

INTERACCION DE ASISTENTE PERSONAL DIGITAL CON PERSONIFICACIONES Y MULTIMEDIA RICA EN RESPUESTAS

ANTECEDENTES

5

La personificación de interacción de computadora es una práctica muy común en el área de inteligencia artificial. Los marcadores de asistentes personales digitales pueden imbuir a sus asistentes con una personalidad para así mantener el nivel de compromiso con el usuario muy alto y también para establecer un elemento de confianza entre el usuario y el asistente, de esa forma facilitando el uso continuo y la interacción con el asistente personal digital.

Una técnica para mantener el nivel de alto compromiso conlleva a hacer las respuestas proporcionadas por el asistente personal digital divertidas y entretenidas. Algunas implementaciones convencionales de los asistentes personales digitales se programan para generar respuestas de texto divertidas cuando un usuario hace preguntas que tienen un tono casual. Sin embargo, los asistentes personales digitales convencionales típicamente no aprovechan la flexibilidad completa de los lienzos digitales cuando presentan respuestas alegres. Ninguno de ellos aprovecha la energía de texto por voz de día moderno que sintetiza las técnicas para sonar gracioso o diferente cuando proporciona las respuestas. Además, no existe representación visual del asistente digital que puede utilizarse

para transmitir cualquier aspecto de una respuesta alegre.

BREVE DESCRIPCION DE LA INVENCION

5 Se describen aquí sistemas, métodos, aparatos, y productos de programa de computadora para implementar un asistente personal digital. El asistente personal digital es capaz de determinar que un usuario ha preguntado una pregunta o hace una declaración que pretende involucrarse un personaje del asistente personal digital en
10 oposición a, por ejemplo, solicitar que el asistente personal digital obtenga información o realice alguna otra tarea en beneficio del usuario. En respuesta a determinar que el usuario ha preguntado dicha pregunta o hace tal declaración, el asistente personal digital proporciona una respuesta de la misma al presentar o reproducir
15 después un objeto multimedia asociado con una referencia de cultura popular dentro o por una interfaz de usuario del asistente personal digital. Adicional o alternativamente, en respuesta a determinar que el usuario ha preguntado a manera de una pregunta o hace tal declaración, el asistente personal digital proporciona la respuesta
20 del mismo al generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con la referencia de cultura popular. Incluso además, el asistente personal digital puede proporcionar la respuesta al presentarse dentro de la interfaz de usuario del texto de asistente personal digital que
25 comprende una cita asociada con la referencia de cultura popular,

presentándose dentro de la interfaz de usuario del asistente personal digital como una representación visual del asistente personal digital que invoca la referencia de cultura popular, y/o que presenta dentro de la interfaz de usuario del asistente personal digital un enlace que
5 puede activarse por el usuario para acceder al contenido asociado con la referencia de cultura popular.

En particular, se describe aquí un método para implementar un asistente personal digital. De acuerdo con el método, se recibe una representación digital de una pronunciación de un usuario del
10 asistente personal digital. Basándose en al menos un análisis de la representación digital de la pronunciación, se determina que la pronunciación comprende una pronunciación prevista para comprometer a un personaje del asistente personal digital. En respuesta a al menos la determinación de que la pronunciación
15 comprende una pronunciación prevista para comprometer a la persona del asistente personal digital, el asistente personal digital se causa que genere una respuesta a la pronunciación. La respuesta a la pronunciación incluye un objeto multimedia (por ejemplo, imagen, contenido de vídeo o contenido de audio) que se presenta
20 dentro o se reproduce después por una interfaz de usuario del asistente personal digital y que está asociado con una referencia de cultura popular. Alternativa o adicionalmente, la respuesta a la pronunciación incluye voz que se genera o reproduce después por el asistente personal digital y que comprende una personificación de
25 una voz una persona asociada con la referencia de cultura popular.

En una modalidad del método anterior, la respuesta a la pronunciación además incluye texto que se presenta dentro de la interfaz de usuario del asistente personal digital, el texto comprende una cita asociada con la referencia de cultura popular.

- 5 En otra modalidad, la voz comprende una cita asociada con la referencia de cultura popular.

Incluso en otra modalidad, la respuesta a la pronunciación adicional incluye una representación visual del asistente personal digital que se presenta dentro de la interfaz de usuario del mismo y
10 que invoca la referencia de cultura popular. Además de acuerdo con tal modalidad, la representación visual del asistente personal digital puede comprender una animación del asistente personal digital.

Incluso en otra modalidad, la respuesta a la pronunciación adicional incluye un enlace que se presenta dentro de la interfaz de
15 usuario del asistente personal digital que puede activarse por el usuario para acceder al contenido asociado con la referencia de cultura popular.

Dependiendo de la implementación del método anterior, que causa que la respuesta a la pronunciación se genere puede incluir
20 identificar una pluralidad de respuestas elegibles a la pronunciación y selección de la respuesta para la pronunciación dentro de entre la pluralidad de respuestas elegibles para la pronunciación.

De acuerdo con otra implementación, que causa que la respuesta a la pronunciación se genere incluye coincidir la
25 pronunciación a un tipo de pronunciación particular dentro de un

árbol jerárquico de tipos de pronunciación, cada tipo de pronunciación en el árbol jerárquico de tipos de pronunciación tiene una o más respuestas asociadas del mismo, y que seleccionan la respuesta la pronunciación de entre la respuesta(s) asociada con el tipo de pronunciación particular.

De acuerdo con una implementación adicional, que causa que la respuesta a la pronunciación se genere incluye determinar que la pronunciación está asociada con una tópico de tendencia y en respuesta a determinar que la pronunciación está asociada con la tópico de tendencia, seleccionar la respuesta a la pronunciación de entre una o más respuestas asociadas con la tópico de tendencia.

De acuerdo incluso con una implementación adicional, que causa que la respuesta a la pronunciación se genere incluye determinar que la pronunciación es una pronunciación para la cual existe una o más respuestas predefinidas previstas para transmitir el personaje del asistente personal digital, y en respuesta a determinar que la pronunciación es una pronunciación para la cual existen una o más respuestas predefinidas previstas para transmitir el personaje del asistente personal digital, seleccionar la respuesta a la pronunciación de entre una o más respuestas predefinidas.

En otra modalidad del método anterior, que causa que la respuesta a la pronunciación se genere incluye enviar un archivo de audio que incluye la voz o información que identifica el archivo de audio para un dispositivo de cómputo que ejecuta el asistente personal digital.

Incluso en otra modalidad, que causa que la respuesta a la pregunta de la pronunciación se genere incluye proporcionar texto de un dispositivo de cómputo que ejecuta el asistente personal digital, el texto será procesado por un componente de texto a voz del asistente personal digital para generar la voz.

También se describe aquí un dispositivo de cómputo. El dispositivo de cómputo incluye al menos un procesador y una memoria. La memoria almacena lógica de programa de computadora para ejecución por al menos un procesador. La lógica de programa de computadora incluye uno o más componentes configurados para realizar operaciones cuando se ejecutan por al menos un procesador. El uno o más componentes incluyen un asistente personal digital. El asistente personal digital está configurado para capturar audio que representa una pronunciación de un usuario y para transmitir el audio a un respaldo de asistente personal digital. La pronunciación comprende una pronunciación prevista para comprometerse con un personaje de asistente personal digital. El asistente personal digital además está configurado para proporcionar una respuesta a la pronunciación basándose al menos en la información recibida del respaldo de asistente personal digital. Proporcionar la respuesta incluye generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con una referencia de cultura popular.

En una modalidad del dispositivo de cómputo, el asistente personal digital se configura para reproducir después un archivo de

audio que incluye la voz. En otra modalidad del dispositivo de cómputo, el asistente personal digital está configurado para aplicar conversión de texto a voz al texto para generar la voz.

En otra modalidad del dispositivo de cómputo, proporcionar la
5 respuesta además incluye al menos uno de presentar o reproducir
después un objeto multimedia por una interfaz de usuario del
asistente personal digital, el objeto multimedia estando asociado con
la referencia de cultura popular, que presenta texto dentro de la
interfaz de usuario del asistente personal digital, el texto comprende
10 una cita asociada con la referencia de cultura popular, que presenta
una representación visual del asistente personal digital que invoca la
referencia de cultura popular dentro de la interfaz de usuario del
asistente personal digital, y que presenta un enlace dentro de la
interfaz de usuario del asistente personal digital que puede activarse
15 por el usuario para acceder al contenido asociado con la referencia
de cultura popular.

