

# **Interfaces dirigidas por voz aplicadas a vehículos autónomos.**

Memoria de Proyecto Final de Grado/Máster  
**Grado en Multimedia**  
Área usabilidad e interfaces

**Autor: Diego Marquina Trasobares**

Consultor: Natàlia Herèdia López  
Profesor: Ferran Gimenez Prado

18 de Junio de 2018



Esta obra está sujeta a una licencia de Reconocimiento [3.0 España de Creative Commons](https://creativecommons.org/licenses/by/3.0/)

*Any sufficiently advanced technology is indistinguishable from magic.*  
Arthur C. Clarke

*Cualquier tecnología lo suficientemente avanzada es indistinguible de la magia.*  
Arthur C. Clarke

## **Abstract**

Este trabajo busca investigar la aplicación de interfaces dirigidas por voz a la industria de los vehículos autónomos. Comandos analizados a través de reconocimiento de voz, asistentes virtuales o interfaces conversacionales son los principales candidatos para mejorar la experiencia de usuario de los vehículos autónomos. Las interfaces de voz no solo se plantean como apoyo a las interfaces GUI extensamente utilizadas, sino que en determinados momentos liberan al usuario de la necesidad de usar las manos para dedicarlas a otros propósitos.

Palabras clave: vehículo autónomo, interfaz conversación, voz, automoción, ux, ui, asistentes virtuales, reconocimiento de voz, síntesis de voz.

## **Abstract**

This job aims to investigate the application of voice driven interfaces to the self-driving vehicles industry. Analysed by voice recognition commands, virtual assistants or conversational interfaces are main candidates for improving user experience in self-driving vehicles. Voice interfaces are not only proposed as Graphical User Interfaces support, but also sometimes they free the user from the necessity to use their hands for driving so they can use them for anything else.

Keywords: self-driving car, conversational interface, voice, automotive, UX, UI, virtual assistants, voice recognition, voice synthesis.

# Índice

1. Introducción .....	8
2. Descripción .....	10
3. Objetivos.....	12
3.1 Principales .....	12
3.2 Secundarios.....	12
4. Marco teórico .....	13
Cuestiones previas al diseño de una interfaz dirigida por voz .....	15
5. Metodología .....	17
6. Diseño de interacción .....	18
7. Planificación .....	21
8. Analisis de interacción de vehiculos .....	25
9. Perfiles de usuario .....	29
Carmen – La práctica .....	30
Antonio – El padre de familia .....	30
Paula – La adolescente .....	31
Pelayo – El conductor premium .....	31
10. Propuesta de diseño .....	33
Roles.....	34
Objetivos .....	35
Seguridad.....	35
Memoria contextual .....	36
Integración.....	36
Configuración .....	36
Entradas.....	37
Salidas .....	37
Ayuda .....	37
Entidades .....	38
Intents.....	38
Ampliaciones .....	39
11. Analisis de usabilidad.....	40
Heuristico .....	40
Paseo cognitivo .....	42
Resultados de los Paseos cognitivos.....	45
12. Conclusión .....	47

Anexo 1. Entregables del proyecto .....	49
Bibliografía .....	50

# 1. Introducción

Durante años la ciencia ficción nos ha mostrado en numerosas ocasiones vehículos capaces de conducirse solos y con los que es posible una interacción utilizando la voz. Un ejemplo presente en la cultura popular es el que encontramos, por ejemplo, en KITT<sup>1</sup>, el coche fantástico. La sociedad de los años 80 no estaba tan familiarizada con los ordenadores como estamos en nuestros días y, pensando de un modo realista, le hubiera costado un tiempo adaptarse a un coche como el presentado en la serie de televisión. De cualquier modo, resultó de utilidad para que la visión de un coche independiente se asentara en la cultura popular y, como en otros casos provenientes de la ciencia ficción, facilitó un escenario sobre el que reflexionar y poder materializar esa visión en el futuro.

En la actualidad una nueva clase de pequeños electrodomésticos tales como Google Home<sup>2</sup> o Amazon Echo<sup>3</sup> han comenzado a hacerse un hueco en nuestros hogares. Podemos hablar con ellos como si de una persona se tratara y pedirles información, así como solicitar que realicen pequeñas acciones no físicas. Estos electrodomésticos desarrollan una adaptación a diferentes circunstancias a través de los *skills*<sup>4</sup>, capacidades que se programan para integrarse con otros servicios.

Si contemplamos la breve evolución temporal entre el asistente de Apple, Siri<sup>5</sup>, y estos nuevos electrodomésticos, así como el precio moderado que los acerca al público general, podemos observar un desarrollo positivo ante el asentamiento de las interfaces dirigidas por voz en la sociedad actual.

Existe otro concepto que ha florecido en los últimos años y es el de el Internet de las Cosas<sup>6</sup> (Internet of Things). Esta idea persigue crear herramientas dotadas de un pequeño circuito con conexión a internet y una lógica configurable. De esta forma podemos controlar dichos aparatos desde la oficina a través de una página web o desde el sofá de casa a través del móvil sin demandar un gran esfuerzo. Dispositivos como los ya mencionados Amazon Echo o Google Home nos permitirán activar los aparatos del internet de las cosas con nuestra propia voz.

La voz es un rasgo significativo del ser humano que ayuda a la identificación entre iguales y diferenciación de otros seres. Es probable que nos encontremos ante uno de los cambios más revolucionarios de todos los tiempos donde con la ayuda de la interacción vocal, se personifica y da vida a nuestros aparatos electrónicos dotándolos de sentimientos y personalidad, características que las marcas buscan que el consumidor asuma para sus productos.

La democratización de los coches autónomos ha acercado una realidad a los consumidores que hasta hace unos pocos años solo estaba al alcance de las fuerzas armadas. Más allá de la utilidad que presentan los vehículos autónomos, éstos todavía tienen que ser comandados por los usuarios.

---

<sup>1</sup> El coche fantástico. Obtenido de Wikipedia: [https://es.wikipedia.org/wiki/Knight\\_Rider](https://es.wikipedia.org/wiki/Knight_Rider)

<sup>2</sup> Google Home. Obtenido de Google: <https://support.google.com/googlehome/#topic=7029677>

<sup>3</sup> Amazon Echo. Obtenido de Amazon: <https://www.amazon.com/Amazon-Echo-And-Alexa-Devices/b?node=9818047011>

<sup>4</sup> Amazon Alexa Skills. Obtenido de Amazon: <https://www.amazon.com/b?&node=13727921011>

<sup>5</sup> Apple Insider (23-1-2018). Siri now actively used on more than 500M devices, up from 375M in June. Obtenido de Apple Insider: <https://appleinsider.com/articles/18/01/24/siri-now-actively-used-on-more-than-500m-devices-up-from-375m-in-june>

<sup>6</sup> Internet of things. Obtenido de Wikipedia: [https://en.wikipedia.org/wiki/Internet\\_of\\_things](https://en.wikipedia.org/wiki/Internet_of_things)

Éstos también han experimentado un avance notable y se han adaptado a nuevas formas de interacción, tal como los asistentes virtuales o las interfaces conversacionales. La industria precisa de la intersección de nuevos avances en el campo de las interfaces de usuario y el de los vehículos autónomos.

El futuro nos presenta ciertas cuestiones como si seremos capaces de controlar los vehículos a través de la voz. En caso afirmativo, ¿Cómo será la interacción? ¿Qué tipo de acciones podremos desencadenar en el vehículo? ¿Hasta donde llegará la inteligencia artificial?

Las interfaces de voz son un campo poco explotado en los aparatos de consumo, hecho probablemente ligado a las limitaciones tecnológicas. En la actualidad nos encontramos con que los principios de diseño gráfico y de interacción basados en interfaces gráficas no aplican a este nuevo soporte, pero sí que lo hacen los principios de diseño sobre los que se sustentan. También sería un reto crear un único interfaz para la diversidad de usuarios que existen, así como las diferentes formas en que esta interfaz se integraría en el producto final.

En la actualidad existen más retos que realidades sobre los vehículos autónomos que van más allá de una cuestión de interacción. Pueden ser tanto socioculturales como cibernéticas y solo el paso del tiempo resaltarán una solución. Un ejemplo donde queda reflejada la cuestión sociocultural de carácter más ético es cuando un conductor haga un gesto con la mano cediéndole el paso a un vehículo autónomo y éste no arranque creando un bucle indefinido. La voz es una gran ayuda ante la multitarea a la que estamos sometidos en nuestro día a día. Nos permitirá realizar otras actividades y ajustar parámetros en el automóvil, de forma similar a cuando estamos supervisando a un niño y realizando otra tarea al mismo tiempo.

## 2. Descripción

Dado que la autonomía de un vehículo y las interfaces dirigidas por voz pueden ser algo ambiguas y pueden aplicar a una única parte sin que afecte su inclusión dentro de la familia es importante aclarar los siguientes conceptos.

La SAE International (Sociedad Internacional de Ingenieros Automovilísticos) divide en 6 niveles los vehículos autónomos según sus capacidades de automatización<sup>7 8</sup>:

Conductor monitoriza el entorno		
Nivel 0	Sin automatización	El conductor realiza en todo momento la conducción
Nivel 1	Asistencia al conductor	El vehículo es capaz de acelerar, decelerar o girar atendiendo al entorno cuando se activa un modo de conducción
Nivel 2	Automatización parcial	El vehículo es capaz de acelerar, decelerar y girar atendiendo al entorno cuando se activa un modo de conducción
Sistema Automático de Conducción monitoriza el entorno		
Nivel 3	Automatización condicional	El vehículo cuando se activa el modo de conducción automática es capaz de realizar todas las tareas de conducción pero el conductor tiene que responder ante una petición de intervención.
Nivel 4	Automatización alta	El vehículo es capaz de realizar todas las tareas de conducción incluso cuando el conductor no responde ante una petición de intervención.
Nivel 5	Automatización completa	El vehículo es capaz de realizar todas las tareas de conducción en cualquier escenario que podría conducir un conductor humano.

Este trabajo toma como vehículos autónomos el segundo grupo que son capaces de monitorizar el entorno por sí mismos, especialmente los dos últimos niveles que trasladan al conductor a una posición de viajero con unas determinadas intenciones.

Las interfaces de voz también tienen varios niveles de profundidad. Las interfaces más simples son capaces de reconocer comandos simples de voz y actuar ante ellos. Un nivel más complejo son los

<sup>7</sup> SAE (16-1-2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. Obtenido de SAE:

[https://www.sae.org/standards/content/j3016\\_201401/](https://www.sae.org/standards/content/j3016_201401/)

<sup>8</sup>SAE (16-1-2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. Obtenido de SAE:

[https://www.sae.org/standards/content/j3016\\_201401/preview/](https://www.sae.org/standards/content/j3016_201401/preview/)

asistentes virtuales cotidianos que, dotados de inteligencia artificial, son capaces de buscar información o realizar tareas simples. El siguiente nivel se da a través de interfaces conversacionales que proveen de cierto contexto en la interacción y se aproximan más a la interacción cotidiana persona-persona. Finalmente existe un último nivel que recoge lo positivo de cada uno de los anteriores. Se trata de Asistentes virtuales que interactúan con el usuario a través de inteligencia artificial de forma similar a como el usuario lo haría con otra persona.

Las interfaces de voz son simples en funcionamiento. Lo que desde el exterior podría parecer la mimetización de una persona esconde un mecanismo bastante sencillo. En primer lugar la petición del usuario es captada por el micrófono. Posteriormente dicha entrada de voz se procesa y se extraen las palabras. Después las palabras se analizan con el fin de obtener la intención(intent) del usuario. Esta intención se realiza ya sea buscando información en recursos externos como servicios de internet o realiza la acción adecuada. Finalmente si fuera necesario se genera el sonido con la información necesaria a través de síntesis de voz.

En el artículo de opinión de OSNews<sup>9</sup> se plantea la interacción con el usuario como un complejo reto que llevará tiempo perfeccionar. Por un lado la propia interacción es difícil de realizar por ejemplo en la selección de destino. El Volvo XC90 incorpora una pequeña pantalla táctil con reconocimiento de texto, pero según la impresión del autor no da buenos resultados. También se identifican problemas de la propia experiencia de usuario no contemplados como por ejemplo de caso de uso en el que queremos ir a un destino pero no recordamos el sitio exacto. No está contemplada una funcionalidad del tipo “llevame cerca de este destino y recorre el vecindario”. Con la sinergia de la inteligencia artificial y las interfaces dirigidas por voz estos y otros problemas pueden solucionarse y así mejorar la experiencia de usuario global.

