz es de un nivel bajo. Por todo esto, vemos una tendencia muy negativa hacia querer configurar esta voz en los dispositivos de los encuestados.

Por otro lado, se encuentra la voz 1, la cual situamos entre medias de las voces 2 y 3 ya que en este caso el sentimiento generalizado es neutro. Aun así, no consigue agradar del todo y por tanto la confianza generada y la intención de configuración no son de un nivel muy positivo.

En cuanto al segundo cuestionario se refiere, la voz 4 es la que peores resultados obtiene. Genera un sentimiento desagradable y, por tanto, no consigue agradar ni generar confianza entre los encuestados; todo esto explica la baja intención de configuración que se demuestra hacia esta voz.

En el lado opuesto, se encuentra la voz 6 cuyos resultados son realmente positivos. Palabras como “alegre” o “dulce” denotan el sentimiento positivo generado por esta voz, siendo por tanto la voz que más gusta y más confianza transmite dentro de este segundo cuestionario. Pese a ello, la intención de configuración es ambigua ya que mucha de la gente está segura de querer configurar esta voz, pero, al mismo tiempo, hay un gran número de “seguro que no”.

En cuanto a la voz 5, tal como pasaba con la voz 1 en el primer cuestionario, destaca su neutralidad. Nos encontramos ante una voz que no ha tenido una gran repercusión positiva entre los encuestados, pero tampoco destaca por su negatividad. Dada su neutralidad, es una voz en la que se deposita confianza, pero no se demuestra una intención de configuración clara.

Una vez se han analizado los sentimientos generados y la predisposición de los encuestados hacia cada una de las voces, observamos los diferentes resultados obtenidos por los adjetivos previamente definidos. Para esto, nos centramos primero en observar las posibles correlaciones entre los diferentes rasgos entre los que destacan las correlaciones positivas entre muchos de ellos (aunque no de valores muy elevados) y alguna correlación negativa. Todo esto, ayuda a entender el posterior análisis de componentes principales.

En este análisis, las diferentes voces quedan reflejadas en un mapa perceptual en base a las valoraciones recibidas por los encuestados. En primer lugar, podemos destacar una polaridad entre las voces donde unas destacan por sus valoraciones positivas (natural, creíble etc.) mientras que otro grupo queda marcado por su artificialidad y agresividad. En cuanto a las voces negativas se refiere, nos encontramos ante la voz 1, voz 3 y voz 4 siendo la voz 3 la que más se aleja del promedio. Tras esta, se encuentra la voz 4 cuyo

nivel de artificialidad es menor, aunque aumenta su nivel de agresividad aumenta. Por último, dentro de este bloque se encuentra la voz 1 que, aunque demuestra ser menos artificial que las anteriores, se encuentra mucho más representada por la cualidad agresividad.

En cuanto al bloque positivo se refiere, nos encontramos de forma destacada con las voces 2 y 6. Esta primera, se ve más representada por cualidades como la amabilidad, la credibilidad o el agrado mientras que la voz 6 se ve más representada por las cualidades naturalidad y familiaridad. Por último, se encuentra la voz 5 la cual se ve representada como una voz promedio.

Con la interpretación de todos los datos obtenidos y con ayuda de las descripciones de cada una de las voces utilizadas, hemos llegado a las diferentes conclusiones mostradas a continuación.

En primer lugar, se puede destacar que la cualidad más excluyente por parte de los encuestados ha sido la artificialidad. Esta característica ha sido muy mencionada a lo largo de todo el análisis ya que desde el inicio se ha visto reflejada en el pensamiento espontáneo de los encuestados y posteriormente ha sido utilizada para diferentes preguntas y como adjetivo predefinido. Además, se ha evidenciado que causa un efecto totalmente negativo en las voces ya que, a mayor nivel de artificialidad, menores porcentajes en cuanto a valor añadido o intención de configuración. Por todo esto, se concluye que las voces 1, 3 y 4 son las que generan un mayor rechazo tratándose las tres de voces femeninas sin alto grado de entonación.

Otra de las conclusiones sacadas del estudio es que lo que pueda llegar a gustar una voz o lo que se pueda confiar en ella no conlleva la intención de los encuestados a configurarla; esto se puede apreciar para el caso de las voces 2 y 6. En el caso de la voz 2 estamos ante una voz que demuestra transmitir más confianza y gusta más a los encuestados, sin embargo, no llega a generar una intención de configuración segura tan alta como muestra la voz 6 cuyo número de respuestas “seguro que sí la configuraría” es el más elevado de todos. Esto se da a pesar de que esta voz no consigue unos niveles de agrado o de confianza tan elevados como en el caso de la voz 2.