Se describe aquí otro dispositivo de cómputo. El dispositivo de
cómputo incluye al menos un procesador y una memoria. La memoria
almacena lógica de programa de computadora para ejecución por al
20 menos un procesador. La lógica de programa de computadora incluye
uno o más componentes configurados para realizar operaciones
cuando se ejecutan por al menos un procesador. El uno o más
componentes incluyen un asistente personal digital. El asistente
personal digital está configurado para capturar audio que representa
25 una pronunciación de un usuario y para transmitir el audio a un

respaldo de asistente personal digital. La pronunciación comprende una pronunciación prevista para comprometerse con un personaje del asistente personal digital. El asistente personal digital además está configurado para proporcionar una respuesta a la pronunciación basándose al menos en la información recibida del respaldo de asistente personal digital. Proporcionar la respuesta incluye presentar o reproducir después un objeto multimedia asociado con una referencia de cultura popular por una interfaz de usuario del asistente personal digital.

10 En una modalidad del dispositivo de cómputo, el objeto multimedia comprende una imagen, contenido de vídeo, o contenido de audio.

 En otra modalidad del dispositivo de cómputo, proporcionar la respuesta además incluye al menos uno de generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con la referencia de cultura popular, presentar el texto dentro de la interfaz de usuario del asistente personal digital, el texto comprende una cita asociada con la referencia de cultura popular, presentar una representación visual del asistente personal digital que invoca la referencia de cultura popular dentro de la interfaz de usuario del asistente personal digital, y presentar un enlace dentro de la interfaz de usuario del asistente personal digital que puede activarse por el usuario al acceder al contenido asociado con la referencia de cultura popular.

25 Esta breve descripción se proporciona para introducir una

selección de conceptos en una forma simplificada que además se describe a continuación en la descripción detallada. Esta breve descripción no pretende identificar características clave o características esenciales del tema reclamado, ni pretende utilizarse para limitar el alcance del tema reclamado. Además, se observará que el tema reclamado no se limita a las modalidades específicas descritas en la descripción detallada y/u otras secciones de este documento. Tales modalidades se presentan aquí para propósitos ilustrativos únicamente. Modalidades adicionales se apreciarán por aquellos expertos en la técnica(s) relevante basándose en las técnicas aquí contenidas.

BREVE DESCRIPCION DE LOS DIBUJOS

Los dibujos acompañantes, los cuales se incorporan aquí y forman parte de la especificación, ilustran modalidades de la presente invención y, junto con la descripción, sirven además para explicar los principios de la invención y para permitir a un experto en la técnica(s) relevante hacer y utilizar la invención.

La Figura 1 es un diagrama de bloques de un sistema ilustrativo que implementa un asistente personal digital que utiliza personificaciones y multimedia en respuesta a la pronunciación de tipo charla de acuerdo con una modalidad ilustrativa.

La Figura 2 es un diagrama de bloques que muestra cómo varios componentes del sistema de la Figura 1 operan para permitir a

un asistente personal digital determinar que un usuario haya hecho una pronunciación de tipo charla y para proporcionar alguna respuesta del mismo.

La Figura 3 describe un árbol jerárquico ilustrativo de tipos de pronunciación que pueden utilizarse para seleccionar una respuesta
5 para una pronunciación de tipo charla de acuerdo con una modalidad.

La Figura 4 es un diagrama de bloques que muestra una respuesta ilustrativa que puede proporcionarse por un asistente personal digital en respuesta a una pronunciación de tipo charla de
10 acuerdo con una modalidad.

Las Figuras 5-10 proporcionan ejemplos ilustrativos de respuestas para pronunciaciones de tipo charla que pueden entregarse a través de la interfaz de usuario de un asistente personal digital de acuerdo con varias modalidades.

La Figura 11 describe un cuadro de flujo de un método para
15 implementar un asistente personal digital que utiliza personificaciones y/o multimedia en respuesta a las pronunciaciones de tipo charla de acuerdo con una modalidad.

La Figura 12 describe un cuadro de flujo de un método por el
20 cual un asistente personal digital proporciona una respuesta para una pronunciación de tipo charla que incluye una personificación por voz de acuerdo con una modalidad.

La Figura 13 describe un cuadro de flujo de un método por el
cual un asistente personal digital proporciona una respuesta para
25 una pronunciación de tipo charla que incluye un objeto multimedia de

acuerdo con una modalidad.

La Figura 14 es un diagrama de bloques de un dispositivo móvil ilustrativo que puede utilizarse para implementar varias modalidades.

La Figura 15 es un diagrama de bloques de un sistema de
5 computadora basado en procesador ilustrativo que puede utilizarse para implementar varias modalidades.

Las características y ventajas de la presente invención se volverán más evidentes de la descripción detallada descrita a continuación cuando se toman en conjunto con los dibujos, en donde
10 caracteres de referencia similares identifican elementos correspondientes a través de éste. En los dibujos, números de referencia similares generalmente indican elementos de funcionalidad idéntica, similar, y/o estructuralmente similar. El dibujo en el cual un elemento aparece primero se indica por el dígito(s) del
15 extremo izquierdo en el número de referencia correspondiente.

DESCRIPCION DETALLADA

1. Introducción

20 La siguiente descripción detallada se refiere a los dibujos acompañantes que ilustran modalidades ilustrativas de la presente invención. Sin embargo, el alcance de la presente invención no se limita a estas modalidades, pero en vez de eso se define por las reivindicaciones anexas. De esa forma, las modalidades más allá de
25 aquellas mostradas en los dibujos acompañantes, tal como versiones

modificadas de las modalidades ilustradas, pueden sin embargo abarcarse por la presente invención.

Las referencias en la especificación para ñuna modalidadò, ñmodalidadò, ñuna modalidad ilustrativaò, o similares, indican que la
5 modalidad descrita puede incluir un aspecto, estructura, o característica particular, pero cada modalidad no necesariamente puede incluir el aspecto, estructura, o característica particular. Además, tales frases no necesariamente se refieren a la misma modalidad. Además, cuando un aspecto, estructura, o característica
10 particular se describe en conexión con una modalidad, esto se omite ya que está dentro del conocimiento de los expertos en la técnica para implementar tal aspecto, estructura, o característica en conexión con otras modalidades ya sea o no se describe explícitamente.

15 Se describen aquí sistemas, métodos, aparatos, y productos de programa de computadora para implementar un asistente personal digital. El asistente personal digital es capaz de determinar que un usuario ha hecho una pregunta o hace una declaración que pretende involucrarse con un personaje del asistente personal digital en
20 oposición a, por ejemplo, solicitar que el asistente personal digital obtenga información o realice alguna otra tarea en beneficio del usuario. En respuesta a determinar que el usuario ha preguntado tal pregunta o ha hecho tal declaración, el asistente personal digital proporciona una respuesta del mismo para presentar o reproducir
25 después un objeto multimedia asociado con una referencia de cultura

popular dentro o por una interfaz de usuario del asistente personal digital. Adicional o alternativamente, en respuesta a determinar que el usuario ha hecho tal pregunta o ha hecho tal declaración, el asistente personal digital proporciona la respuesta del mismo al

5 generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con la referencia de cultura popular. Incluso además, el asistente personal digital puede proporcionar una respuesta al presentar dentro de la interfaz de usuario texto de asistente personal digital que comprende

10 una cita asociada con la referencia de cultura popular, presentar dentro de la interfaz de usuario del asistente personal digital una representación del asistente personal digital que invoca la referencia de cultura popular, y/o presentar dentro de la interfaz de usuario digital del asistente personal digital un enlace que puede activarse

15 por el usuario para acceder al contenido asociado con la referencia de cultura popular.

Al activar el asistente personal digital para responder a las pronunciaciones del usuario que pretenden involucrarse con su personaje (también se refiere aquí como "pronunciaciones de tipo

20 charla") en la forma anterior, las modalidades aquí descritas pueden aumentar de manera ventajosa el nivel de compromiso entre el usuario y el asistente personal digital y también establecer un elemento de confianza entre el usuario y el asistente, de esa forma facilitar el uso continuo de interacción con el asistente personal

25 digital. Por ejemplo, al proporcionar respuestas que incluyen objetos

multimedia, personificaciones por voz, citas, y enlaces asociados con las referencias de cultura populares que probablemente se reconocerán y/o apreciarán por el usuario, el asistente personal digital puede tanto entretener como establecer un sentido de
5 uniformidad con el usuario.

La sección II describe un sistema ilustrativo que implementa un asistente personal digital que utiliza personificaciones y/o multimedia en respuesta a las pronunciaciones de tipo charla de acuerdo con las modalidades. La sección III describe métodos
10 ilustrativos para implementar un asistente personal digital que utiliza las personificaciones y/o multimedia en respuesta a las pronunciaciones tipo charla de acuerdo con las modalidades. La sección IV describe un dispositivo móvil ilustrativo que puede utilizarse para implementar un asistente personal digital de acuerdo
15 con las modalidades aquí descritas. La sección V describe una computadora de escritorio ilustrativa que puede utilizarse para implementar un asistente personal digital de acuerdo con las modalidades aquí descritas. La sección VI proporciona algunas observaciones finales.