---

<sup>9</sup> David Adams. OS News (3-3-2016). Self-driving Cars: User Interface Will Be The Key To Success. Obtenido de OS News: [http://www.osnews.com/story/29107/Self-driving\\_Cars\\_User\\_Interface\\_Will\\_Be\\_The\\_Key\\_To\\_Success](http://www.osnews.com/story/29107/Self-driving_Cars_User_Interface_Will_Be_The_Key_To_Success)

## 3. Objetivos

### 3.1 Principales

Objetivos clave del TF.

- Encontrar posibles necesidades y nichos en la industria de los vehículos autónomos.
- Plantear soluciones basadas en interfaces de voz para mejorar la experiencia de usuario en los vehículos autónomos.

### 3.2 Secundarios

Objetivos adicionales que enriquecen el TF y que pueden sufrir variaciones.

- Dar a conocer la situación actual de las interfaces en los vehículos autónomos
- Facilitar el encuentro de nuevas sinergias aplicables a las interfaces de voz en los vehículos autónomos
- Reconocer las deficiencias de los vehículos autónomos que podrían ser solventadas a través de interfaces de usuario dirigidas por voz

## 4. Marco teórico

Podemos considerar los años 20 del siglo pasado como el inicio de la automatización vehicular. El ocho de diciembre de 1926 el Sentinel de Wisconsin<sup>10</sup> recogía la noticia de un un coche “fantasma” autoconducido de la marca Achen Motor que recorrería las calles de Milwaukee encendiendo su propio motor, girando el volante, embragando o accionando el claxon. Este acontecimiento anecdótico dio paso a una etapa de automatización donde mediante elementos mecánicos se reemplaza el esfuerzo humano.

Los siguientes avances reseñables fueron la electrónica, la robótica y las tecnologías de la información. Estos desarrollos permiten introducir lógicas simples y detección del entorno mediante sensores. John McCarthy, precursor de la inteligencia artificial, plantea a finales de los años 60 en Computed Controlled Cars<sup>11</sup> la construcción de un “chofer automatico” que controla el coche en función de las imágenes producidas por cámaras de video que posee el coche.

En la actualidad el abaratamiento de costes y el perfeccionamiento de la tecnología han permitido dotar a los vehículos de avances tales como la inteligencia artificial o sensores LIDAR<sup>12</sup>. Estos avances suponen un gran progreso y múltiples empresas han comenzado una carrera en la construcción de vehículos autónomos. Primordialmente, el esfuerzo se había centrado en la industria de la agricultura (tractores y cosechadoras autónomas) y logística (camiones semi-autónomos) pero en la actualidad parece que el sector de consumo es el que ha tomado protagonismo.

Los coches autónomos han heredado los controles y su posicionamiento de los vehículos convencionales para que los usuarios puedan reutilizar su conocimiento previo. El cuadro de mandos ahora tiende a ser una pantalla digital representando en una pantalla velocímetro y tacómetro analógicos. También existe un área en el salpicadero, entre los dos asientos, que ha sido habilitada en algunos casos con pantallas táctiles. Pero no se ha integrado de forma sistemática el control por voz.

Existen problemas como la falta de una legislación que permita la circulación de vehículos autónomos en la mayoría de los países. En la actualidad los principales avances los lideran empresas como:

Uber (Volvo XC90), Waymo-Google (Chrysler, Jaguar), Ford (Ford Fussion Hybrid), Apple, Tesla o Nissan (SAM)<sup>13</sup>

Otra cuestión importante es que no necesariamente un nivel superior de autonomía representa un modelo mejor. Existen diferentes métricas y gustos por parte de los usuarios en la elección de los

---

<sup>10</sup>The Milwaukee Sentinel (8-12-1926). ‘Phantom auto’ will tour city. Obtenido de Google News: <https://news.google.com/newspapers?nid=1368&dat=19261208&id=unBQAAAIBAJ&sjid=QQ8EAAAIBAJ&pg=6190,3634062>

<sup>11</sup> John McCarty (1968). Computer Controlled Cars. Obtenido de Universidad de Stanford: <http://www-formal.stanford.edu/jmc/progress/cars/cars.html>

<sup>12</sup> Alex Davies. Wired (2-6-2018). WHAT IS LIDAR, WHY DO SELF-DRIVING CARS NEED IT, AND CAN IT SEE NERF BULLETS?. Obtenido de Wired: <https://www.wired.com/story/lidar-self-driving-cars-luminar-video/>

<sup>13</sup>Leadercar Levante. Las 6 mejores marcas de coches autónomos. Obtenido de Leadercar Levante: <http://leadercarlevante.com/las-6-mejores-marcas-de-coches-autonomos/>

vehículos. No obstante, probablemente influenciados por la ciencia ficción se puede constatar un patrón de diseño minimalista en estos vehículos autónomos, donde se tiende a minimizar la interacción entre la persona responsable de la conducción y el vehículo. En el video realizado por el medio Car Throttle en el que se prueba un prototipo de vehículo autónomo de nivel 4 de la casa Renault se puede apreciar como es el propio coche sin intervención humana quien ajusta la velocidad y realiza adelantamientos.<sup>14</sup>

Respecto a las interfaces basadas en voz la primera de ellas fue desarrollada por los laboratorios Bell en 1952. Audrey solo era capaz de reconocer dígitos y no fue hasta 1962 cuando IBM mostró su Shoebox que era capaz de entender 16 palabras inglesas. Con esos referentes es sencillo observar cómo la evolución de otras disciplinas tales como la inteligencia artificial repercuten positivamente en la posibilidad de crear sistemas más avanzados.

En Octubre de 2011 hizo aparición Siri de Apple. Este asistente virtual acercó a multitud de usuarios la posibilidad de responder a preguntas simples y desencadenar acciones en su teléfono móvil a través de la voz. También recientemente ha prosperado el uso de interfaces conversacionales en la que un usuario habla con un agente para interactuar con el producto. El agente en cuestión es un sistema informático y la interfaz conversacional no necesariamente funciona con la voz. Inicialmente la interacción se producía a través de chatbots cuyo comportamiento era más o menos mecánico, pero en la actualidad se ha evolucionado hasta comportamientos más humanos de interacción pura con voz como en la demostración producida por Google de su producto Duplex<sup>15</sup>.

Como ya se ha puntualizado con anterioridad el mercado de los vehículos autónomos es incipiente y todavía no se ha asentado un modelo de producto concreto que el usuario tome como referencia. Es por ello que cada fabricante diferencia experiencias de usuario distintas apoyándose en sus estrategias y target de mercado. Por ejemplo Volvo con sus XC90 hace hincapié en que el usuario tiene el control del vehículo y no hace referencia a sus capacidades autónomas sino de forma puntual cuando describe sus capacidades de “asistencia de conducción”. Es una apuesta conservadora que aparentemente incorporará mejoras en el campo de los vehículos autónomos de forma más cautelosa. Diametralmente opuesta se encuentra la compañía Tesla que ofrece su conducción autónoma a través de Autopilot como funcionalidad principal, aunque ésta no sea totalmente autónoma y requiere de supervisión del conductor.

Los modelos de Tesla disponen de una gran pantalla táctil interactiva en el salpicadero. Así mismo algunas funcionalidades pueden utilizarse desde el teléfono móvil tras vincularse al coche.<sup>16 17</sup>

Waymo centra su valor en ser capaz de detectar todos los elementos de su entorno y sus características (velocidad, distancia...) y ser capaz de mostrarlos en el interior de sus vehículos. Waymo crea vehículos destinados a ser usados bajo suscripción por lo que diseña una experiencia

---

<sup>14</sup> Car Throtte (27-12-2017). Testing The World's Smartest Autonomous Car (NOT A Tesla). Obtenio de Youtube: <https://www.youtube.com/watch?v=I3ELVACR2VY>

<sup>15</sup> Cris Welch. The Verge (8-5-2018). Google just gave a stunning demo of Assistant making an actual phone call. Obtenio de The Verge: <https://www.theverge.com/2018/5/8/17332070/google-assistant-makes-phone-call-demo-duplex-io-2018>

<sup>16</sup> Tesla. Tesla Autopilo. Obtenido de Tesla: [https://www.tesla.com/es\\_ES/autopilot](https://www.tesla.com/es_ES/autopilot)

<sup>17</sup> Tesla. Tesla software. Obtenido de Tesla: [https://www.tesla.com/es\\_ES/software](https://www.tesla.com/es_ES/software)

minimalista y genérica que no permite al usuario interacción con el vehículo, sino que una vez seleccionado el destino éste toma las decisiones de forma autónoma.<sup>18 19 20</sup>

El volvo XC90 es capaz de conducir de forma autónoma, pero espera que el usuario indique acciones como el adelantamiento para que este se realice.<sup>21</sup>

Dado que las interfaces comandadas por voz solo han visto un esplendor en los últimos años no es extraño ver que no se han implementado en vehículos autónomos. La compañía Nuance Communications que desarrolla productos de reconocimiento de voz para diferentes industrias y público ha desarrollado un módulo de su producto llamado Dragon Drive. Dado el carácter generalista del producto la integración con un vehículo autónomo parece superficial. Y su funcionalidad parece bastante limitada permitiendo elegir destino, cambiar la temperatura del climatizador y otras de carácter más general como reproducir música. Este producto dispone de una palabra de despertar que es “Hola Dragon”.<sup>22 23</sup>

## Cuestiones previas al diseño de una interfaz dirigida por voz

En primer lugar, las características de las interfaces dirigidas por voz resultan muy específicas. Aunque la interacción es rápida podría no ser suficientemente rápida en el desarrollo de algunas tareas debido a las latencias, que pueden ser tanto del usuario al ejecutar una petición como el tiempo de procesamiento o incluso de red si fuera necesario, con lo que se descarta la implementación de funcionalidades como la frenada de emergencia con esta interfaz. Este sistema es especialmente exitoso para peticiones complejas o poco claras ya que se simplificarán mediante inteligencia artificial o se realizarán sugerencias.

Si el reconocimiento de voz y síntesis de voz funcionara de manera continua se podrían producir efectos no deseados ya sea por interpretar una conversación dentro del vehículo como una petición al propio coche o por interpretación de ruidos como palabras humanas. Este problema se ha solucionado tradicionalmente con una palabra para despertar (wake word) por lo que será necesario implementar dicha palabra para interactuar con el sistema. Es igualmente importante poder activar o

---

<sup>18</sup> Darrell Etherington. Tech Crunch (31-8-2017) Waymo focuses on user experience, considers next steps. Obtenido de Tech Crunch: <https://techcrunch.com/2017/10/31/waymo-self-driving-ux>

<sup>19</sup> Waymo. Waymo Tech. Obtenido de Waymo <https://waymo.com/tech/>

<sup>20</sup> Kirsten Korosec. Fortune(31-8-2017).Here's the First Look Inside Waymo's Self-Driving Minivans. Obtenido de Fortune: <http://fortune.com/2017/10/31/waymo-self-driving-minivan-inside/>

<sup>21</sup> Volvo. Coche XC90. Obtenido de Volvo: <https://www.volvocars.com/intl/cars/new-models/xc90>

<sup>22</sup> Lynn Walford. Connected car news(15-1-2017). Autonomous & Connected Car News: new software & platforms. Obtenido de Connected car news: <http://www.autoconnectedcar.com/2017/01/autonomous-connected-car-news-new-software-platforms/>

<sup>23</sup> Wayne Cunningham. Road Show (1-2-2017).Car voice commands won't suck with Nuance's assistant. Obtenido de Road Show: <https://www.cnet.com/roadshow/news/voice-command-in-cars-sucks-nuances-virtual-assistant-makes-it-better/>

desactivar la síntesis de voz ya que esto podría afectar a las actividades que estén realizando los usuarios y acabar siendo molesto. Adicionalmente se podría reemplazar esta activación del reconocimiento de voz y de la síntesis de voz mediante sendos botones físicos. De esta forma la activación del sistema sería más rápida y podremos seguir utilizando la voz para realizar acciones como una conversación cotidiana con el resto de los pasajeros.