Además de lo ya comentado, observando el tipo de voces ante las cuales nos encontramos cabe destacar que no hay una clara predilección en cuanto al género de las voces y en cuanto a su agudeza o gravedad. A lo largo del estudio, se ha visto que

aunque las voces masculinas han demostrado generar un sentimiento más positivo, el hecho de encontrarnos antes una mayoría de voces femeninas hace que no sea posible demostrar el hecho de que hay una cierta predilección por algún género en concreto. Lo mismo sucede para el caso de lo aguda o grave que pueda llegar a ser cada voz ya que esta no ha sido una variable decisoria en el estudio.

5.1 Extrapolación al ámbito del marketing

Con lo analizado a lo largo de todo el estudio, hemos llegado a la conclusión de que una voz corporativa construye identidad de marca, las voces son elementos reconocibles que asociados a las marcas permiten a los consumidores recordar, conocer y diferenciarlas mejor. Llegamos a la conclusión que una voz corporativa puede funcionar igual que un logo de una marca, pero afectando al sentido auditivo de los clientes; de la misma manera que las personas reconocen el logo de Google, han sido capaces de reconocer la voz que usa Google en diferentes páginas web. Además de esto, las voces también aportan a generar experiencias de marca ya que tal y como se ha podido comprobar en este estudio las voces son capaces de generar reacciones en los consumidores que a su vez terminan asociando esos sentimientos a la marca.

La creación de I.A. capaz de recrear conversaciones fluidas y con sentido está aun en una fase muy primeriza, además, también se hace difícil conseguir que una voz reproducida por un aparato electrónico sea percibida con carácter humano. Aun así, se observan en el mercado primeros prototipos de lo que parece va a ser una revolución del modelo de compra actual hacía un mercado más individualizado y cómodo para el cliente. Creemos que la implantación de una voz corporativa para la empresa, aunque sea un prototipo inicial, puede suponer cierta ventaja competitiva para la susodicha empresa, no solo por el hecho de aprovechar las ventajas que puede ofrecer tener una voz corporativa, sino también crear cierto know-how para así disponer de los conocimientos y mecanismos necesarios para cuando la tecnología permita el desarrollo de asistentes de calidad.

A lo largo del estudio se ha podido ver que hay voces que, aunque generen sentimientos positivos, éstos no siempre van dirigidos a lo mismo; como es el caso de gustar y querer configurar. Por tanto, dependiendo del objetivo para el cual es diseñada cada una de las voces, es importante tener en cuenta unos factores u otros a la hora de realizar ese

diseño. Además, conseguir que una voz suene natural, se ha visto que es un claro signo de diferenciación y podría suponer una ventaja competitiva dentro del mercado, la gente confía más en voces que suenan humanas porque son más cercanas a ellos y desconfían de las voces robóticas porque no consiguen transmitir cercanía ni imitar el trato humano que se reproduce en las tiendas físicas.

Es por ello que el “Voice Marketing” en muchos casos, un mundo aún por descubrir, es una herramienta relevante a la hora de buscar crear una identidad de marca que lleve a las empresas a diferenciarse y a destacar sobre otras del mismo ámbito.

6. Bibliografía

Aaker, D. A., & Joachimsthaler, E. (2000). The brand relationship spectrum: The key to the brand architecture challenge. *California management review*, 42(4), 8-23.

Ballester, M. E. D., & Sabiote, E. F. (2011). Las experiencias de marca: estudio exploratorio de sus consecuencias en la relación consumidor-marca1. *European Journal of Management and Business Economics*, 20(3), 121.

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech communication*, 46(3-4), 252-267.

Carrillo, M. V., & Castillo, A. (2005). La Nueva Publicidad Digital (NPD): Servicios Digitales y Contenidos Interactivos que Generen 'Experiencias' en los Consumidores. *Razón y palabra*, 10(45).