20

II. Sistema ilustrativo que implementa un asistente digital personal que utiliza personificaciones y/o multimedia en respuestas

La Figura 1 es un diagrama de bloques de un sistema
25 ilustrativo 100 que implementa un asistente personal digital que

utiliza personificaciones y multimedia en respuesta a las pronunciaciones de tipo charla de acuerdo con una modalidad ilustrativa. Como se muestra en la Figura 1, el sistema 100 incluye un dispositivo de cómputo de usuario final 102 que se conecta de manera comunicativa a un respaldo de asistente personal digital 106 a través de una o más redes 104. Cada uno estos componentes ahora se describirán.

El dispositivo de cómputo de usuario final 102 pretende representar un dispositivo electrónico basado en procesador que es capaz de ejecutar un asistente personal digital 130 basado en *software* que se instala en el mismo. El asistente personal digital 130 puede ejecutarse en beneficio de un usuario y del dispositivo de cómputo de usuario final 102. En una modalidad, el dispositivo de cómputo de usuario final 102 comprende un dispositivo de cómputo móvil tal como un teléfono móvil (por ejemplo, un teléfono inteligente), una computadora laptop, una computadora de tableta, una netbook, una computadora portátil tal como un reloj inteligente o una computadora montada en la cabeza, un reproductor de medios portátil, una consola de juegos portátil, un asistente de navegación personal, una cámara, o cualquier dispositivo móvil capaz de ejecutar un asistente personal digital en beneficio de un usuario. Un ejemplo de un dispositivo móvil que puede incorporar la funcionalidad de dispositivo de cómputo de usuario final 102 se discutirá más adelante con referencia a la Figura 14. En otra modalidad, el dispositivo de cómputo de usuario final 102 comprende

una computadora de escritorio, una consola de juegos, u otra plataforma de cómputo no móvil que es capaz de ejecutar un asistente personal digital en beneficio de un usuario. Una computadora de escritorio ilustrativa que puede incorporar la funcionalidad de dispositivo de cómputo de usuario final 102 se discutirá más adelante con referencia a la Figura 15.

El dispositivo de cómputo de usuario final 102 es capaz de comunicarse con el respaldo de asistente personal digital 106 a través de la red 104. El respaldo de asistente personal 106 comprende una o más computadoras (por ejemplo, servidores) que están programadas para proporcionar servicios en soporte de las operaciones del asistente personal digital 130 y otros asistentes personales digitales que se ejecutan en otro dispositivo de cómputo de usuario final. Por ejemplo, como se discutirá aquí, el respaldo de asistente personal digital 106 incluye una o más computadoras configuradas para proporcionar servicios para el asistente personal digital 130 que se refiere al reconocimiento por voz y entendimiento y respuesta de consulta. En particular, como se muestra en la Figura 1, estos servicios se proporcionan respectivamente por un servicio de reconocimiento por voz 132 y un sistema de entendimiento de consulta y respuesta 136. Se observará que el respaldo de asistente personal digital 106 puede realizar cualquier número de otros servicios en beneficio del asistente personal digital 130 aunque tales servicios adicionales no pueden describirse aquí de manera explícita.

En una modalidad, el respaldo de asistente personal digital 106 comprende un respaldo basado en nube en el cual cualquiera de un gran número de máquinas configuradas adecuadamente puede seleccionarse de manera arbitraria para prestar uno o más servicios
5 deseados en el soporte del asistente personal digital 130. Como se apreciará por aquellos expertos en la técnica(s) relevante, tal implementación basada en nube proporciona una estructura legible y escalable para proporcionar servicios de respaldo para los asistentes personales digitales, tal como el asistente personal digital 130.

10 La red(es) 104 pretende representar cualquier tipo de red o combinación de redes adecuadas para facilitar la comunicación entre los dispositivos de cómputo usuario final, tal como el dispositivo de cómputo de usuario final 102, y el respaldo de asistente personal digital 106. La red(es) 104 puede incluir, por ejemplo y sin
15 limitación, una red de área ancha, una red de área local, una red privada, una red pública, una red de paquete, una red conmutada de circuito, una red por cable, y/o una red inalámbrica.

Además, como se muestra en la Figura 1, el dispositivo de cómputo de usuario final 102 incluye una pluralidad de componentes
20 interconectados, que incluyen una unidad de procesamiento 110, memoria no volátil 120, memoria volátil 112, uno o más dispositivos de entrada de usuario 116, una presentación 118, y una o más interfaces de red 114. Cada uno de estos componentes se describirá ahora.

25 La unidad de procesamiento 110 pretende representar uno o

más microprocesadores, cada uno de los cuales puede tener una o más unidades de memoria de procesamiento central (CPU) o núcleos de microprocesador. La unidad de procesamiento 110 opera en una forma bien conocida para ejecutar programas de computadora (también se refieren aquí como lógica de programa de computadora). La ejecución de tales programas de computadora causa que la unidad de procesamiento 110 realice operaciones que incluyen operaciones que se describirán aquí. Cada una de la memoria no volátil 120, memoria volátil 112, dispositivo(s) de entrada 116, presentación 118, e interfaz(ces) de red 114 se conectan a la unidad de procesamiento 110 a través de una o más interfaces adecuadas.

La memoria no volátil 120 comprende uno o más dispositivos de memoria legibles por computadora que operan para almacenar programas de computadora y datos en una forma persistente, de manera que la información almacenada no se perderá incluso cuando el dispositivo de cómputo de usuario final 102 está sin energía o en un estado de apagado. La memoria no volátil 120 puede implementarse utilizando cualquiera de una variedad amplia de dispositivos de memoria legibles por computadora no volátiles, que incluyen pero no se limitan a, dispositivos de memoria de sólo lectura (ROM), unidades de estado sólido, unidades de disco duro, medios de almacenamiento magnético tales como discos magnéticos y unidades asociadas, medios de almacenamiento ópticos tales como discos ópticos y unidades asociadas, y dispositivos de memoria flash tales como unidades flash de USB.

La memoria volátil 112 comprende uno o más dispositivos de memoria legibles por computadora que operan para almacenar programas de computadora y datos en una forma no persistente, de manera que la información almacenada se perderá cuando el dispositivo de cómputo de usuario final 102 está sin energía o en un estado de apagado. La memoria volátil 112 puede implementarse utilizando cualquiera de una amplia variedad de dispositivos de memoria legibles por computadora volátiles que incluyen, pero no se limitan a, dispositivos de memoria de acceso aleatorio (RAM).

La presentación 118 comprende un dispositivo para que el contenido, tal como texto de imágenes, pueda reproducirse para que sea visible a un usuario de un dispositivo de cómputo de usuario final 102. Algunas o todas de las operaciones de reproducción requeridas para presentar tal contenido pueden realizarse al menos en parte por la unidad de procesamiento 110. Algunas o todas de las operaciones representadas también pueden realizarse por una interfaz de dispositivo de presentación tal como un chip o tarjeta de vídeo o gráficos (no mostrada en la Figura 1) que se acoplan entre la unidad de procesamiento 110 y la presentación 118. Dependiendo de la implementación del dispositivo de cómputo de usuario final 102, la presentación 118 puede comprender un dispositivo que está integrado dentro de la misma estructura física o alojamiento como unidad de procesamiento 110 o puede comprender un monitor, proyector, u otro tipo dispositivo que se separa físicamente de una estructura o un alojamiento que incluye la unidad de procesamiento

110 y se conecta al mismo a través de una conexión por cable y/o inalámbrica adecuada.

El dispositivo(s) de entrada usuario 116 comprende uno o más dispositivos que operan para generar la información de entrada de usuario en respuesta a una manipulación del usuario o control del mismo. Tal información de entrada de usuario se pasa a través de una interfaz adecuada para la unidad de procesamiento 110 para procesar el mismo. Dependiendo de la implementación, un dispositivo(s) de entrada de usuario 116 puede incluir una pantalla táctil (por ejemplo, una pantalla táctil integrada con presentación 118), un teclado, un teclado numérico, un ratón, una almohadilla táctil, una seguidora, una palanca de mandos, un dispositivo de puntero, un guante por cable, un sensor de rastreo por movimiento, un controlador o almohadilla de juego del juego, o un dispositivo de captura de video tal como una cámara. Sin embargo, estos ejemplos no pretenden limitarse y el dispositivo(s) de entrada de usuario 116 puede incluir otros tipos de dispositivos que los otros enlistados aquí. Dependiendo de la implementación, cada dispositivo de entrada de usuario 116 puede integrarse dentro de la misma estructura física o alojamiento como la unidad de procesamiento 110 (tal como una pantalla táctil integrada, almohadilla táctil, o teclado en un dispositivo móvil) o separarse físicamente de una estructura física o alojamiento que incluye la unidad de procesamiento 110 y conectada al mismo a través de una conexión por cable y/o inalámbrica adecuada.