El diseño de interfaces para vehículos autónomos requiere de interrelación el campo de la psicología. ¿Qué siente el usuario cuando percibe que no conduce él? ¿Se debe indicar de alguna forma que el coche está siendo conducido de forma autónoma? ¿Se debe sugerir al conductor que acciones realizar cuando no está al control del volante? ¿Se debe plantear la percepción de un asistente en lugar de conducción autónoma pura?

Otro elemento muy importante de la conducción autónoma a tener en cuenta es la confianza. Al conducir un vehículo convencional el conductor comienza a sentirse seguro cuando se familiariza con la distancia de frenado, el tamaño del vehículo etc. En relación a los vehículos autónomos el conductor se entrega al propio vehículo que podría colisionar y hacerle perder la vida. Transmitir seguridad y confianza al conductor y permitirle retomar el control en cada momento es muy importante. Para dar más confianza es recomendable mostrar información justificando las acciones realizadas por el vehículo incluso anticipando la información a la decisión. Un ejemplo de ellos es indicar en el propio vehículo que un semáforo al que se va a llegar está en rojo y que se va a proceder a frenar progresivamente. De esta forma el usuario se sentirá seguro y no experimentará ansiedad.

Respecto a la interacción de voz existen dos grupos de interacción. En primer lugar, se encuentra la interacción de voz “de una vez” que representa un comando transaccional e interfaces conversaciones capaces de preguntar al usuario.

En las interfaces conversacionales se puede crear un flujo conversacional que tiene en cuenta las respuestas de la propia interacción para acotar las respuestas y el discurrir de la conversación.

El diseño técnicamente viene determinado por la palabra de despertar, utterances e intents. Las utterances son las expresiones mínimas que el sistema de reconocimiento de voz es capaz de analizar. Y los intents son las intencionalidades en que se traducen las diferentes utterances. El conjunto de diferentes intents y utterances se puede modelar a través de un lenguaje XML específico para la tarea llamado Speech Synthesis Markup Language

Es necesario mostrar al usuario tanto que una petición ha sido ejecutada con éxito o no como que se está procesando la información y en un futuro próximo el usuario recibirá el resultado de la operación.

## 5. Metodología

En el presente trabajo se realiza un análisis de la situación actual de los vehículos autónomos y de cómo está implementada la interacción con el usuario. También se lleva a cabo un análisis de la situación actual de las interfaces de usuario basadas en voz. A partir de dichos análisis se buscan componentes o soluciones basadas en voz para ser aplicadas a la industria de los vehículos autónomos. Se desarrollará un diseño basado en Diseño Centrado en el Usuario<sup>24</sup> y se realizará un test heurístico de usabilidad y un paseo cognitivo por cada una de las personas identificadas.

Dado que los coches completamente autónomos (nivel 4 y 5 SAE) todavía no son una realidad de nuestra sociedad<sup>25</sup> es complejo analizar cómo estos funcionan o funcionarán. Existen documentos académicos sobre determinados aspectos de los vehículos autónomos, pero debido al gran avance y a la naturaleza generalmente privada y hermética de éstos no es posible limitar la información al ámbito académico.

El trabajo parte de un análisis actual de la industria de los vehículos autónomos y de las interfaces dirigidas por voz. Se muestra su evolución y la situación actual. Más adelante se realiza un análisis de interacción<sup>26</sup> de un vehículo autónomo genérico. Para ello se analizan los vehículos autónomos desarrollados por las principales compañías del sector, así como funcionalidades que podrían mejorar la experiencia de usuario. Se analizan también diferentes perfiles que podrían ser usuarios de vehículos autónomos.

En el núcleo del trabajo se plantea la convergencia de los vehículos autónomos y las interfaces de voz. Se realiza un diseño que posteriormente se validará con métodos heurísticos<sup>27</sup> y con paseos cognitivos<sup>28</sup>.

Inicialmente se plantea la convergencia de vehículos autónomos e Interfaces dirigidas por voz. Para validar la viabilidad y determinar como se ha de concretar la convergencia se realiza un analisis de la situación actual de los vehículos autónomos y de las interfaces basadas en voz. Para contextualizar el escenario donde se articula la sinergia de los vehículos autónomos y las interfaces dirigidas por voz se establece un Marco teorico. Para concretar la interacción se determinan cuales son los actores y elementos principales y se crea un diseño de interacción con sus respectivos diagramas de flujo.

---

<sup>24</sup> Usability Gov. User Centered Design. Obtenido de Usability Gov <https://www.usability.gov/what-and-why/user-centered-design.html>

<sup>25</sup> Luis Carlos. Auto 10. 8-1-2018. Coches Autonomos que ya se venden en España. Obtenido de Auto 10: <http://www.auto10.com/reportajes/coches-autonomos-que-ya-se-venden-en-espana/16088>

<sup>26</sup> Augusto Cortez Vásquez, Carlos Yáñez Durán, Luzmila Pró Concepción Universidad Nacional Mayor de San Marcos. Técnicas de análisis para el diseño de interface de usuario. Obtenido de RevistasInvestigacion:

<http://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/download/5639/4888>

<sup>27</sup> Mònica Zapata Lluch. UOC. Métodos de evaluación sin usuarios. Obtenido de UOC: [http://materials.cv.uoc.edu/daisy/Materials/PID\\_00176621/pdf/PID\\_00176613.pdf](http://materials.cv.uoc.edu/daisy/Materials/PID_00176621/pdf/PID_00176613.pdf)

<sup>28</sup> Mònica Zapata Lluch. UOC. Métodos de evaluación sin usuarios. Obtenido de UOC: [http://materials.cv.uoc.edu/daisy/Materials/PID\\_00176621/pdf/PID\\_00176613.pdf](http://materials.cv.uoc.edu/daisy/Materials/PID_00176621/pdf/PID_00176613.pdf)

El siguiente objetivo es realizar una propuesta de interacción. Para conseguirlo se realizará un Análisis de Interacción para esclarecer cuales son los puntos clave en el proceso de diseñar una experiencia adecuada dirigida por voz en el contexto de los vehículos autónomos. Adicionalmente y para saber cuál será el público objetivo de nuestro diseño se busca mediante datos estadísticos la creación de personas<sup>29</sup>. Dichos perfiles servirán para realizar una evaluación de paseo cognitivo. Se validarán los principales valores de usabilidad mediante un test heurístico. Finalmente, se sintetizarán los resultados en una conclusión.

## 6. Diseño de interacción.

La interacción se desarrolla a través del sonido. El usuario se comunica con el vehículo mediante su voz y el vehículo con el usuario gracias a voces generadas y emisión de sonidos en forma de alertas. En ocasiones puntuales como el proceso de configuración o la muestra de errores severos es necesario otro soporte de interacción con el usuario

La entrada del sistema es la propia voz del usuario que es recogida por un micrófono y procesada por un pequeño ordenador a bordo del vehículo. Este pequeño ordenador a través de técnicas de Inteligencia Artificial procesa la entrada<sup>30</sup> e interactúa con el vehículo o con recursos externos en caso de ser necesario.

Cuando el sistema de Inteligencia Artificial quiere comunicarse con el usuario lo puede hacer a través de sonidos de diferentes características para notificar o a través de una voz generada por el pequeño ordenador de a bordo. Finalmente, un altavoz emite las ondas que llegan hasta el usuario.

En este trabajo se han planteado tres niveles de interacción. A continuación, se procede a describir como sucede la interacción desde el punto de vista de la Inteligencia Artificial.

### Nivel 0

Cada entrada se identifica con una Acción. Es la más simple de las interacciones y por ello es concisa y requiere de precisión.

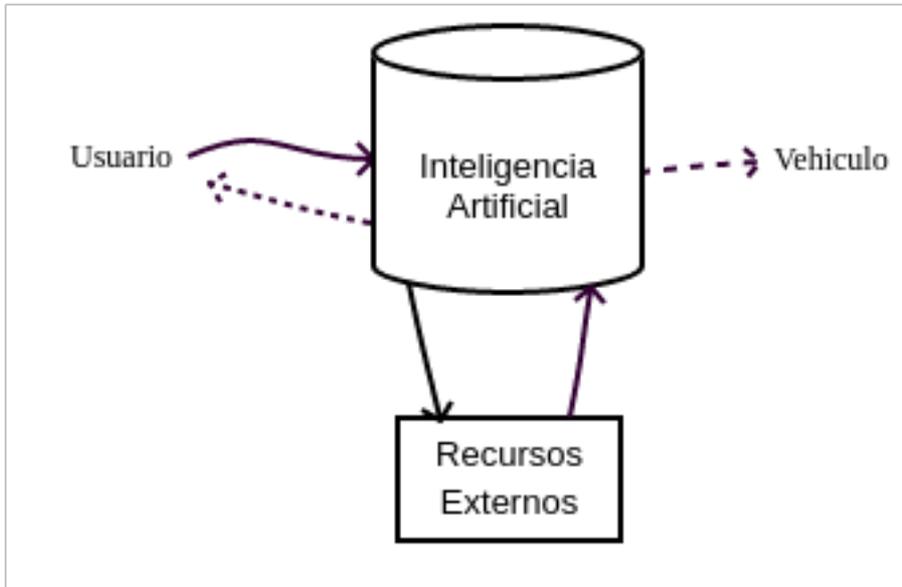


<sup>29</sup>Wikipedia. Persona(user experience). Obtenido de Wikipedia: [https://en.wikipedia.org/wiki/Persona\\_\(user\\_experience\)](https://en.wikipedia.org/wiki/Persona_(user_experience))

<sup>30</sup>Wikipedia. Reconocimiento del habla. Obtenido de Wikipedia: [https://es.wikipedia.org/wiki/Reconocimiento\\_del\\_habla](https://es.wikipedia.org/wiki/Reconocimiento_del_habla)

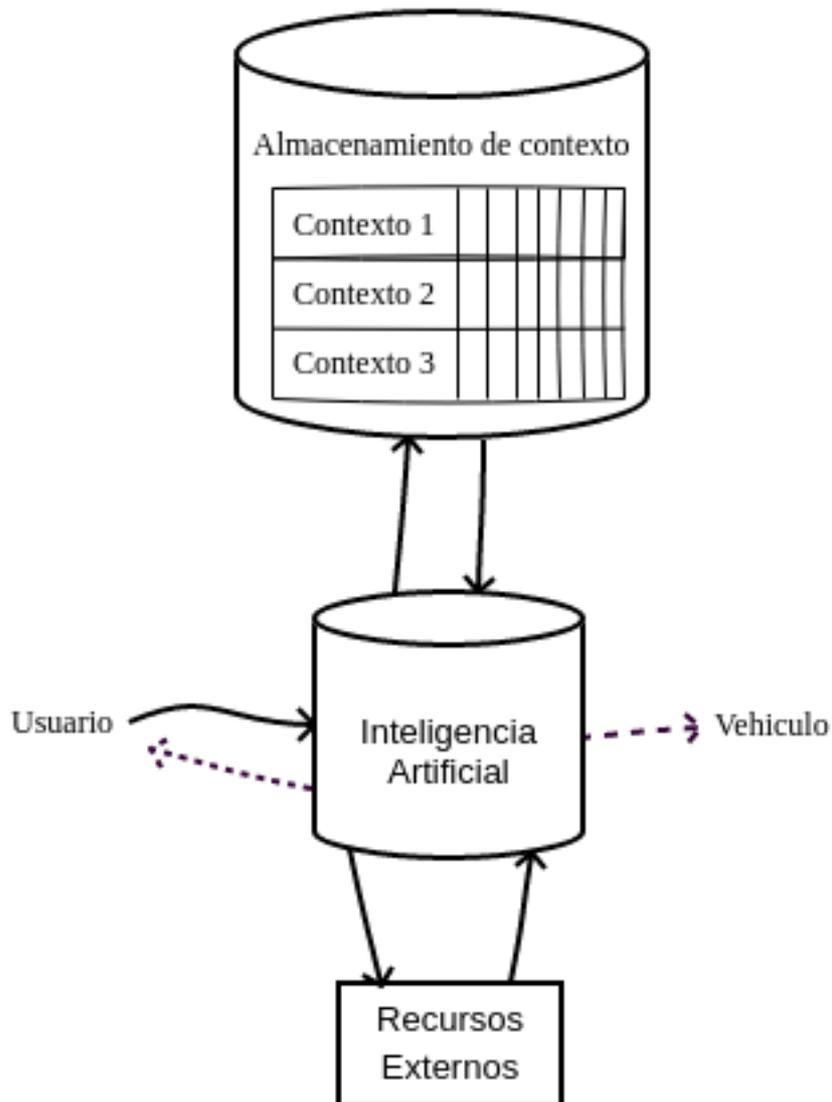
### Nivel 1

El usuario interactúa con el vehículo a través de frases sencillas que están preestablecidas. El vehículo es capaz de identificar frases similares si no existe una coincidencia exacta.