Casado, J. C. A., & Merino, M. J. (2011). Comunicación experiencial y sensorial: algunos ejemplos de aplicación. *Harvard Deusto business review*, 198, 62-79

Cherniss, C., & Goleman, D. (2001). The emotionally intelligence workplace. *How to select for measure and improve emotional intelligence in individuals, groups and organizations san Francisco: Jossey-Bass.*

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

Dunteman, G. H. (1989). *Principal components analysis* (No. 69). Sage.

Ebrahim, R., Ghoneim, A., Irani, Z., & Fan, Y. (2016). A brand preference and repurchase intention model: The role of consumer experience. *Journal of Marketing Management*, 32(13-14), 1230-1259.

Estrada Barojas, C. O. (SF) Espectro de la voz humana y la música.

Fisher, R. A. (1992). Statistical methods for research workers. In *Breakthroughs in statistics* (pp. 66-70). Springer, New York, NY.

de Garcillán López-Rúa, M. (2015). Persuasión a través del marketing sensorial y experiencial. *Opción*, 31(2), 463-478.

Ghodeswar, B. M. (2008). Building brand identity in competitive markets: a conceptual model. *Journal of product & brand management*.

Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of consumer research*, 9(2), 132-140.

Knapp Duane, E. (2000). The Brand Mindset. Éditions MacGraw-Hill, New York, NY, États-Unis, Chapitre, 4, 69-93.

Melewar, T. C., Lim, L. L., & Petruzzellis, L. (2010). Mobile phone choice: technology versus marketing. The brand effect in the Italian market. *European Journal of marketing*.

Mellado, A., Tomicic, A., Martínez, C., Reinoso, A., & Bauer, S. (2017). La relevancia de la voz en el proceso psicoterapéutico: un estudio a dos voces. *Revista Argentina de Clínica Psicológica*, 26(3), 261-273.

Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard business review*, 85(2), 116.

Pine II, B. J. y Gilmore, J. H. (1998). "The Experience Economy: work is theatre and every business a stage". Editorial Harvard Business School Press. Boston.

Ramos-Serrano, M. (2007). La emoción como valor estratégico de la marca. De la inteligencia emocional al diseño Kansei. *Telos: Cuadernos de comunicación e innovación*, 71, 22-28.

Ramsden, A., & Bate, A. (2008). Using word clouds in teaching and learning. *University of Bath*. Retrieved December, 18, 2009.

Schmitt, B. (1999). Experiential marketing. *Journal of marketing management*, 15(1-3), 53-67.

Schmitt, B. H. (2010). *Customer experience management: A revolutionary approach to connecting with your customers*. John Wiley & Sons.

Tynan, C. & McKechnie, S. (2009). Experience Marketing: a review and reassessment, *Journal of Marketing Management*, 25(5/6), 501-517.

Vierna Carrasco, L. (2014). Marketing sensorial BMW: Análisis de campaña y precepción de los usuarios.

Zarantonello, L., & Schmitt, B. H. (2010). Using the brand experience scale to profile consumers and predict consumer behaviour. *Journal of Brand Management*, 17(7), 532-540.

Zhao, N., Dong, Q., Addis, M., Holbrook, M. B., Bennett, R., Hartel, C. E., ... & David, L. R. (2001). Fuzzy Sentiment Inference (FSI) of online reviews for consumer recommendation. *Information Technology Journal*, 12(20), 50-66.

Definición del concepto “Voz”: <https://dle.rae.es/voz?m=form>

7. Anexo

7.1 Anexo 1: Cuestionario

Gracias por acceder a responder a nuestro test, a continuación escuchará diferentes voces, nos gustaría que respondiera a las preguntas con sinceridad. La finalidad del cuestionario es única y exclusivamente académica, todos los datos adelantados serán usados de manera orientativa para los investigadores.

q1 Escucha la voz número uno y, usando solo una palabra o expresión danos tu opinión sobre ella



q1.2 ¿Algo más que pueda añadir? (Opcional)

q1.3 Voz 1

- | | |
|---------------------------------|-----------------------|
| Esta voz me parece agresiva | <input type="radio"/> |
| Esta voz me parece amable | <input type="radio"/> |
| Esta voz me transmite confianza | <input type="radio"/> |
| Esta voz me agrada | <input type="radio"/> |
| Esta voz me parece natural | <input type="radio"/> |
| Esta voz me parece familiar | <input type="radio"/> |
| La voz me parece creíble | <input type="radio"/> |
| Esta voz me parece artificial | <input type="radio"/> |

q1.4 ¿Confiarías en una marca que use esta voz?