La interfaz(ces) de red 114 comprende una o más interfaces que permiten al dispositivo de cómputo de usuario final 102 comunicarse a través de una o más redes 104. Por ejemplo, la interfaz(es) de red 114 puede comprender una interfaz de red por cable tal como una interfaz por Ethernet o una interfaz de red inalámbrica tal como una interfaz IEEE 802.11 (Wi-Fi) o una interfaz de telecomunicación 3G. Sin embargo, estos son únicamente ejemplos y no pretenden ser limitantes.

Como además se muestra en la Figura 1, la memoria no volátil 120 almacena un número de componentes de *software* que incluyen una pluralidad de aplicaciones 122 y un sistema operativo 124.

Cada aplicación en la pluralidad de aplicaciones 122 comprende un programa de computadora que un usuario de un dispositivo de cómputo de usuario final 102 puede hacer que se ejecute por la unidad de procesamiento 110. La ejecución de cada aplicación causa que ciertas aplicaciones se realicen en beneficio del usuario, en donde el tipo de operaciones realizadas variarán dependiendo de cómo la aplicación se programa. Las aplicaciones 122 pueden incluir, por ejemplo y sin limitación, una aplicación de telefonía, una aplicación de correo electrónico, una aplicación de mensajería, una aplicación de navegador web, una aplicación de calendario, una aplicación de utilidad, una aplicación de juego, una aplicación de red social, una aplicación de música, una aplicación de productividad, una aplicación de estilo de vida, una aplicación de referencia, una aplicación de viaje, una aplicación de deportes, una

aplicación de navegación, una aplicación de cuidado de la salud y de forma física, una aplicación de noticias, una aplicación de fotografía, una aplicación de finanzas, una aplicación de negocios, una aplicación de educación, una aplicación de clima, una aplicación de
5 libros, una aplicación médica, o similares. Como se muestra en la Figura 1, las aplicaciones 122 incluyen un asistente personal digital 130, las funciones de las cuales se describirán en mayor detalle aquí.

Las aplicaciones 122 pueden distribuirse para y/o instalarse en
10 el dispositivo de cómputo de usuario final 102 en una variedad de formas, dependiendo de la implementación. Por ejemplo, en una modalidad, al menos una aplicación se descarga desde una tienda de aplicación y se instala en el dispositivo de cómputo de usuario final 102. En otra modalidad en la cual el dispositivo de usuario final 102
15 se utiliza como parte de o en conjunto con una red de empresa, al menos una aplicación se distribuye para un dispositivo de cómputo de usuario final 102 por un administrador de sistema utilizando cualquiera de una variedad de herramientas de manejo de red de empresa y entonces se instalan en el mismo. Incluso en otra
20 modalidad, al menos una aplicación se instala en el dispositivo de cómputo de usuario final 102 por un constructor de sistema, tal como por un fabricante de equipo original (OEM) o fabricante de dispositivo incorporado, utilizando cualquiera de una variedad de utilidades de constructor de sistema adecuadas. En una modalidad
25 adicional, un fabricante de sistema operativo puede incluir una

aplicación a lo largo del sistema operativo 124 que se instala en el dispositivo de cómputo de usuario final 102.

El sistema operativo 124 comprende un grupo de programas que manejan recursos y proporcionan servicios comunes para las aplicaciones que se ejecutan en el dispositivo de cómputo de usuario final 102, tal como las aplicaciones 122. Entre otras características, el sistema operativo 124 comprende una interfaz de usuario de sistema operativo (OS) 132. La interfaz de usuario OS 132 comprende un componente del sistema operativo 124 que genera una interfaz de usuario por la cual un usuario puede interactuar con el sistema operativo 124 para varios propósitos, tal como pero no limitándose a encontrar e iniciar aplicaciones, invocar ciertas funcionalidades del sistema operativo, y configurar ciertas configuraciones del sistema operativo. En una modalidad, la interfaz de usuario OS 132 comprende una interfaz de usuario gráfica basada en pantalla táctil (GUI), aunque esto únicamente es un ejemplo. De acuerdo con tal ejemplo adicional, cada aplicación 122 instalada en el dispositivo de cómputo de usuario final 102 puede representarse como un icono o mosaico dentro del GUI e invocarse por un usuario a través de la interacción de pantalla táctil con el icono o mosaico apropiado. Sin embargo, cualquiera de una amplia variedad de modelos de interfaz de usuario alternativos puede utilizarse por la interfaz de usuario OS 132.

Aunque las aplicaciones 122 y el sistema operativo 124 se muestran como estando almacenados en la memoria no volátil 120,

se entenderá que durante operación del dispositivo de cómputo de usuario final 102, aplicaciones 122, sistema operativo 124, o porciones del mismo pueden cargarse a la memoria volátil 112 y ejecutarse de la misma a medida que se procesa por la unidad de procesamiento 110.

El asistente personal digital 130 comprende un programa de computadora que está configurado para realizar tareas, o servicios para un usuario del dispositivo de cómputo de usuario final 102 basándose en la entrada de usuario además de las características tal como el conocimiento de ubicación y la habilidad para acceder a la información de una variedad de fuentes que incluyen fuentes en línea (tal como condiciones de clima o tráfico, noticias, precios en existencia, programas de usuario, precios al por menor, etc.). Ejemplos de tareas que pueden realizarse por el asistente personal digital 130 en favor del usuario pueden incluir, pero no se limitan a, colocar una llamada telefónica a una persona especificada de usuario, iniciar una aplicación especificada por usuario, enviar un correo electrónico especificado por usuario o mensaje de texto a un receptor especificado por usuario, reproducir música especificada por usuario, programar una reunión, u otro evento en un calendario de usuario, obtener direcciones a una ubicación especificada por usuario, obtener una clasificación asociada con un evento deportivo especificado por usuario, publicar contenido especificado por usuario a un sitio web de medios sociales o servicio de microblogueo, registrar recordatorios o notas específicas por usuario, obtener un

reporte del clima, obtener el tiempo actual, configurar una alarma en un tiempo especificado por usuario, obtener un precio en existencia para una compañía especificada por usuario, encontrar un establecimiento comercial cercano, realizar una búsqueda de Internet, o similares. El asistente personal digital 130 puede utilizar cualquiera de una variedad de técnicas de inteligencia artificial para mejorar su desempeño al momento a través de la interacción continua con el usuario. El asistente personal digital 130 también puede referirse a un asistente personal inteligente, un asistente de *software* inteligente, un asistente personal virtual, o similares.

El asistente personal digital 130 está configurado para proporcionar una interfaz de usuario por la cual un usuario puede presentar preguntas, comandos, u otra entrada verbal y por la cual responde a tal entrada que puede entregarse al usuario. En una modalidad, la entrada puede comprender voz del usuario que se captura por uno o más micrófonos del dispositivo de cómputo de usuario final 102 (cada uno de los cuales puede comprender uno de los dispositivos de entrada de usuario 116), aunque este ejemplo no pretende ser limitante y la entrada de usuario puede proporcionarse en otras formas también. Las respuestas generadas por el asistente personal digital 130 pueden hacerse visibles al usuario en la forma de texto, imágenes, u otro contenido visual mostrado en la presentación 118 dentro de una interfaz de usuario gráfica del asistente personal digital 130. Las respuestas también pueden comprender voz generada por computadora u otro contenido de audio

que se reproduce después o a través de una o más bocinas del dispositivo de cómputo de usuario final 102 (no mostrado en la Figura 1).

De acuerdo con las modalidades, el asistente personal digital 130 es capaz de determinar que un usuario ha hecho una pregunta o ha hecho una declaración que pretende involucrarse con un personaje del asistente personal digital 130 en oposición a, por ejemplo, solicitar que el asistente personal digital obtenga información o realice otra tarea en beneficio del usuario. Tales preguntas o declaraciones frecuentemente son casuales o alegres por naturaleza y pueden incluir, por ejemplo, ¿te casarías conmigo?, ¿Cuál es tu color favorito?, ¿Cántame una canción?, ¿Cuéntame un chiste?, ¿Knock knock?, ¿Cuánta madera podría una marmota arrojar si una marmota pudiera arrojar madera?, ¿Quién hace el mejor teléfono?, ¿En dónde puedo esconder un cuerpo?, ¿Qué aspecto tienes?, ¿Tú eres hermosa?, ¿Cuántos años tienes?, ¿Quién es tu papá?, ¿Tienes novio?, ¿Cuál es el significado de la vida?, ¿Me gustaría conocerte mejor?, o similares. Para el bien de conveniencia, tales preguntas y declaraciones se refieren aquí como pronunciaciones de tipo charla o simplemente charla.