## Nivel 2

La interacción se da a través de una interfaz conversacional<sup>31</sup> más compleja y el sistema de Inteligencia Artificial es capaz de modificar las frases del usuario y ajustarlas a la más similar<sup>32</sup>. La característica principal de este nivel es que las interacciones individuales se almacenan durante toda la sesión y se toman en cuenta cuando se ha de realizar un análisis.



<sup>31</sup> John Brownlee. Fast codesign. (4-4-2016). Conversational Interfaces, Explained. Obtenido de Fast codesign: <https://www.fastcodesign.com/3058546/conversational-interfaces-explained>

<sup>32</sup> Adam Cheyer (et al.) A Case Study in Engineering a Knowledge Base for an Intelligent Personal Assistant. Obtenido de [www.adam.cheyer.com](http://www.adam.cheyer.com) <http://www.adam.cheyer.com/papers/ontology-overview-semantic-desktop.pdf>

## 7. Planificación

Este proyecto se inicia tras una simplificación del alcance del proyecto planteado anteriormente: “Interfaces adaptadas al contexto de usuario”.

Esa propuesta consistía en explorar una interfaz de usuario genérica a través de diferentes soportes aunque finalmente se decidió únicamente centrarse en la interacción mediante voz.

Se planteaba un caso práctico destinado a los vehículos autónomos, pero se decidió dar más protagonismo a éstos para acotar su aplicación.

Debido a la limitación de tiempo no era posible realizar la cantidad suficiente de pruebas con usuarios que garantizara una muestra representativa de usuarios.

Se han dividido las tareas en diferentes grupos y algunas de ellas son dependientes entre sí. Los diferentes tipos de tareas se aprecian especialmente en los diagramas de Gantt. Cada tipo de tarea se identifica con un color diferente:

Azul: Revisión.

Amarillo: Análisis.

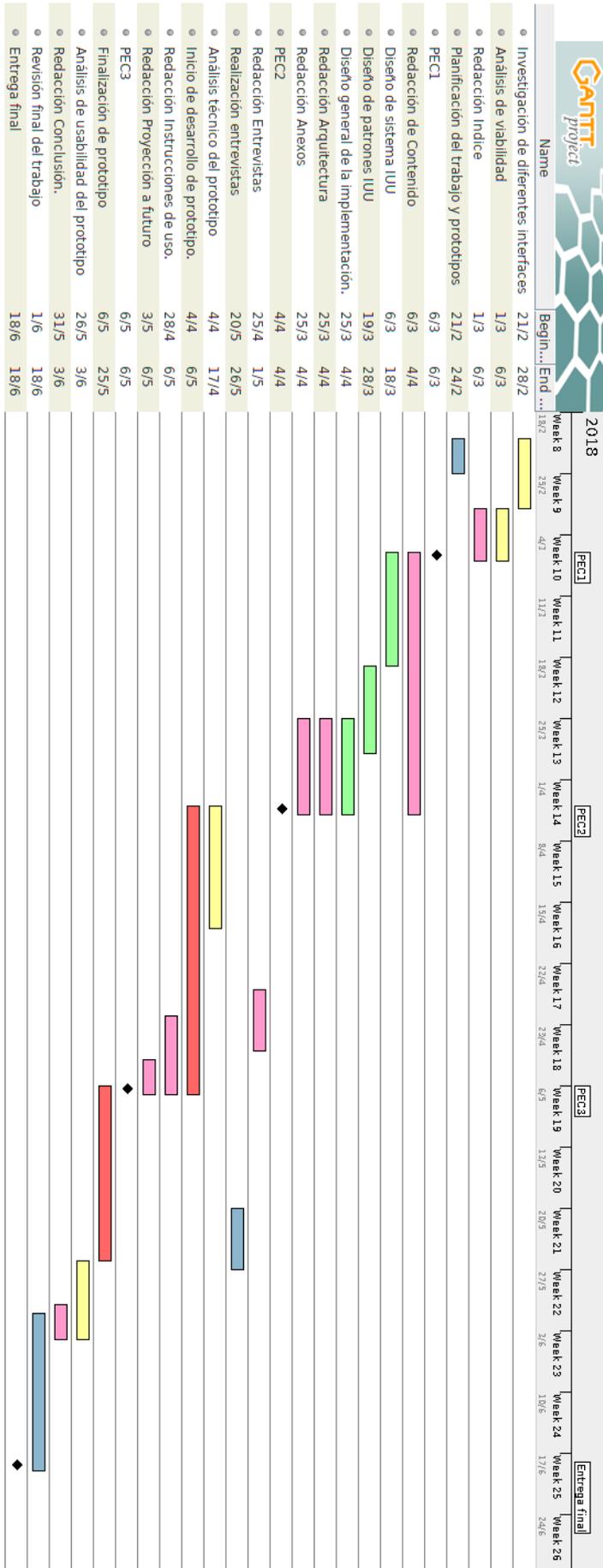
Rosa: Redacción.

Verde: Diseño.

Naranja: Desarrollo.

Rojo: Hito.

El primer diagrama muestra la planificación planteada para el primer trabajo y la segunda se identifica con el presente.



**Semana 1 (18 - 25 mayo)**

Propuesta del proyecto,  
Estructura del documento  
Redacción de:  
Abstract  
Introducción  
Descripción  
Estado de la cuestión  
Objetivos  
Metodología  
Planificación

**Semana 2 (25 mayo - 1 junio)**

Realización de análisis de interacción en un coche  
Redacción de:  
Proceso de trabajo  
Análisis de interacción en un vehículo

**Semana 3 (1 junio - 8 junio)**

Redacción de:  
Interfaces de voz en vehículos autónomos  
Comandos  
Asistentes virtuales  
Interfaces conversacionales  
Diseño de interacción.  
Perfiles de usuario  
Realización de prototipo de baja resolución

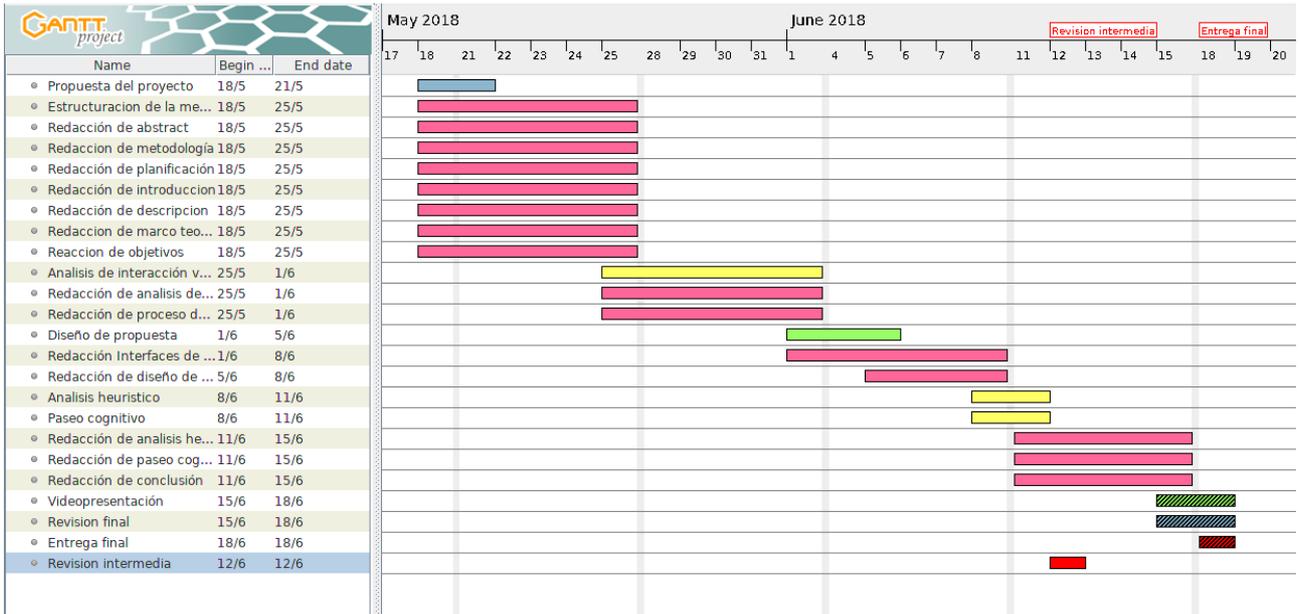
**Semana 4 (8 junio - 15 junio)**

Análisis heurístico de usabilidad  
Redacción de:  
Análisis heurístico de usabilidad  
Conclusión  
Video presentaciones  
Autoevaluación  
Bibliografía

**Revisión final (15 junio - 18 junio)**

Grabación de presentaciones  
Edición de presentaciones

Interfaces dirigidas por voz aplicadas a vehículos autónomos, Grado en Multimedia. Diego Marquina Trasobares



## 8. Analisis de interacción de vehiculos

Los vehículos autónomos por lo general se conducen de forma manual y es necesario activar el modo de conducción automática explícitamente para que éste conduzca de forma automática. Así mismo en determinadas ocasiones el vehículo anticipará situaciones en las que solicitará al usuario que tome el control del vehículo. En las interfaces de voz también existe este mismo patrón (llamado “Wake word” o palabra para despertar) donde tras enunciar una palabra el sistema de reconocimiento de voz comienza a interpretar las diferentes órdenes.

Los vehículos autónomos realizan una transición suave desde el modo manual al automático y viceversa por lo que cuando el usuario quiere rectificar una acción del vehículo no resulta en un cambio brusco. Al analizar vídeos de pruebas realizadas con gente podemos apreciar que los usuarios no saben cómo interactuar con el coche. El volante comienza a girar solo y ellos no saben si deben seguir agarrándolo o qué movimiento realizar. Dado que no se asimila a ninguna otra experiencia previa los usuarios que contemplan cómo el coche tiene “vida propia” comienzan a hablarle como si realmente se tratara de otro ser. Le dicen “Gracias” o le preguntan al conductor del test qué es lo que tienen que hacer. Aparentemente la forma natural de interacción es mediante la voz.

Tras analizar la experiencia de usuario de un vehículo convencional y un vehículo autónomo podemos ver que ésta está orientada a dos diferentes niveles de detalle.

En el primero el conductor interactúa constantemente con el vehículo girando el volante y los pedales principalmente y eventualmente con otros elementos como la radio, climatización etc. En el segundo caso la interacción es menos detallada y se le indica al vehículo una intención que este posteriormente resuelve: “Quiero ir a mi casa” o “Realiza un adelantamiento”

La tendencia de la interacción de vehículos autónomos vemos que es minimalista y se busca la funcionalidad de transporte global. Una persona podría estar acostumbrada a una determinada ruta pero su vehículo autónomo podría elegir otra que fuera más adecuada en ese determinado momento pudiendo producir cierta incomodidad al usuario.

Parte del diseño de interacción de estos nuevos vehículos autónomos consiste en saber qué decisiones preguntar al usuario y que decisiones asumir. Si tras cada comando el vehículo nos presenta alternativas o solicita confirmación entre varias opciones la experiencia de usuario puede ser exhausta. También existen cuestiones culturales que deberían de ser incorporadas al sistema de toma de decisiones del vehículo como es por ejemplo no conducir por pequeñas calles en suburbios por la noche. A continuación se puede ver cómo es la interacción con diferentes vehículos:

Continental<sup>33</sup>

Tesla<sup>34</sup>

Volvo<sup>35</sup>

---

<sup>33</sup> Continental Automotive Global. Continental Automotive Global. Obtenido de [www.youtube.com](https://www.youtube.com/watch?v=xBICfWQbXiE)  
<https://www.youtube.com/watch?v=xBICfWQbXiE>

<sup>34</sup> Tesla Self-Driving Demonstration. Obtenido de [www.tesla.com](https://www.tesla.com/es_ES/videos/autopilot-self-driving-hardware-neighborhood-long)  
[https://www.tesla.com/es\\_ES/videos/autopilot-self-driving-hardware-neighborhood-long](https://www.tesla.com/es_ES/videos/autopilot-self-driving-hardware-neighborhood-long)

<sup>35</sup> YOUCAR. Introducing Volvo Interface for Self-Driving Cars. Obtenido de [www.youtube.com](https://www.youtube.com/watch?v=19Zf6BWbBFA)  
<https://www.youtube.com/watch?v=19Zf6BWbBFA>

Mercedes.<sup>36</sup>

En la actualidad una de las compañías que más avances ha realizado en la materia es Tesla. Por lo que se procede a analizar las funcionalidades de interacción de su modelo X<sup>37 38 39</sup>. En los siguientes párrafos se describen las diferentes características que el vehículo incorpora:

#### Climatización automatizada

El vehículo es capaz de controlar la climatización. La principal diferencia es que está automatizada por lo que podría controlarse a través de cualquier interfaz de usuario del vehículo o incluso de forma autónoma.