- ☐ Si
☐ No

q1.5 ¿Configurarías esta voz como tu asistente virtual?

- ☐ Seguro que si
☐ Probablemente si
☐ No lo se
☐ Probablemente no
☐ Seguro que no

q2 Ahora te presentaremos la voz número dos, por favor, indica con una palabra o expresión que te sugiere esta voz.



q2.1 ¿Algo más que pueda añadir? (Opcional)

q2.2 Voz 2

- | | |
|---------------------------------|-----------------------|
| Esta voz me parece agresiva | <input type="range"/> |
| Esta voz me parece amable | <input type="range"/> |
| Esta voz me transmite confianza | <input type="range"/> |
| Esta voz me agrada | <input type="range"/> |
| Esta voz me parece natural | <input type="range"/> |
| Esta voz me parece familiar | <input type="range"/> |
| La voz me parece creíble | <input type="range"/> |
| Esta voz me parece artificial | <input type="range"/> |

q2.3 ¿Confiarías en una marca que use esta voz?

- ☐ Si
☐ No

q2.4 ¿Configurarías esta voz como tu asistente virtual?

- ☐ Seguro que si
☐ Probablemente si
☐ No lo se
☐ Probablemente no
☐ Seguro que no

q3 Por último te presentamos la voz numero tres, por favor indica con una palabra o expresión tu opinión sobre ella.



q3.1 ¿Algo más que pueda añadir? (Opcional)

q3.2 Voz 3

- | | |
|---------------------------------|-----------------------|
| Esta voz me parece agresiva | <input type="range"/> |
| Esta voz me parece amable | <input type="range"/> |
| Esta voz me transmite confianza | <input type="range"/> |
| Esta voz me agrada | <input type="range"/> |
| Esta voz me parece natural | <input type="range"/> |
| Esta voz me parece familiar | <input type="range"/> |
| La voz me parece creíble | <input type="range"/> |
| Esta voz me parece artificial | <input type="range"/> |

q3.3 ¿Confiarías en una marca que use esta voz?

- ☐ Si
☐ No

q3.4 ¿Configurarías esta voz como tu asistente virtual?

- ☐ Seguro que si
- ☐ Probablemente si
- ☐ No lo se
- ☐ Probablemente no
- ☐ Seguro que no

q4 ¿que voz o voces te han gustado más?

- ☐ Voz 1
- ☐ Voz 2
- ☐ Voz 3

q4.1 ¿En cual de las voces confiaría más?

- ☐ Voz 1
- ☐ Voz 2
- ☐ Voz 3

q4.2 ¿Cuál de las voces le parece más artificial?

- ☐ Voz 1
- ☐ Voz 2
- ☐ Voz 3

q5 Edad

- ☐ Menor de 18
- ☐ Entre 18 y 34 años
- ☐ Mayor de 34 años

q6 Género

- ☐ Femenino
- ☐ Masculino
- ☐ Otro

q7 País de origen

¡Muchas gracias por su aportación!

7.2 Anexo 2: Valores correlaciones

Column 1	Column 2	Correlation	Column 1	Column 2	Correlation
Amable	Confianza	0.79918294187	Artificial	Natural	-0.66958651558
Agrada	Confianza	0.79872152096	Artificial	Creible	-0.6101024583
Agrada	Amable	0.79591265706	Agrada	Artificial	-0.59200058108
Agrada	Creible	0.75640680441	Amable	Artificial	-0.53953887738
Creible	Natural	0.72174567539	Artificial	Confianza	-0.51474081019
Confianza	Creible	0.71534504216	Artificial	Familiar	-0.41822070731
Confianza	Natural	0.69180755816	Agresiva	Amable	-0.22588752841
Agrada	Natural	0.66248018847	Agrada	Agresiva	-0.1725348087
Amable	Creible	0.65475417337	Agresiva	Natural	-0.15123704061
Amable	Natural	0.64876375023	Agresiva	Creible	-0.13165523572
Creible	Familiar	0.60445911414	Agresiva	Confianza	-0.12437405294
Familiar	Natural	0.58479214599	Agresiva	Familiar	-0.080510012959
Agrada	Familiar	0.55493997029			
Confianza	Familiar	0.53540374561			
Amable	Familiar	0.5065853888			
Agresiva	Artificial	0.2272578126			

7.3 Anexo 3: Mapa perceptual