Como se discutirá a continuación, el asistente personal digital 130 además está configurado para tomar ciertas acciones en respuesta a determinar que el usuario ha hecho una pronunciación de tipo charla. Por ejemplo, en respuesta a determinar que el usuario ha hecho una pronunciación de tipo charla, el asistente personal

digital 130 puede proporcionar una respuesta al mismo al presentar un objeto multimedia asociado con una referencia de cultura popular dentro de su interfaz de usuario (cuando el objeto multimedia es visual en naturaleza) o reproducir después por su interfaz de usuario tal objeto multimedia (cuando el objeto multimedia es auditado en naturaleza). Como se utiliza aquí, el término "referencia de cultura popular" pretende abarcar ampliamente una referencia a cualquier tema asociado con las personalizaciones, técnicas y/o interacciones sociales de una porción grande de una población. Por ejemplo, y sin cualquier limitación, una referencia de cultura popular puede incluir una referencia para una película bien conocida, programa de televisión, novela, historia, pintura, juego de video, imagen, video, caricatura, celebridad, actor o actriz, político u otra figura pública, estereotipo, meme, evento actual, evento histórico, o similares.

Adicional o alternativamente, en respuesta a determinar que el usuario ha hecho una pronunciación de tipo charla, el asistente personal digital 130 puede proporcionar la respuesta del mismo al generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con la referencia de cultura popular. Incluso además, el asistente personal digital 130 puede configurarse para proporcionar la respuesta al presentarse dentro de su texto de interfaz de usuario que comprende una cita asociada con la referencia de cultura popular, que presenta dentro de su interfaz de usuario una representación visual del asistente personal digital e invoca la referencia de cultura popular,

y/o que presenta dentro de su interfaz de usuario un enlace que puede activarse por el usuario para acceder al contenido asociado con la referencia de cultura popular.

Una forma en la cual el asistente personal digital 130
5 determina que el usuario ha hecho una pronunciación de tipo charla proporciona una respuesta del mismo que ahora se describirá con referencia al diagrama de bloques 200 de la Figura 2. En particular, diagrama de bloques 200 muestra cómo varios componentes del sistema 100 operan juntos para permitir al asistente personal digital
10 130 a determinar que el usuario ha hecho una pronunciación de tipo charla y para proporcionar una respuesta al mismo.

Con referencia a la Figura 2, el proceso comienza después de que el asistente personal digital 130 ha iniciado en un dispositivo de cómputo de usuario final 102. Para interactuar con el asistente
15 personal digital 130, un usuario habla en uno o más micrófonos del dispositivo de cómputo de usuario final 102. La pronunciación del usuario se captura por el micrófono(s) y se convierte de forma análoga a digital en una forma bien conocida. El asistente personal digital 130 causa que la representación digital de la pronunciación se
20 transmita como una transmisión de audio para el servicio de reconocimiento por voz 132 (el cual es parte del respaldo de asistente personal digital 106) a través de la red(es) 104. Por ejemplo, en una modalidad, un usuario habla en el micrófono(s) del dispositivo de cómputo de usuario final, así el asistente personal
25 digital 130 periódicamente causa que una representación digital de

una porción de la pronunciación del usuario sea empacada y transmitida a un servicio de reconocimiento por voz 132 a través de la red(es) 104.

El servicio de reconocimiento por voz 132 opera para recibir la
5 transmisión de audio transmitido al mismo por el asistente personal digital 130 y para analizar la transmisión de audio para determinar el contenido fonético del mismo. Una vez que el servicio de reconocimiento por voz 132 ha determinado el contenido fonético de la transmisión de audio, entonces se asigna el contenido fonético a
10 una o más palabras, que tomadas juntas comprenden una pronunciación reconocida. El servicio de reconocimiento por voz 132 entonces pasa la pronunciación reconocida al sistema de entendimiento de consulta y respuesta 136.

Como se muestra en la Figura 2, el servicio de reconocimiento
15 por voz 132 también puede transmitir la pronunciación reconocida de regreso al asistente personal digital 130 a través de la red(es) 104. El asistente personal digital 130 puede presentar una versión de texto de la pronunciación reconocida dentro de su interfaz de usuario gráfica (visible a través de la presentación 118) para que el usuario
20 pueda ver la pronunciación reconocida y determinar si o no la pronunciación reconocida representa adecuadamente que él/ellas dicen. De acuerdo con ciertas modalidades, el asistente personal digital 130 además puede proporcionar medios por los cuales el usuario puede editar la pronunciación reconocida si él/ella determina
25 que la pronunciación reconocida no representa precisamente que

él/ella dice y transmite la versión editada de la pronunciación para el sistema de entendimiento de consulta y respuesta 136 para procesamiento adicional del mismo.

El sistema de entendimiento de consulta y respuesta 136 recibe
5 la pronunciación reconocida o corregida y analizan las palabras de la misma para determinar cómo tal pronunciación podría manejarse. Por ejemplo, el sistema de entendimiento de consulta y respuesta 136 puede determinar que la pronunciación reconocida o corregida comprende una invocación de una tarea particular dentro de un grupo
10 predefinido de tareas. Por ejemplo y sin cualquier limitación, la tarea puede comprender colocar una llamada telefónica a una persona especificada por usuario (por ejemplo, ñllamar a Bryanò), enviar un correo electrónico especificado por usuario o mensaje de texto a un receptor especificado por usuario (por ejemplo, ñtexto de Carol que
15 estoy corriendo tardeò), o crear un recordatorio (por ejemplo, ñrecordarme de verificar la estufa en una horaò). Si el sistema de entendimiento de consulta y respuesta 136 determina que la pronunciación reconocida o corregida comprende una invocación de una tarea particular dentro del grupo predefinido, entonces causará
20 lógica especializada (por ejemplo, lógica especializada dentro del dispositivo de cómputo de usuario final 102) para realizar la tarea.

Sin embargo, si el sistema de entendimiento de consulta y respuesta 136 falla al coincidir con la pronunciación reconocida o corregida para cualquiera de las tareas dentro del grupo predefinido,
25 entonces el sistema de entendimiento de consulta y respuesta 136

además puede analizar las palabras de la pronunciación al determinar cómo tal pronunciación podría manejarse por el mismo. Por ejemplo, el sistema de entendimiento de consulta y respuesta 136 pueden determinar que la pronunciación podría manejarse al conducir una búsqueda web o al ofrecer al usuario una oportunidad de conducir una búsqueda web. En este caso, la pronunciación puede manejarse por lógica especializada para facilitar la búsqueda web que es interna y/o externa al sistema de entendimiento de consulta y respuesta 136. Alternativamente, el sistema de entendimiento de consulta y respuesta 136 puede determinarse basándose en un análisis de las palabras de la pronunciación que la pronunciación comprende una pronunciación de tipo charla, que como se observó anteriormente es una pronunciación que pretende comprometerse con un personaje del asistente personal digital 130.

Dependiendo de las implementaciones, el sistema de entendimiento de consulta y respuesta 136 pueden determinar que la pronunciación comprende una pronunciación de tipo charla basándose en un análisis de factores distinto o en adición a un análisis de palabras de la pronunciación. Por ejemplo, el sistema de entendimiento de consulta y respuesta 136 puede determinar que la pronunciación comprende una pronunciación de tipo charla basándose en parte en un análisis de una entonación de la pronunciación, en pistas contextuales obtenidas de una historia de conversación del usuario, o en cualquiera de otros factores que pueden considerarse útiles al determinar que la pronunciación

comprende una pronunciación de tipo charla.

Si el sistema de entendimiento de consulta y respuesta 136 determinan que la pronunciación comprende una pronunciación de tipo charla, entonces la pronunciación se manejará por un sistema de entendimiento de consulta y respuesta para charla 138, el cual es una parte del sistema de entendimiento de consulta y respuesta 136. El sistema de entendimiento de consulta y respuesta para charla 138 se configura para determinar el tema de la pronunciación de tipo charla y entonces, basándose en el tema determinado, tomar pasos para causar una respuesta apropiada para la pronunciación de tipo charla para acceder a la salida por el asistente personal digital 130. Como se muestra en la Figura 2, esto puede involucrar enviar todo o parte de una respuesta del sistema de entendimiento de consulta y respuesta para charla 138 al asistente personal digital 130 a través de la red(es) 104. La composición de la respuesta y la forma en la cual se transporta a y/o genera por el asistente personal digital 130 se discutirán en mayor detalle a continuación.