#### Retrovisores autoajustables para ver bordillos y bolardos con marcha atrás

Cuando se activa la marcha atrás los retrovisores se ajustan para mostrar bordillos y bolardos que se encuentran en torno al vehículo

#### Conexión a móvil para interacción a través de él

El vehículo permite ser vinculado a móvil para poder controlar algunas de sus funcionalidades. En la actualidad este sistema solo funciona mientras el móvil se encuentra a una distancia no muy alejada del vehículo.

#### Ajuste de suspensión

La suspensión puede ajustarse desde el interior tan solo presionando la configuración adecuada.

#### Función de convocación

Esta función permite desplazar el coche hacia adelante y hacia atrás para el caso en que no podamos acceder a él porque no se pueden abrir las puertas. Es necesario realizar dicha acción a través del móvil.

#### Ubicación sincronizada con dispositivos

La ubicación se sincroniza con tus dispositivos esto ayuda a ver el estado actual del vehículo y dónde se encuentra. Es especialmente útil cuando no recordamos donde lo hemos aparcado.

#### Bloqueo y arranque

Estas funciones se realizan al apretar un botón por lo que podrían ser utilizadas potencialmente a través de comandos de voz.

#### Destello, claxon, arranque, apertura maletero delantero y trasero.

Estas otras funciones secundarias así como las dos anteriores se pueden ejecutar desde el móvil para facilitar el uso del vehículo.

---

<sup>36</sup> CARJAM TV. Mercedes S Class Self Driving Car Is Here 2015 Autonomous Car Real Roads S Class W222 CARJAM TV. Obtenido de [www.youtube.com](http://www.youtube.com)

<https://www.youtube.com/watch?v=VDwMhSobaOg>

<sup>37</sup> <https://www.tesla.com/modelx>

<sup>38</sup> Wikipedia. Tesla Model X. Obtenido de Wikipedia [https://en.wikipedia.org/wiki/Tesla\\_Model\\_X](https://en.wikipedia.org/wiki/Tesla_Model_X)

<sup>39</sup> clipset. Tesla X review - el INCREIBLE coche con ALAS. Obtenido de [www.youtube.com](http://www.youtube.com)  
<https://www.youtube.com/watch?v=CSL37ix-jSk>

### Información de carga

Cuando el vehículo, que es eléctrico, está en estado de carga nos muestra información como cuál es la capacidad de carga actual y el tiempo restante.

Cámaras traseras para ayuda así como información de sensores.

El coche dispone de cámaras traseras así como otros sensores que nos indican información del entorno.

Cuando se plantea la realización de una interfaz de usuarios para estos vehículos se tiene a realizar experiencias minimalistas. Es el caso de UX Studio que ha realizado una propuesta de experiencia de usuario con un número reducido de componentes.<sup>40 41</sup>

Conecta el vehículo a través de una aplicación móvil para indicarle información y aumentar las posibilidades de interacción. Es a través de esta aplicación como solicita el vehículo (de alquiler). Una vez dentro del coche comienza la interacción a través de voz y selecciona el destino. La pantalla ofrece información adicional y comienza el viaje. De forma similar a Waymo muestra información en la pantalla sobre el entorno.

Este tipo de interfaces aparentan ser sencillas de utilizar pero no muestran cómo resuelven posibles conflictos o cómo interactúan con el usuario en casos excepcionales como cuando este tiene que tomar los mandos del vehículo.

También hay que tener en cuenta el diferente trasfondo de los diferentes usuarios y las diferentes prioridades y características de los usuarios. En una publicación de la universidad de Eindhoven<sup>42</sup> se reevalúa el análisis de MDSI que incluía ocho factores (o perfiles) de conducción y crea cinco diferentes grupos:

#### Conducción enfadada (Angry driving)

Son conductores que hacen sonar el claxon, insultan a otros conductores o conducen demasiado cerca del coche que se encuentra delante

#### Conducción riesgosa (Risky driving)

Estos conductores disfrutan la conducción peligrosa, llevan las situaciones al límite y sienten una subida de adrenalina al incumplir la ley.

#### Conducción cuidadosa (Careful driving)

Esta conducción es la que se realiza respetando las normas y conduciendo con seguridad.

---

<sup>40</sup>Dávid Pásztor. Forget about owning a car – the UX of future transportation. uxstudio (03-05-2016) Obtenido de [www.uxstudioteam.com](http://www.uxstudioteam.com) <https://uxstudioteam.com/ux-blog/the-ux-of-fully-autonomous-cars/>

<sup>41</sup> UX studio. The UX of Fully Autonomous Cars. Obtenido de [www.youtube.com](http://www.youtube.com) <https://www.youtube.com/watch?v=sC8jo874PRE>

<sup>42</sup>Hooft van Huysduynen (et al.) Measuring driving styles : a validation of the multidimensional driving style inventory. AutomotiveUI '15 Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications. Obtenido de [pure.tue.nl](http://pure.tue.nl) <https://pure.tue.nl/ws/files/3918433/585491304113011.pdf>

### Conducción con reducción de angustia (Distress-Reduction Driving)

Estos conductores realizan ejercicios de relajación muscular mientras conducen o otro tipo de actividades relajantes. Meditan y desean que termine el viaje mientras conducen.

### Conducción ansiosa (Anxious Driving)

Los conductores que se sienten nerviosos o se sienten frustrados al conducir pertenecen a esta categoría. Se preocupan cuando conducen con mala climatología y se sienten incómodos al conducir.

### Conducción disociada (Dissociated Driving)

Se produce cuando los conductores cometen deslices como intentar salir de los semáforos teniendo la tercera marcha engranada, olvidan que tienen activadas las luces largas hasta que otro conductor les hace una señal o accionan los limpiaparabrisas cuando quieren activar los intermitentes.

Estos grupos han sido formados por comparación y solo tienen sentido si se comparan entre ellos. No se intenta describir una realidad absoluta. Aunque también se identifican otros grupos la gran mayoría de los conductores participantes encaja en una de estas categorías. Se deben contemplar los diferentes estilos de conducción para evitar frustración en los usuarios y mejorar su experiencia con el producto.

Además de los diferentes perfiles de conducción<sup>43 44</sup> las interfaces de los vehículos autónomos tienen unas características específicas derivadas de la naturaleza de la conducción. Las interacciones deben ser rápidas por lo que las opciones se tienen que distinguir rápidamente.

Las opciones están contextualizadas para un acceso más rápido (menú climatización, menú entretenimiento, etc)

El salpicadero o la pantalla adecuada muestra la información necesaria en todo momento para poder tomar decisiones con la información adecuada.

Los menús ofrecen pocos niveles de profundidad siendo el promedio 3. Facilitan que el conductor vuelva su foco a la tarea de conducción rápidamente.

Las diferentes opciones ofrecen feedback generalmente visual para que el conductor identifique que la acción se ha realizado con éxito y no tenga dudas.

---

<sup>43</sup> Myoungsoon Jeon. UX Challenges and Opportunities of Autonomous Vehicles regarding Driving Styles and Automation Levels. Michigan Technological University. Obtenido de [www.auto-ui.org https://www.auto-ui.org/15/p/workshops/4/UX%20Challenges%20and%20Opportunities%20of%20Autonomous%20Vehicles%20regarding%20Driving%20Styles%20and%20Automation%20Levels.pdf](https://www.auto-ui.org/15/p/workshops/4/UX%20Challenges%20and%20Opportunities%20of%20Autonomous%20Vehicles%20regarding%20Driving%20Styles%20and%20Automation%20Levels.pdf)

<sup>44</sup> Stephen Moyers. Current UX Design Challenges for Driverless Cars. Digital Agency Network (5-12-2017) Obtenido de [digitalagencynetwork.com https://digitalagencynetwork.com/current-ux-design-challenges-for-driverless-cars/](https://digitalagencynetwork.com/current-ux-design-challenges-for-driverless-cars/)

## 9. Perfiles de usuario

Para poder orientar el diseño de una interfaz y para poder encontrar necesidades y nichos de mercado relacionados con los vehículos autónomos y la interacción dirigida por voz analizaremos cuál es el mercado actual de vehículos en el territorio nacional. Para soportar el trabajo posterior se sintetizará el mercado en diferentes personas<sup>45</sup> que se tomarán como referencia en el diseño de la propuesta.

España cuenta con 26.514.026<sup>46</sup> conductores y una población total de 46.549.045<sup>47</sup>. El 57% de la población tiene algún tipo de permiso de conducir y al 51% de los conductores no les gusta conducir<sup>48</sup>

El 56% de las nuevas matriculaciones según el informe de ANFAC<sup>49</sup> se produce en vehículos utilitarios y compactos.

Desde el anterior informe se puede apreciar un notable incremento de matriculaciones de SUVs (entre un 20% y un 40% de incremento)

También se aprecia un 15% de incremento en matriculaciones<sup>50</sup> de vehículos eléctricos de donde se puede extrapolar un interés en nuevos avances tecnológicos como podría ser el presentado en este trabajo.

La información anterior nos muestra un usuario promedio que dispone de un vehículo utilitario para uso diario cotidiano o SUV para uso familiar.

A partir de los datos y basándose en la característica de los vehículos con autonomía 5 obtenemos un nuevo grupo de usuarios (43% de la población) que podrían hacer uso de vehículos sin poseer un permiso de circulación

Principalmente se encuentra el usuario promedio, que no disfruta especialmente de la experiencia de conducción y utiliza el vehículo de forma pragmática.

---

<sup>45</sup>Wikipedia. Persona (user experience). Obtenido de Wikipedia

[https://en.wikipedia.org/wiki/Persona\\_\(user\\_experience\)](https://en.wikipedia.org/wiki/Persona_(user_experience))

<sup>46</sup>Diputación General de Tráfico. Estadísticas. Obtenido de [www.dgt.es](http://www.dgt.es)

<http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/censo-conductores/tablas-estadisticas/>

<sup>47</sup>Instituto Nacional de Estadística. Obtenido de [www.ine.es](http://www.ine.es)

[http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736176951&idp=1254735572981](http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176951&idp=1254735572981)

<sup>48</sup> Teresa Rubio. A 13 millones de conductores españoles no les gusta conducir. Cadena Ser. (02-11-2010) Obtenido de [cadenaser.com](http://cadenaser.com)

[http://cadenaser.com/ser/2010/11/02/sociedad/1288667609\\_850215.html](http://cadenaser.com/ser/2010/11/02/sociedad/1288667609_850215.html)

<sup>49</sup>Asociación Española de Fabricantes de Automóviles y Camiones. Informe anual. Obtenido de [www.anfac.com](http://www.anfac.com) <http://www.anfac.com/documents/tmp/MemoriaANFAC2017.pdf>

<sup>50</sup>Diputación General de Tráfico. Estadísticas. Obtenido de [www.dgt.es](http://www.dgt.es)

<http://www.dgt.es/es/seguridad-vial/estadisticas-e-indicadores/parque-vehiculos/tablas-estadisticas/>

Se han recogido dos personas que aunque no representan a la mayoría de usuarios es importante recoger como usuarios de los vehículos autónomos.

Una persona que no sabe conducir pero es un potencial usuario de vehículos con autonomía 5.

Una persona que disfruta de la conducción y por tanto podría no querer automatizar su experiencia.

Cada una de las personas podría clasificarse potencialmente dentro de un tipo de conducción relativo a la investigación de MDSI y aparece en la ficha.