En una modalidad, el sistema de entendimiento de consulta y respuesta para charla 138 determina el tema de la pronunciación de tipo charla y entonces identifica una pluralidad de respuestas elegibles que son adecuadas para responder a la pronunciación. El sistema de entendimiento de consulta y respuesta para charla 138 entonces selecciona uno de la pluralidad de respuestas elegibles como la respuesta al proporcionarse por el asistente personal digital 130. Tal selección puede realizarse al azar, en cierta secuencia, o al

utilizar algunas otras metodologías de selección. Al mantener una pluralidad de respuestas elegibles para un tipo de pronunciación de charla dado, el sistema de entendimiento de consulta y respuesta para charla 138 puede asegurar que el asistente personal digital 130
 5 no proporcionará la misma respuesta a la misma pronunciación en cada instancia, por lo tanto, proporcionará alguna variedad de imprevisibilidad para la interacción del usuario con el asistente personal digital 130.

En una modalidad, el sistema de entendimiento de consulta y
 10 respuesta para charla 138 opera para coincidir con la pronunciación de tipo charla para un tipo de pronunciación particular dentro de un árbol jerárquico de tipos de pronunciación que tienen una o más respuestas asociadas del mismo. El sistema de entendimiento de consulta y respuesta para charla 138 entonces selecciona la
 15 respuesta para la pronunciación de tipo charla de entre la respuesta(s) asociada del mismo.

Para ayudar a ilustrar este concepto, la Figura 3 describe un árbol jerárquico ilustrativo 300 de los tipos de pronunciación que pueden utilizarse para seleccionar una respuesta para una
 20 pronunciación de tipo charla de acuerdo con una modalidad. Como se muestra en la Figura 3, el nodo de raíz del árbol jerárquico 300 es el tipo de pronunciación de charla general. Cada tipo de pronunciación por debajo de este nodo de raíz comprende una pronunciación de tipo charla. En un nivel por debajo de este nodo de raíz son
 25 pronunciaciones de tipo charla que son afirmaciones (ñafirmaciónò),

comandos (ñComandoò), coqueteos en naturaleza (ñCoqueteoò), que solicita información sobre el asistente personal digital 130 (ñInformación de sistemaò), o que solicita una opinión del asistente personal digital 130 (ñOpinión del sistemaò). Por debajo de cada uno

5 de estos nodos están categorías adicionales y subcategorías de tipos de pronunciación de charla. De esa forma, los tipos de pronunciación generalmente van desde lo más amplio en el tipo del árbol jerárquico 300 más estrecho en la parte inferior del árbol jerárquico 300.

En una modalidad, basándose en un análisis de la

10 pronunciación de tipo charla, el sistema de entendimiento de consulta y respuesta para charla 138 atraviesa el árbol jerárquico 300 y coincide la pronunciación para uno de los nodos. Por ejemplo, el sistema de entendimiento de consulta y respuesta para charla 138 puede generar una clasificación de confianza que una cierta

15 pronunciación de tipo charla podría coincidir para ñafirmarò, ñcomandoò, ñcoqueteosò, ñinformación de sistemaò y ñopinión de sistemaò. El sistema de entendimiento de consulta y respuesta para charla 138 entonces selecciona el modo para el cual la clasificación de coincidencia más alta sea obtenida (asumir que alguna

20 clasificación de confianza mínima sea obtenida por al menos un valor de los nodos). Asumir para el bien de ilustración que la clasificación de confianza más alta se genera para la ñopinión de sistemaò. Entonces, el sistema de entendimiento de consulta y respuesta para charla 138 atravesará el árbol jerárquico 300 hacia el

25 nodo ñopinión de sistemaò y generará una clasificación de

confidencia que la pronunciación de tipo charla podría coincidirse para cada uno de los nodos secundarios de "opinión de sistema" principalmente, "Microsoft", "tendencia" y "aviso". El sistema de entendimiento de consulta y respuesta para charla 138 entonces
 5 selecciona el nodo secundario para el cual la clasificación de confidencia más alta se ha logrado (de nuevo, asumiendo alguna clasificación de confidencia mínima que se ha obtenido por al menos uno de los nodos secundarios). Si la clasificación de confidencia para cada uno de los nodos secundarios es menor que alguna
 10 clasificación de confidencia mínima predefinida, entonces el recorrido del árbol jerárquico 300 se detiene en el nodo "opinión de sistema".

Una o más respuestas están asociadas con cada nodo en el árbol jerárquico 300. Cuando el sistema de entendimiento de
 15 consulta y respuesta para charla 138 se hace al recorrer el árbol jerárquico 300 y ha seleccionado un nodo de coincidencia para una pronunciación de tipo charla, entonces el sistema de entendimiento de consulta y respuesta para charla 138 pueden seleccionar una respuesta de entre la una o más respuestas asociadas con el nodo
 20 de coincidencia.

El aspecto anterior para identificar respuestas adecuadas para las pronunciaciones de tipo charla es ventajoso ya que permite las respuestas al definirse para ambos grupos amplios de las pronunciaciones de tipo charla además de más grupos estrechos
 25 dentro de los grupos más amplios. A manera de ejemplo, para el

nodo "Microsoft" dentro del nodo "opinión de sistema", respuestas muy específicas para las pronunciaciones de tipo charla pueden manipularse (por ejemplo, "¡yo creo que Microsoft es grandioso!"), ya que el sistema tiene un nivel alto de confianza que el usuario está preguntando por la opinión del asistente personal digital 130 sobre Microsoft. En contraste, para el nodo de "opinión de sistema", una respuesta más genérica para las pronunciaciones de tipo charla pueden manipularse (por ejemplo, "sin comentarios" o "me gustaría no decirlo"), desde que el sistema tiene un nivel alto de confianza que el usuario está preguntando por la opinión del asistente personal digital 130, pero no puede determinar el tema sobre el cual una opinión está siendo buscada.

En una modalidad, los tipos de pronunciaciones que pueden incluirse en el árbol jerárquico 300 pueden determinarse a través de examen humano de registros de las pronunciaciones de usuario y etiquetando cada pronunciación con un tipo de pronunciación apropiada. Por ejemplo, un grupo de plataformas de fuentes tal como el sistema de relevancia humana universal (UHRS), desarrollado por Microsoft Corporation de Redmond, Washington, pueden utilizarse para obtener examen humano y etiquetas de miles de pronunciaciones de usuarios. Esta información de fuente múltiple entonces puede utilizarse para generar el árbol jerárquico 300. Aún otros métodos para generar un árbol jerárquico de tipos de pronunciación tal como el árbol jerárquico 300 pueden utilizarse.

En otra modalidad, el sistema de entendimiento de consulta y

respuesta para charla 138 está configurado para mantener una o más
 respuestas asociadas con cada uno de una o más tendencias tópicas.
 Generalmente hablando, las tendencias tópicas son tópicos que se
 vuelven populares o se han vuelto recientemente populares con
 5 usuarios y pueden identificarse automáticamente (por ejemplo,
 pronunciaciones automáticamente verificadas presentadas a los
 asistentes personales digitales, consultas de procesador de
 búsqueda, microblogs tal como TWITTER, sitios de red social tal
 como FACEBOOK, publicaciones de noticias, u otras fuentes) o
 10 manualmente (por ejemplo, a través de la observación humana de
 cualquiera o todas de estas fuentes). Cuando el sistema de
 entendimiento de consulta y respuesta para charla 138 determina que
 la pronunciación de tipo charla está asociada con una tópico de
 tendencia particular, entonces el sistema de entendimiento de
 15 consulta y respuesta para charla 138 puede seleccionar la respuesta
 a la pronunciación de tipo charla de entre la una o más respuestas
 asociadas con la tópico de tendencia particular.

De acuerdo con tal modalidad adicional, las tendencias tópicas
 pueden representarse dentro de un árbol jerárquico de tipos de
 20 pronunciación que se utilizan por el sistema de entendimiento de
 consulta y respuesta para charla 138 al seleccionar una respuesta
 adecuada para la pronunciación de tipo charla. Por ejemplo, como se
 muestra en la Figura 3, uno de los nodos bajo la opinión de sistema
 es la tendencia. Este nodo puede utilizarse para almacenar
 25 respuestas a las pronunciaciones de tipo charla que solicitan una

opinión del asistente personal digital 130 con respecto a una o más tendencias tópicas. Incluso con esta modalidad adicional, el nodo de tendencia puede tener nodos secundarios múltiples asociados con el mismo, en donde cada nodo secundario está asociado con un tópico de tendencia particular y tiene una o más respuestas asociadas con la misma.