## **Carmen – La práctica**

### **Tareas típicas**

Comprar ropa en el centro comercial  
Ir a eventos de la provincia

### **Datos**

34 años.  
Vive en Sevilla  
Soltera  
Oficinista

Carmen trabaja para el ayuntamiento local y posee estudios superiores. Dispone de un coche utilitario que utiliza para visitar a sus padres, ir a centros comerciales y realizar pequeñas excursiones.

Carmen no comprende cómo funcionan algunas partes de su coche y su conducción es muy tranquila. Aprobó el carnet de conducir a la primera, pero tiene dudas de algunas reglas de conducción.

No comprará otro coche hasta que sea necesario y espera que sea al menos dentro de 15 años.

Carmen disfruta cenando con sus amigas en modernos restaurantes de la ciudad.

Conducción con reducción de angustia (Distress-Reduction Driving)

## **Antonio – El padre de familia**

### **Tareas típicas**

Llevar a Juan al partido los domingos  
Realizar la compra mensual  
Llevar a Pedro y sus amigos al cine

### **Datos**

48 años

Madrid  
Casado con Maria y con dos hijos (Juan y Pedro)  
Operario industrial

Antonio ha cambiado recientemente su vehículo turismo a uno de tipo SUV que se acomode a las necesidades de la familia. Antonio está orgulloso de su SUV que no solo es práctico sino también bonito.

Aunque a Antonio no le disgusta conducir a veces se ve frustrado de tener que llevar a cada miembro de la familia a sus diferentes destinos. Maria dice que no se siente segura conduciendo un coche tan grande.

Al contrario que María, Antonio conduce con confianza, quizá demasiada. Su forma de conducir es más agresiva que el promedio y es el primero en incorporarse a las rotondas y el primero en salir en los semáforos.

Conducción riesgosa (Risky driving)

## **Paula – La adolescente**

### **Tareas típicas**

Volver a casa tarde  
Salir de viaje

### **Datos**

19 años  
Valencia  
Soltera  
Estudiante

Paula estudia diseño gráfico y todavía no ha conseguido aprobar el permiso de circulación. Se pone nerviosa y no es capaz de controlar el coche. Aunque rinde bien bajo presión en los estudios ve muy complicado coordinar todos los mandos de un coche.

Paula a veces se siente mal por tener que llamar a sus padres para que la vayan a recoger algunos fines de semana. También le gustaria conocer otros sitios como la montaña.

A Paula le gusta disfrutar del paisaje cuando su padre la lleva a sitios.

Conduccion disociada (Dissociated Driving).

## **Pelayo – El conductor premium**

### **Tareas típicas**

Hacer una salida el domingo por la mañana  
Ir a ferias de automovilismo

Presumir de su coche delante de conocidos

**Datos**

40 años

Segovia

Soltero

Empresario

Pelayo posee un coche de alta gama. Disfruta de su conducción e incluso de pensar cómo se está moviendo cada engranaje del coche a cada momento. Es un asiduo a las ferias automovilísticas, está suscrito a una revista de automoción y comparte afición con un grupo de amigos.

Pelayo disfruta de cada curva y cada juego de pedales. Se diría de su conducción que es normal.

A Pelayo le gusta estar al día de los avances tecnológicos que se incorporan al sector de la automoción.

Conduccion cuidadosa (Careful driving)

## 10. Propuesta de diseño

Se plantean tres niveles de interacción que pueden funcionar para todos los casos descritos anteriormente tanto niveles de autonomía como estilos de conducción.

### **Commandos - Nivel 0**

Es el nivel de abstracción más sencillo. Existe un mapeo 1:1 entre comandos y acciones a realizar.

Es un automatismo complejo. El vehículo no realiza acciones de forma autónoma.

Un ejemplo de interacción sería “Sube la ventanilla”

### **Asistente virtual - Nivel 1**

En este nivel nos encontramos cierta abstracción pero no se guarda el contexto de la interacción y las peticiones se tienen que encapsular en sí mismas y no afectan a posteriores interacciones.

“Dime la temperatura.

O dirígete a Barcelona

### **Interfaz conversacional - Nivel 2**

Es una interfaz conversacional cargada de inteligencia artificial. Este nivel podría reemplazar a un asistente personal con el que interactuar de forma más humana.

Un ejemplo:

Tengo calor

¿Quiere que suba la ventanilla?

Sí

Estos niveles de abstracción son configurables dependiendo de la experiencia que elija el usuario.

Pero tomaremos como referencia para el estudio la Interfaz conversacional (Nivel 2) dado que es la más compleja de ellas y que puede producir una experiencia más inmersiva.

La propuesta del prototipo puede activarse mediante un botón pulsador o una palabras para despertar configurables. Los botones centran la atención en la conducción y no en la propia interacción.

Después de haber analizado las funcionalidades de los coches autónomos y después de haber visto qué características determinan la interacción de voz en vehículos autónomos en la propuesta permite realizar las siguientes funcionalidades:

Encender y apagar el motor

Controlar la climatización

Bajar y subir ventanillas

Controlar la suspensión

Elegir destino

Controlar el reproductor de música y video

Controlar el teléfono móvil y realizar llamadas

Estacionamiento

Convocar

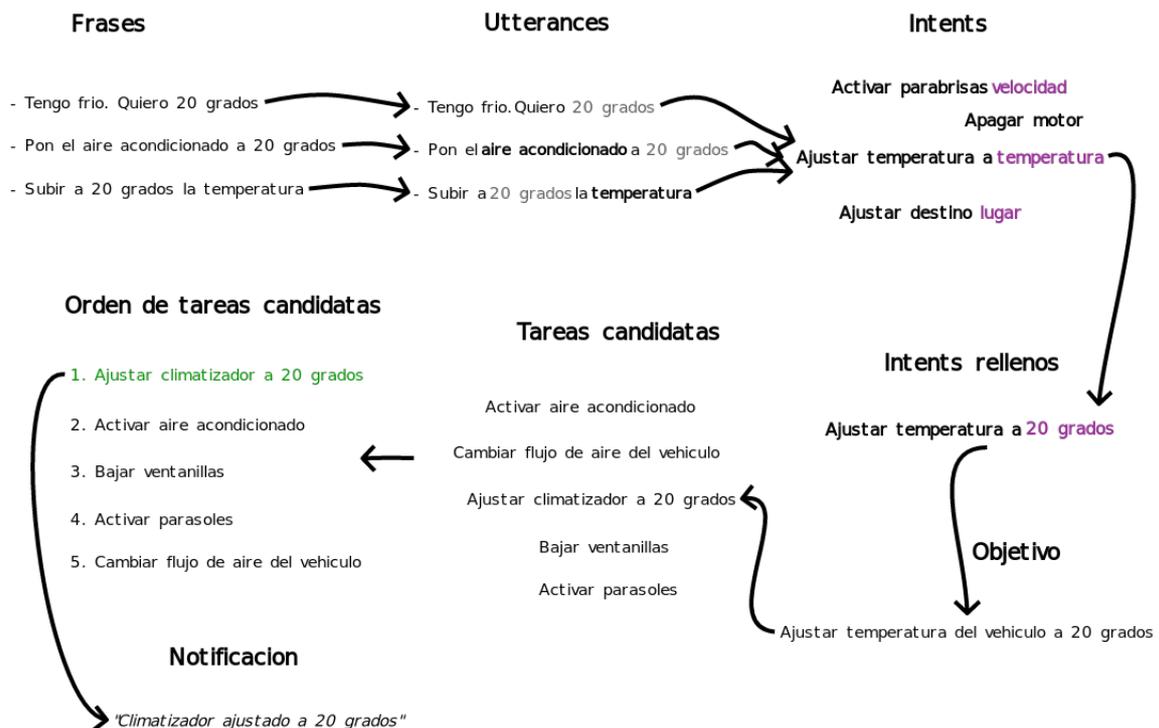
Mostrar información de los sensores

Posibilidad de modificación de la ruta con puntos intermedios

Accionamiento del claxon

Frenada de emergencia  
 Información del estado de los diferentes componentes del vehículo  
 Mostrar sugerencias durante el viaje

Aunque la interfaz de voz podría ser implementada en vehículos con cualquier nivel de autonomía el prototipo plantea una interfaz para vehículos con niveles de autonomía superiores al 3, especialmente 4 y 5.



## Roles

El prototipo canaliza la interacción a través de tres asistentes diferentes cuyo nombre es configurable, ya que este será la palabra de despertar de cada uno de ellos. Cada asistente es una voz diferente y emula una persona diferente. Cada uno de ellos se puede activar o desactivar de forma independiente para adaptar la experiencia a las diferentes personas.

### *Asistente de viaje*

Es la encargada de controlar el propio vehículo. Recoge las diferentes peticiones del usuario relativas al destino y al condicionamiento del vehículo como la climatización. Podría evocar a una azafata de vuelo.

### *Mecánico*

Es el encargado de informarnos sobre el estado actual del vehículo. Estado de la batería, capacidad de carga o de depósito, nivel de presión de los neumáticos etc. También recoge peticiones avanzadas como de ajuste de la suspensión. Algunas personas podrían ayudarse de ver esta voz como un mecánico independiente al resto de los asistentes.

## ***Informador***

Este rol se encarga de informarnos sobre el contexto exterior del vehículo y se le puede pedir información general. Podría informarnos por ejemplo de que estamos detenidos porque se ha producido un accidente en nuestra carretera o le podríamos preguntar a qué distancia encontraremos una gasolinera.

Cada uno de los roles posee diferente conocimiento y vocabulario que influirá en la propia interacción.

## **Objetivos**

En *Designing Human-Machine Interface for Autonomous Vehicles*<sup>51</sup> (International Federation of Automatic Control - 2016) aparece un modelo de toma de decisiones basado en objetivos. La propuesta comparte dicho modelo. Una alerta o estímulo genera una pregunta. A través de la información del entorno y del estado del sistema se diagnostica el estímulo y se predicen posibles consecuencias y posibles objetivos. Tras evaluar los diferentes objetivos y realizar una toma de decisiones se genera un estado objetivo hacia el que realizar una transición para que se genera una tarea y un procedimiento que la materializa.

Como también aparece recogido en el documento es importante que el usuario conozca este procedimiento para así poder tomar decisiones en el momento adecuado y de forma justificada. En la medida de lo posible la priorización de la toma de decisiones debería de ser priorizada por el usuario para que se ajuste a sus necesidades.

Es significativo ver como aparecen los objetivos como diferentes posibilidades y que solo se ejecutan procedimientos una vez que se ha evaluado toda la información disponible y se ha determinado que ese objetivo es el adecuado. Posteriormente se realiza de nuevo un proceso similar en el que varias tareas candidatas para la consecución del objetivo son analizadas hasta encontrar la más adecuada.

## **Seguridad**

El prototipo tiene como prioridad la seguridad del vehículo y sus ocupantes por lo que no realiza acciones que pudieran cuestionar la seguridad. Por ejemplo no ejecutará la acción de apagar el motor o abrir las puertas cuando el coche no esté detenido. Bajo la prioridad de la seguridad si algún sensor dejara de funcionar o el vehículo no supiera cómo responder éste solicitará al usuario que tome el control del vehículo seguida de una cuenta atrás tras la cual se desactivaba la conducción automática. En el caso de ser un vehículo con autonomía 5 que no necesariamente dispone de volante este se detendría o avanzaba despacio hasta poder proseguir el viaje de nuevo en condiciones de seguridad.

---

<sup>51</sup> S.Debernard (et al) *Designing Human-Machine Interface for Autonomous Vehicles*. IFAC. Obtenido de [www.sciencedirect.com](http://www.sciencedirect.com)  
<https://www.sciencedirect.com/science/article/pii/S2405896316322418>

## Memoria contextual

Se plantea la *memoria contextual* del prototipo como una memoria relativa a los contextos<sup>52</sup>. En la más sencilla de las interacciones conversacionales podemos ver como el gráfico de flujo conversacional y las futuras interacciones dependen de la presente y pasadas, pero cuando se realiza la acción deseada esta conversación desaparece y ya no vuelve a influir en las demás interacciones. Podría decirse que ha existido memoria intercontextual puesto que la inteligencia artificial ha recordado nuestra interacción, pero solo dentro de un contexto (una interacción). La memoria intercontextual se da cuando se recuerda información relativa a otra interacción e influye la interacción actual. Como ejemplos se plantea un usuario inseguro y que se ve abrumado cuando se le plantean demasiadas opciones. O en un periodo corto de tiempo se le ordena al vehículo a subir las ventanillas (porque existe un mal olor en el exterior) y posteriormente en otra interacción se le informa de que el usuario tiene calor. El vehículo podría plantear la opción de bajar las ventanillas pero al “recordar” que han sido subidas recientemente decide optar por sugerir activar el aire acondicionado.

## Integración

La integración se hace aumentando la interfaz actual. Los componentes mecánicos siguen funcionando por lo que no se le impide al usuario realizar las acciones por si mismo por seguridad. Las interfaces convencionales deben de funcionar de forma autónoma. Un fallo eléctrico o mecánico podría hacer que el sistema no funcionara adecuadamente por lo que se da la posibilidad de que el usuario pueda hacer uso de forma manual. Del mismo modo una información determinada podría no esta disponible en un determinado vehículo o un determinado comportamiento.

## Configuración

La configuración de la interacción por voz debería de realizarse a través de otra interfaz. Preferiblemente una pantalla táctil. Es una configuración que una vez se encuentra la adecuada debería de permanecer fija por un tiempo. La configuración contiene parámetros como:

### *Roles*

Posibilidad de habilitar y deshabilitar roles (Asistentes virtuales) según nuestras preferencias. Este contexto también nos permitiría configurar las palabras de despertar de los diferentes roles.

### *Niveles de interacción*

Posibilidad de elegir que tipo de interacción por voz queremos usar.

---

<sup>52</sup>Carlos E. Perez. The Next AI Milestone: Bridging the Semantic Gap. (6-4-2017) Obtenido de medium.com <https://medium.com/intuitionmachine/the-first-rule-of-agi-is-bc8725d21530>

### ***Prioridad en toma de decisiones***

Configuración relativa al estilo de conducción deseado y la prioridad en la toma de decisiones ante posibles conflictos

### ***Actualizaciones***

Permite añadir nuevos roles y actualizaciones

### ***Sonido***

Volumen  
Habilitar y deshabilitar alertas sonoras

### **Entradas**

Las entradas de la interacción vienen determinadas por la activación del Sistema de reconocimiento de voz y las diferentes frases que se digan en voz alta. En el apartado de Intents se puede ver un listado de posibles intents a las que se reducen las interacciones.

### **Salidas**

La principal salida de la interacción serán voces generadas por síntesis de voz. Puesto que tienen que coexistir diferentes roles es crucial que no se solapen sus voces produciendo ruido. Para ello deberán de turnarse para comunicarse con el usuario y que no se pierdan los mensajes. Si un mensaje de salida perdiera relevancia al permanecer largo tiempo en la cola de salida podría suprimirse.

Adicionalmente a las voces existen alertas sonoras que nos informan de diferentes situaciones. Configurando la cadencia, longitud, frecuencia, repeticiones y volumen se puede configurar un lenguaje sonoro propio con el que comunicar información, alerta, necesidad de toma de decisión etc al usuario. Un ejemplo sería cuando el vehículo no es capaz de continuar de forma autónoma. Se podría generar un sonido con el que el usuario ya este familiarizado y le transmita que tiene que tomar control del vehículo.

### **Ayuda**

No está contemplado que la interfaz de voz muestre ayuda de sí misma a través de la voz por la complejidad que esto implicaría. Si se contempla un análisis de los intents y cuando el intent que se solicita es muy parecido a alguno existente pero no hay una coincidencia exacta se le sugiere al usuario para que entienda que esa es la acción posible más parecida.

## Entidades

Las entidades serán los entes sobre los que pivotarán las interacciones del reconocimiento de voz. Aunque por su propia naturaleza son muy amplios los principales se recogen a continuación. Vienen determinados principalmente por los intents

- **Temperatura**
- **Numero**
- **TipoConduccion**
- **Lugar**
- **EstiloMusica**
- **ArtistaMusica**
- **Persona**
- **Posicion** (Izquierda|Derecha|Delante|Detras)
- **VariableEstado**
- **Contexto**

## Intents

Estos intents en esencia podrían ser los mapeos que se realizarían en el nivel 0 de interacción (comandos). Pero se aplican de la misma forma en el resto de los niveles. La inteligencia artificial reduce las estructuras de utterances hasta un intent concreto por lo que se pueden reutilizar para los diferentes niveles de interacción. A demás de los que aparecen a continuación existirían intents intermedios cuyo único fin sería interactuar con los diferentes roles para recabar información o vinculados a servicios externos.

- Encender motor
- Apagar motor
- Encender aire acondicionado
- Apagar aire acondicionado
- Ajustar temperature <Temperatura>
- Subir ventanilla
- Bajar ventanilla
- Subir ventanilla <Numero> (%)
- Bajar ventanilla <Numero> (%)
- Ajustar suspensión <Numero> (%)
- Ajustar suspensión <TipoConduccion>
- Ir a <Lugar>
- Reproducir musica
- Reproducir musica <EstiloMusica> | <ArtistaMusica>
- Parar musica
- Llamar a <Persona>
- Estacionar <Posición>
- Convocar
- Informacion <VariableEstado>
- Cambiar ruta pasando por <Lugar>
- Claxon

- Parar
- Sugerencias
- Sugerencias <Contexto>

## **Ampliaciones**

La propuesta esta abierta a posibles ampliaciones y debido al diseño modular es algo sencillo. Se podrían incluir nuevos roles con tareas específicas o “personalidades” específicas, nuevos componentes de hardware con los que interactuar e incluso nuevos servicios de terceros con información que nos permitan interactuar con otras plataformas.

# 11. Analisis de usabilidad

## Heurístico

Como analisis de usabilidad se realiza un analisis heurístico bajo los principios heurísticos de Nielsen. Aunque el objeto a tratar no es un producto web los principios se adecuan al contexto. Es posible adaptar las acciones como "mostrar" al contexto de la voz.

Cada principio heurístico se analiza y se relaciona en el contexto de la propuesta.

### *Visibilidad del estado del sistema.*

El estado del sistema se puede solicitar en cualquier momento. Así mismo si el estado del sistema se encuentra en un estado peligroso, o que requiere algún tipo de intervención humana también se muestra dicho estado a partir de pitidos identificables a estados.

Se muestra información también de forma visual justificando las acciones que se realizarán a continuación. No es viable incluir esta información por voz ya que ocuparían constantemente todo el escenario sonoro imposibilitando su utilización.

### *Adecuación entre el sistema y el mundo real.*

Este punto podría ser conflictivo ya que existen multitud de diversidades culturales y los lenguajes y palabras son diferentes en todas ellas por lo que no sería posible satisfacer a todas las personas. Dado que el vehículo debe de satisfacer a todas las personas se presentarían dos alternativas. Por un lado, permitir la configuración de las diferentes utterances. Esta solución es muy costosa a nivel de implementación puesto que las utterances serían dinámicas, habría que crear interfaces para que el usuario pudiera editarlas y además los usuarios deberían de tener conocimientos avanzados. La otra alternativa pasa por la normalización. En el mundo real la gente no utiliza palabras como "estacionamiento" o "claxon" pero sí que lo hace el reglamento de circulación. Este último aproximamiento es el que se ha realizado buscando un lenguaje neutro y fácilmente alcanzable por todas las personas.

La información aparece de forma que emula una conversación humana con la que el usuario está familiarizado.

### *Libertad y control por parte del usuario.*

Aunque el usuario es libre de repetir conversaciones o revertir las consecuencias actualmente no es posible cancelar una conversación una vez que se ha iniciado. Ésta podría ser una debilidad desde el punto de vista de la usabilidad. En las interfaces de voz más extendidas (las que se han analizado con anterioridad) no existe una solución común para detener los comandos de voz. Una solución aplicable sería mimetizar el comportamiento humano en que cuando una persona detiene el habla se interpreta que se ha confundido y comienza una nueva oración.

### *Consistencia y estándares.*

Se ha seguido una convención y aparece en la sección de intents, por ejemplo. A parte de eso las palabras son propias del modelo de negocio

Sí que es un problema que la propia convención no sea extensa por lo que potencialmente los usuarios deberían de hacerse con ella antes de comenzar a utilizar el producto.

### ***Prevención de errores.***

Dado que la interfaz de voz no bloquea los actuadores mecánicos no existe el problema de poder llegar a bloquear el vehículo. En el caso de que se produzcan errores éstos no realizan acciones de corrección, sino que se muestran de forma informativa. También el prototipo es capaz de sugerir el intent adecuado en el caso de que la similitud sea elevada o se carezca de algún parámetro (entidades).

### ***Reconocimiento antes que recuerdo.***

El sistema de interacción por voz se apoya para esta tarea en sistemas visuales que producen menor cansancio físico ya que es opcional buscar dicha información con la mirada y alternarla con otra información. Podría llegar a plantearse como un problema en sí mismo si solo dispusiéramos de la plataforma sonora. También en la medida de lo posible la interacción se ha intentado diseñar de formar similar a la interacción humana.

### ***Flexibilidad y eficiencia en el uso.***

Se contempla la posibilidad de acortar las conversaciones hasta el tamaño de los intents o realizar frases mas largas que nos resulten mas comodas. En cualquier caso el usuario tiene la decisión de elegir cual es su interacción.

### ***Diseño estético y minimalista.***

La interacción se ha diseñado ligada a los intents y aunque se contemplan comportamientos mas complejos para hacer mas realista la interacción, éstos están vinculados a los intents y no se contemplan interacciones carentes de objetivo.

### ***Ayuda a los usuarios a reconocer, diagnosticar y recuperarse de los errores.***

Los usuarios serán capaces de identificar los errores y determinar que acción fue la que los originó. De nuevo en algunos casos complejos será necesario apoyarnos en un sistema visual en el que mostrar información relativa al error si el sistema de síntesis de voz falla.

### ***Ayuda y documentación.***

Mas allá de la propia asistencia contextual que se aporta cuando faltan datos para completar un intent o las diferentes sugerencias que aportan los roles no existe una ayuda ya que puede hacer que el usuario se centre en dicha ayuda y pierda la atención de la acción de conducir si fuera necesaria en ese determinado momento. Si que se contempla incluir un manual de usuario en formato físico de papel con las diferentes posibilidades de interacción que se brindan.

## **Paseo cognitivo**

Dado que se han identificado los principales usuarios (personas) que harán uso de la propuesta se ha creído oportuno realizar paseos cognitivos con las diferentes personas. Este método nos desvelará posibles problemas de usabilidad o elementos de interacción que sean poco claros. Este método es especialmente útil en esta etapa donde todavía el prototipo se encuentra en fase de desarrollo. El objetivo no es evaluar el prototipo sino encontrar posibles problemas relacionados con una de las personas generadas.

Se van a realizar dos tareas:

### **Climatización: Acción habitual existente vehículo convencional.**

Ajustar la temperatura del vehículo

### **Conducción: Acción adaptada no existente en vehículo convencional**

Elección de destino: Centro comercial "El Aguila"

Durante el trayecto el prototipo notifica de falta de combustible.

Cambio de rumbo con parada en gasolinera.

Antes de usar el prototipo se realiza una sesión informativa con los usuarios dado que no es una experiencia común para éstos el haber utilizado un vehículo comandado por voz. Como simplificación y para evitar confusión cada rol no se ha reemplazado por el nombre propio que tendría en cada caso. Y no se han incluido las palabras de despertar por el mismo motivo, pero las personas si que las hubieran tenido que utilizar.

## ***Carmen – La práctica***

*Carmen ajusta el nivel de interacción a 1.*

### **Tarea 1: Ajustar temperatura**

Carmen: "Poner la temperatura a 20 grados".

Asistente de viaje: "Quiere 'Ajustar la temperatura a 20 grados'".

Carmen: "Sí".

### **Tarea 2: Conducción**

Carmen: "Arrancar motor"

Carmen: "Fijar destino Centro comercial "El Aguila".

...

Mecanico: “El vehículo necesita repostar porque ha comenzado el depósito de reserva”.

Carmen: “Para”.

*El coche se detiene lentamente en condiciones de seguridad.*

Carmen: “Ir a la gasolinera”.

Asistente de viaje: “No le he entendido”

*Carmen se pone nerviosa*

Carmen: “¿Cuál es la gasolinera mas cercana?”.