Incluso en otra modalidad, el sistema de entendimiento de consulta y respuesta para charla 138 está configurado para mantener una o más respuestas para ciertas pronunciaciones de tipo charla que pretenden transportar al personaje del asistente personal digital 130. Por ejemplo, puede ser de interés en asegurar que el asistente personal digital 130 tenga algo que decir sobre una palabra particular, frase, o tópico que está asociado con su personaje. En este caso, un equipo editorial puede generar respuestas predefinidas a ciertas pronunciaciones de tipo charla para asegurar que el asistente personal digital 130 proporcione respuestas características cada vez que se discute tal tópico. En este caso, cuando el sistema de entendimiento de consulta y respuesta para charla 138 determina que la pronunciación de tipo charla es una pronunciación para el cual existen una o más respuestas predefinidas que pretenden transportar al personaje del asistente personal digital 130, el sistema de entendimiento de consulta y respuesta para charla 138 seleccionará la respuesta para la pronunciación de tipo charla de entre la una o más respuestas predefinidas.

La Figura 4 es un diagrama de bloques que muestra una

respuesta ilustrativa 400 que puede proporcionarse por el asistente personal digital 130 en respuesta a la pronunciación de tipo charla de acuerdo con una modalidad. Como se muestra en la Figura 4, la respuesta 400 incluye un número de componentes, que incluyen una
5 cadena de presentación 402, contenido de voz 404, componente de personificación por voz 406, un componente de emoción por voz 408, una animación de asistente personal digital 410, un componente multimedia 412, y un enlace a contenido 414.

Dependiendo de la implementación, cada uno de los
10 componentes dentro de la respuesta 400 puede almacenarse y/o generarse por el respaldo de asistente personal digital 106 y transmitirse al asistente personal digital 130 por el sistema de entendimiento de consulta y respuesta para charla 138 en el momento de la respuesta que se proporciona a un usuario.
15 Alternativamente, uno, más de uno, o todos de los componentes de respuesta 400 pueden almacenarse en y/o generarse por el dispositivo de cómputo de usuario final 102 (por ejemplo, en la memoria no volátil 120) y el sistema de entendimiento de consulta y respuesta para charla 138 puede enviar al asistente personal digital
20 130 suficiente información para identificar u obtener el componente(s) en el momento de la respuesta que será proporcionada a un usuario, para que el asistente personal digital 130 pueda obtener el componente(s) localmente.

Adicionalmente, aunque la respuesta 400 incluye siete
25 componentes diferentes, se entenderá que una respuesta para una

pronunciación de tipo charla puede incluir menos que todos los componentes mostrados en la Figura 4.

La cadena de presentación 402 comprende texto que se presenta dentro de la interfaz de usuario del asistente personal digital 130. El texto puede comprender una respuesta verbal para la pronunciación de tipo charla del usuario. En una modalidad, la cadena de presentación 402 puede comprender una cita que está asociada con una referencia de cultura popular.

El contenido por voz 404 comprende voz que se genera o reproduce después por la interfaz de usuario o asistente personal digital 130. El asistente personal digital 130 puede generar tal voz al aplicar conversión de texto a voz para el texto que comprende parte del contenido por voz 404. Alternativamente, el asistente personal digital 130 puede generar tal voz al reproducir después un archivo de audio que se incluye dentro o se identifica por el contenido por voz 404. En una modalidad, el contenido por voz 404 comprende una versión audible del contenido incluido en la cadena de presentación 402, aunque esto no necesariamente es el caso. Por ejemplo, el contenido por voz 404 puede comprender información verbal que es completamente diferente a la información verbal incluida en la cadena de presentación 402. Similar a la cadena de presentación 402, el contenido de contenido por voz 404 puede comprender una cita que se asocia con una referencia de cultura popular.

El componente de personificación por voz 406 es un componente que indica que el asistente personal digital 130 debe

generar o reproducir después contenido por voz 404 en una forma que personifique una voz de una persona, tal como una persona asociada con una referencia de cultura popular. El componente de personificación por voz 406 puede incluir o identificar un archivo de audio que debe reproducirse después por el asistente personal digital 130 para realizar la personificación. Alternativamente, el componente de personificación por voz 406 puede indicar que un convertidor de texto a voz debe utilizarse por el asistente personal digital 130 para generar contenido por voz 404 en una forma que personifique la voz de la persona deseada.

El componente de emoción por voz 408 comprende un elemento emocional que debe aplicarse al contenido por voz 404 cuando la conversión de texto a voz se aplica a tal contenido para generar voz.

La animación de asistente personal digital 410 comprende una animación de un avatar que representa el asistente personal digital 130 que está para presentarse dentro de su interfaz de usuario. La animación puede designarse de manera que invoque una referencia de cultura popular. Aunque la descripción de respuesta 400 se refiere a la animación de asistente personal digital 410, se debe apreciar que tipos de representaciones visuales del avatar de otras animaciones pueden utilizarse para invocar la referencia de cultura popular, incluyendo imágenes estáticas o similares.

El componente multimedia 412 comprende uno o más objetos multimedia que se presentan dentro o reproducen después por la interfaz de usuario del asistente personal digital 130. Cada objeto

multimedia puede asociarse con una referencia de cultura popular. Dependiendo de la implementación, cada objeto multimedia puede comprender, por ejemplo, una imagen al presentarse dentro de la interfaz de usuario del asistente personal digital 130, contenido de vídeo que se presentará dentro de la interfaz de usuario del asistente personal digital 130, o contenido de audio que se reproducirá después por la interfaz de usuario del asistente personal digital 130.

El enlace a contenido 414 comprende un enlace que puede presentarse dentro de la interfaz de usuario del asistente personal digital 130 y que puede activarse por el usuario para acceder a otro contenido. Por ejemplo, el enlace puede activarse por el usuario al acceder al contenido asociado con una referencia de cultura popular.

Las Figuras 5-10 proporcionan varios ejemplos de respuestas para pronunciaciones de tipo charla que pueden entregarse a través de la interfaz de usuario del asistente personal digital 130. Estos ejemplos ayudan a ilustrar a los varios componentes que pueden incluirse en una respuesta para la pronunciación de tipo charla de acuerdo con las modalidades. En los ejemplos de las Figuras 5-10, el dispositivo de cómputo de usuario final 102 es un teléfono inteligente y la presentación 118 es una presentación integrada del teléfono inteligente. Sin embargo, como se observó previamente, el dispositivo de cómputo de usuario final 102 no se limita a teléfonos inteligentes y pueden ser cualquiera de una amplia variedad de dispositivos de cómputo móviles y no movibles.

La Figura 5 ilustra una respuesta que puede proporcionarse por el asistente personal digital 130 a la pronunciación de tipo charla ñyo estoy nervioso sobre la prueba de estadoò. En respuesta a esta pronunciación de tipo charla, una cadena de presentación 502 que
 5 comprende las palabras ñpuede la fuerza estar contigoò se presenta dentro de la interfaz de usuario gráfica del asistente personal digital 130. Este texto de cadena de presentación comprende una cita bien conocida de las películas populares ñGuerra de las Galaxiasò.

Además, en respuesta a la pronunciación de tipo charla, una
 10 representación visual del asistente personal digital 130 se presenta dentro de la interfaz de usuario gráfica. La representación visual 504 incluye un sable de luz, y de esa forma invoca a las películas ñGuerra de las Galaxiasò. En una modalidad, la representación visual 504 puede comprender parte de una animación del avatar del
 15 asistente personal digital 130 que oscila el sable de luz sobre, quizás acompañado por los sonidos de sables de luz audibles que se reproducen después a través de una o más bocinas del dispositivo de cómputo de usuario final 102.

En una modalidad, la respuesta proporcionada en la Figura 5
 20 también incluye voz audible que se reproduce después a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras ñque la fuerza esté contigoò. En una modalidad, la voz comprende una personificación de un personaje famoso de ñGuerra de las Galaxiasò tal como Yoda u
 25 Obi-Wan Kenobi. Como se describió previamente, tal voz puede

presentarse al reproducir después un archivo de audio designado o al aplicar un proceso de conversión de texto a voz especial para el texto "Que la fuerza esté contigo". Dependiendo de la implementación, la personificación puede presentarse de manera que se apreciará por una persona que el personaje de "Guerra de las Galaxias" realiza la personificación (por ejemplo, una voz predeterminada asociada con el asistente personal digital 130 se realiza la personificación). En otras implementaciones, la personificación puede producir una voz que es indistinguible del personaje de la "Guerra de las Galaxias", o puede de hecho ser la voz del actor que representó el personaje de la "Guerra de las Galaxias".

Tomados juntos, la cadena de presentación 502, representación visual 504, y la voz entregada con una personificación no únicamente de manera apropiada la respuesta a la pronunciación de tipo charla del usuario al ofrecer las palabras de ánimo, pero también sirve para invocar fuertemente una referencia de cultura popular ("Guerra de las Galaxias") que probablemente serán familiares de manera instantánea para el usuario y ayudarán a establecer un sentido de uniformidad del mismo. Adicionalmente, ya que la respuesta incluye diversas formas de salida de detección que incluyen la animación de sable de luz y la voz personificada, la respuesta es más probable que comprometa y entretenga al usuario que una respuesta de texto plana.