Informador: “Via de servicio El caminillo”.

Carmen: “Fijar destino Via de servicio El caminillo”.

Asistente de viaje: “Destino fijado”.

*El vehiculo llega allí y Carmen repuesta*

Carmen: “Fijar destino Centro comercial El aguila”.

### ***Antonio – El padre de familia***

*Antonio lee el manual del vehiculo y configura nivel 2*

#### **Tarea 1: Ajustar temperatura**

Antonio: “Activar aire acondicionado”.

Antonio Sonrie

Antonio: “Ajustar la temperatura a 20 grados”.

*Vuelve a sonreir orgulloso*

#### **Tarea 2: Conducción**

Antonio: “Fijar destino Centro comercial El Aguila”.

Asistente de viaje: “Destino fijado”.

*Antonio espera durante un rato y tras unos minutos y comienza a buscar interruptores. Finalmente arranca el coche de forma manual.*

...

Antonio: "Acelera"

Asistente de viaje: "No está permitido. Velocidad máxima de la vía".

Antonio: "Ves mas rapido"

Asistente de viaje: "No está permitido. Velocidad máxima de la vía".

*Antonio desactiva la conducción autónoma y continúa el recorrido de forma manual.*

### ***Paula – La adolescente***

*Configura interacción a nivel 2*

#### **Tarea 1: Ajustar temperatura**

Paula: "Aire acondicionado".

Asistente de viaje: "¿Quiere encender el aire acondicionado?".

Paula: "20 grados".

Asistente de viaje: "¿Quiere Ajustar la temperatura a 20 grados?".

Paula: Si.

#### **Tarea 2: Conducción**

Paula: "Arrancar motor"

Paula: "Quiero ir a Centro comercial El Aguila"

Asistente de viaje: "Destino fijado".

...

Mecanico: "El vehiculo necesita repostar porque ha comenzado el deposito de reserva".

Paula: ¿Qué tengo que hacer?

Mecanico: "Debe repostar en la próxima estación de servicio"

Paula: "¿Cuál es la próxima estación de servicio?"

Informador: "Vía de servicio El caminillo".

Paula: “Repostar en Via de servicio El caminillo?”

...

*El vehículo llega allí y Paula repuesta*

Asistente de viaje: “Fijar destino Centro comercial El aguila”.

*Paula se sorprende y se acomoda*

### ***Pelayo – El conductor premium***

*Pelayo ajusta el nivel de interacción a 0 y desactiva la conducción autónoma.*

#### **Tarea 1: Ajustar temperatura**

Pelayo: "Ajustar temperatura 20 grados"

#### **Tarea 2: Conducción**

*Pelayo comienza a conducir de forma manual.*

....

Mecánico: “El vehículo necesita repostar porque ha comenzado el depósito de reserva”.

*Pelayo frunce el ceño y se dirige a la estación más próxima que aparece indicada en la señalización de la carretera*

### **Resultados de los Paseos cognitivos**

En la interacción de Carmen podemos ver cómo en un primer momento ésta solicita una acción de parar cuando no hubiera sido necesario. Probablemente para pensar una estrategia alarmada pensando que se trata de un problema de alta urgencia. En este caso si la interfaz hubiera pedido confirmación hubiera sido más exitosa la interacción. Como solución se podría crear un grupo de intents para los que se solicitara confirmación.

Posteriormente cuando quiere reanudar la marcha parece olvidar que no se encuentra ante un ente realmente inteligente. Quizá un nivel 2 de interacción le hubiera proporcionado más información, pero desde luego un nivel 1 no tiene suficiente información como para realizar la acción, por lo que Carmen tiene que consultar el destino.

En la interacción de Antonio vemos como se genera una discursión con el vehículo que genera frustración en Antonio y decide terminar su interacción vocal con el vehículo. Antonio quiere conducir a mayor velocidad de la permitida pero el vehículo no se lo permite porque entiende que esto puede poner en peligro a Antonio.

En la sesión de Paula podemos ver como el nivel de interacción ofrece sugerencias e incluso es capaz de almacenar el destino inicial de Paula. También en la primera tarea vemos como el vehículo ofrece encender el aire acondicionado, pero en su lugar Paula ajusta la temperatura. Quizá Paula no sepa si es la misma acción o son acciones diferentes. En este caso el usuario ha conseguido su objetivo por lo que en este caso concreto la omisión de información no repercute negativamente. Paula interactúa con los diferentes roles y le permite contextualizar la información. Esto le ayuda a tomar decisiones.

La interacción de Pelayo es muy primitiva pues intencionalmente configura el vehículo para una experiencia más manual. No obstante, utiliza la interfaz de voz para ajustar el climatizador y es informado exactamente en el momento exacto cuando el coche comienza a utilizar el depósito de reserva.

## 12. Conclusión

Hemos recorrido en este trabajo el origen de los vehículos autónomos y su evolución, explicado los 6 niveles de autonomía que recoge la SAE International y recorrido los tipos de experiencias existentes, desde las basadas en comandos hasta las más inmersivas interfaces conversacionales.

Se han descrito la situación actual de dispositivos que interactúan mediante la voz, cuáles son los principales impulsores de vehículos autónomos y las características principales que hay que tener en cuenta para diseñar una interfaz de voz para un vehículo autónomo.

Se han mostrado las diferentes formas de interacción que presenta el prototipo, así como una descripción general de cómo se interactúa con los vehículos autónomos dirigidos por voz en el apartado de Diseño de interacción.

Se ha realizado un análisis explorativo de la interacción con vehículos autónomos. Se ha mostrado como los diferentes fabricantes de vehículos autónomos han diseñado la experiencia de interacción y que en la actualidad ninguno de ellos incorpora interacción mediante voz. Se ha mostrado la propuesta de diseño en la que se incluye la voz de UX Studio así como Dragon Drive de Nuance Communications que se adapta como solución de interacción de voz en vehículos. Se han descrito los diferentes tipos de conducción que identificó la Universidad de Eindhoven.

En el apartado de perfiles se han recogido datos estadísticos que perfilan un público objetivo de vehículos autónomos comandados por interfaces de voz. Posteriormente se han desarrollado cuatro personas basadas en los datos estadísticos: Dos perfiles consumidores de vehículos convencionales, otro sin permiso de circulación y otro con un carácter de conducción más conservador.

Se plantea una solución basada en interfaces de voz con objetivo de mejorar la experiencia de los usuarios en los vehículos autónomos. Dicha propuesta es capaz de trabajar con diferentes niveles de interacción ajustables por el usuario. El nivel más avanzado de interacción cuenta con diferentes roles(asistentes) para contextualizar las interacciones. El vehículo autónomo toma las decisiones en función de objetivos y prevalece la seguridad como prioridad.

El nivel más avanzado utiliza la memoria contextual para memorizar las interacciones recientes e incluirlas en la toma de decisiones. Se describen cuáles son las entidades y los intents principales y también se describe posibilidades de ampliación de la interfaz.

En el análisis heurístico de la propuesta se han encontrado puntos débiles como la dependencia de otras interfaces principalmente visuales para mostrar errores críticos en casos excepcionales. Los paseos cognitivos basados en personas han evidenciado comportamientos críticos que requieren modificación de la propuesta. El Diseño Centrado en el Usuario requiere posteriores interacciones para resolver dichas carencias, pero se escapa del alcance de este trabajo.

Como resultado del desarrollo del proceso se han detectado diferentes nichos a explotar en la convergencia planteada en este trabajo: El potencial mercado de personas que no pueden conducir en la actualidad, pero sí que podrían con un vehículo de nivel 5 de autonomía que no requiera de conductor. Y la personalización de la experiencia (roles, servicios externos, idiomas, etc) para diferentes personas y servicios (vehículo propio, alquiler de vehículo, taxi autónomo etc). Así

mismo se han identificado necesidades a satisfacer que no han sido desarrolladas en este trabajo por su complejidad. En uno de los paseos cognitivos se ha visto como se ha llegado a producir una discusión entre la interfaz, intentando preservar la seguridad y legislación vigente, y el usuario. Existe pues una discusión abierta sobre la normalización de este tipo de cuestiones que se solapan con la ética y el derecho. Es una industria compleja puesto que las decisiones y sus acciones derivadas llegan a no ser realizadas por el usuario. También es necesario mejorar las técnicas de inteligencia artificial para acotar los posibles errores derivados de ésta. Las interfaces dirigidas por voz requieren de una mayor inclusión en la cultura, que cree comportamientos y símbolos fácilmente identificables para que la experiencia no sea tan abrupta y resulte más natural.

Este trabajo me ha ayudado a asentar los conocimientos aprendidos durante el estudio del grado, principalmente las relacionadas con el perfil de optatividad de Usabilidad e Interfaces (Comportamiento de usuarios, Usabilidad y Diseño de interacción). He aprendido las características tanto de la investigación como de la edición de documentos académicos. He descubierto la trasdisciplinariedad de la usabilidad y accesibilidad que pueden aplicarse a diferentes medios y soportes.

## **Anexo 1. Entregables del proyecto**

Memoria del proyecto.

Presentacion del proyecto para publico general.

Presentación del proyecto para el tribunal evaluador.

Autoinforme evaluador.

## Bibliografía adicional

[https://es.wikipedia.org/wiki/Interfaz\\_mediante\\_voz\\_del\\_usuario](https://es.wikipedia.org/wiki/Interfaz_mediante_voz_del_usuario)

[https://en.wikipedia.org/wiki/Autonomous\\_car](https://en.wikipedia.org/wiki/Autonomous_car)

<https://www.fastcompany.com/40516928/gm-is-leading-the-self-driving-car-race-while-tesla-lags-far-behind-report-says>

<http://www.businessinsider.com/why-automakers-have-advantage-in-building-self-driving-car-race-2018-1>

<https://www.theverge.com/2018/1/16/16893452/detroit-auto-show-2018-google-gm-waymo-ford-tesla>

<https://arstechnica.com/cars/2018/01/why-analysts-put-gm-and-waymo-far-ahead-of-tesla-in-driverless-car-race/>

<https://www.teslarati.com/tesla-finishes-last-navigant-research-self-driving-tech/>

<https://www.theguardian.com/technology/self-driving-cars>

[https://issuu.com/rudermanfoundation/docs/self\\_driving\\_cars\\_-\\_the\\_impact\\_on\\_p](https://issuu.com/rudermanfoundation/docs/self_driving_cars_-_the_impact_on_p)

[https://es.wikipedia.org/wiki/Automatizaci%C3%B3n\\_vehicular](https://es.wikipedia.org/wiki/Automatizaci%C3%B3n_vehicular)

[https://es.wikipedia.org/wiki/Veh%C3%ADculo\\_aut%C3%B3nomo](https://es.wikipedia.org/wiki/Veh%C3%ADculo_aut%C3%B3nomo)

<https://www.20minutos.es/noticia/2825372/0/clasificacion-coches-autonomos/>

<https://blogthinkbig.com/historia-de-los-coches-autonomos>

[https://www.pcworld.com/article/243060/speech\\_recognition\\_through\\_the\\_decades\\_how\\_we\\_ended\\_up\\_with\\_siri.html](https://www.pcworld.com/article/243060/speech_recognition_through_the_decades_how_we_ended_up_with_siri.html)

<https://stories.devacademy.la/interfaces-conversacionales-y-chatbots-14eeecd0a>

<https://blogthinkbig.com/reconocimiento-de-voz>

<https://www.beeva.com/beeva-view/innovacion/interfaces-controladas-mediante-voz-arquitectura/>

<https://hablandoesalud.wordpress.com/2017/02/20/la-voz-el-interfaz-universal/>

<http://uxpanol.com/teoria/interfaces-conversacionales-aplicacion-y-ejemplos/>

<https://www.teamlewis.com/es/lewis-blog/articles/interfaz-conversacional-chatbots>

<http://uxpanol.com/teoria/experiencia-de-usuario-e-interfaces-conversacionales-una-introduccion/>

<https://blogthinkbig.com/interfaces-conversacionales-las-tres-fases-de-la-tecnologia>

<http://www.torresburriel.com/weblog/2016/04/11/introduccion-al-diseno-de-interfaces-conversacionales/>

<https://www.mockplus.com/blog/post/ux-persona-examples>

Programming Voice Interfaces de Walter Quesada & Bob Lautenbach O'Reilly Press 2018