La Figura 6 ilustra una respuesta que puede proporcionarse por

el asistente personal digital 130 para la pronunciación de tipo charla ñ¿Cómo puedo robar un banco?ò En respuesta a esta pronunciación de tipo charla, una cadena de presentación 602 que comprende las palabras ñyo no termine bien con estos chicosò se presenta dentro de

5 la interfaz de usuario gráfica del asistente personal digital 130. La respuesta también puede incluir voz audible que se reproduce después a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras ñyo no termine bien con estos chicosò. El texto de la cadena de

10 presentación 602 se refiere a un objeto multimedia 604 que también se presenta dentro de la interfaz de usuario gráfica del asistente personal digital 130. Como se muestra en la Figura 6, un objeto multimedia 604 comprende una imagen del póster de la película para la película de 1967 ñBonnie & Clydeò, la cual es un drama que

15 concierne a la vida y a la muerte de los ladrones de banco Bonnie Parker y Clyde Barrow bien conocidos.

Tomados juntos, la cadena de presentación 602 (y la voz audible correspondiente) y el objeto multimedia 604 comprenden una respuesta para la pronunciación de charla ñcómo robaré un bancoò ya

20 que ambas respuestas responden apropiadamente a la pronunciación de tipo charla del usuario al señalarlos fuera de los peligros de robar un banco (Bonnie y Clyde dispararon a muerte a los oficiales de policía) y también invoca una referencia de cultura popular (Bonnie y Clyde y la película del mismo nombre) que probablemente es familiar

25 para el usuario y ayuda a establecer un sentido de uniformidad del

mismo. Adicionalmente, ya que la respuesta forma la salida distinta a texto plano, es más probable que comprometa y entretenga al usuario.

La Figura 7 ilustra una respuesta que puede proporcionarse por el asistente personal digital 130 a la pronunciación de tipo charla
5 ¿Cuál es tu coche favorito en el mundo entero? En respuesta a esta pronunciación de tipo charla, una cadena de presentación 702 que comprende las palabras ñyo amo al Deloreans. Especialmente uno que viaja a través del tiempo se presenta dentro de la interfaz
10 de usuario gráfica del asistente personal digital 130. La respuesta también puede incluir voz audible que se reproduce después a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras ñyo amo al Deloreans. Especialmente uno que viaja a través del tiempo. Este
15 texto y voz se refieren a las películas bien conocidas ñVolver al futuro.

Como también se muestra en la Figura 7, la respuesta también incluye un objeto multimedia 704 en la forma de un vídeo de YOUTUBE® llamado ñVolver al futuro - la escena de la torre del
20 reloj.avò que puede reproducirse y verse por el usuario dentro del contexto de la interfaz de usuario gráfica del asistente personal digital 130.

Como además se muestra en la Figura 7, la respuesta incluye un enlace 706 que, cuando se activa por el usuario, permite al
25 usuario buscar en la web la frase ñ¿Cuál es tu coche favorito en el

mundo entero?ò, que es la pronunciación original. Se observará que, en las modalidades alternativas, un enlace puede proporcionarse ya que, cuando se activa por el usuario, permite al usuario buscar en la web contenido asociado con la referencia de cultura popular (por ejemplo, las películas ñVolver al futuroò).

Similar a las respuestas de las Figuras 5 y 6, la respuesta de la Figura 7 invoca fuertemente una referencia cultura popular y de esa forma puede establecer uniformidad con el usuario. Adicionalmente, el contenido de vídeo que es directamente visible desde la interfaz de usuario gráfica del asistente personal digital 130 hace que la respuesta se comprometa altamente.

La Figura 8 ilustra una respuesta que puede proporcionarse por el asistente personal digital 130 a la pronunciación de tipo charla ñYou bitchò. En respuesta a esta pronunciación de tipo charla, una cadena de presentación 802 que comprende las palabras ñtambién soy una amante, una niña y una madreò se presenta dentro de la interfaz de usuario gráfica del asistente personal digital 130. La respuesta también puede incluir voz audible que se reproduce después a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras ñtambién soy una amante, una niña y una madreò. Este texto y voz comprenden una porción de las letras de la canción bien conocida ñBitchò por Meredith Brooks.

También como se muestra en la Figura 8, la respuesta también incluye un objeto multimedia 804 en la forma de una imagen de

Meredith Brooks. La respuesta también incluye una porción de texto 806 que proporciona información sobre la canción "Bitch" y también puede incluir uno o más enlaces que pueden activarse por el usuario al comprar una copia digital de la canción desde una o más fuentes, 5 respectivamente.

La respuesta a la Figura 8 ingeniosamente utiliza una referencia de cultura pop para responder a (y desviar algo) de la pronunciación de tipo charla aparentemente despectiva. Esto también incluye contenido multimedia de interés que pueda ayudar a 10 comprometer al usuario. La respuesta mostrada en la Figura 8 también ilustra como una respuesta a la pronunciación de tipo charla puede servir en algunos aspectos como un anuncio u oportunidad comercial en que el usuario ha permitido la compra de la canción que se refiere a la respuesta.

15 La Figura 9 ilustra una respuesta que puede proporcionarse por el asistente personal digital 130 a la pronunciación de tipo charla "¿Quién es tu papi?" En respuesta a esta pronunciación de tipo charla, una cadena presentación 902 que comprende las palabras "estos chicos" se presenta dentro de la interfaz de usuario gráfica 20 del asistente personal digital 130. La respuesta también puede incluir voz audible que se reproduce después a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras "estos chicos". El texto de la cadena de presentación 902 se refiere a un objeto multimedia 904 25 que también se presenta dentro de la interfaz de usuario gráfica del

asistente personal digital 130. Como se muestra en la Figura 9, el objeto multimedia 904 comprende un vídeo de Bill Gates (presidente de Microsoft) y Steve Bellmer (Director ejecutivo de Microsoft) vestidos como los personajes de Austin Powers y Dr. Evil, respectivamente, de las películas muy populares "Austin Powers". Este video puede reproducirse y verse por el usuario dentro del contexto de la interfaz de usuario gráfica del asistente personal digital 130.

En una modalidad en la cual un asistente personal digital 130 se publica por Microsoft Corporation de Redmond, Washington, la respuesta mostrada en la Figura 9 es apta ya que las Figuras mostradas en el vídeo son personajes bien conocidos asociados con Microsoft. Adicionalmente, el contenido de vídeo se asume y se involucra para el usuario.

La Figura 10 ilustra una respuesta que puede proporcionarse por el asistente personal digital 130 a la pronunciación de tipo charla "¿Muéstrame algo gracioso?". En respuesta a esta pronunciación de tipo charla, una cadena de presentación 1002 que comprende las palabras "un amigo mío tiene una opinión" se presenta dentro de la interfaz de usuario gráfica del asistente personal digital 130. La respuesta también puede incluir voz audible que se reproduce después o a través de una o más bocinas del dispositivo de cómputo de usuario final 102, en donde tal voz también incluye las palabras "Un amigo mío tiene una opinión". El texto de cadena de presentación 1002 se refiere a un objeto multimedia 1004 que

también se presenta dentro de la interfaz de usuario gráfica del asistente personal digital 130. Como se muestra en la Figura 10, el objeto multimedia 1004 comprende una imagen de un gato de aspecto gruñón con la línea de etiqueta: "Una vez me divertí. Esto fue horrible". Esta imagen es una referencia al popular meme de Internet "el gato gruñón", el cual puede ser reconocible instantáneamente por el usuario y el cual también puede servir para divertir e involucrar al usuario.

Los ejemplos anteriores de las Figuras 5-10 se han proporcionado aquí a manera de ejemplo únicamente. Los expertos en la técnica(s) relevante apreciarán que una amplia variedad de respuestas para las pronunciaciones de tipo charla pueden proporcionarse diferentes a aquellas descritas en las Figuras 5-10. Tales respuestas pueden incluir cualquiera de una o más de los componentes de respuesta previamente descritos con referencia a la Figura 4, además de componentes adicionales.

III. Métodos ilustrativos para implementar un asistente personal digital que utiliza personificaciones y/o multimedia en respuestas

La Figura 11 describe un cuadro de flujo 1100 de un método para implementar un asistente personal digital que utiliza personificaciones y/o multimedia en respuesta a las pronunciaciones de tipo charla de acuerdo con una modalidad. El método del cuadro de flujo 1100 puede realizarse, por ejemplo, por el respaldo de

asistente personal digital 106 como se discutió anteriormente con referencia a la Figura 1. Por consiguiente, el método del cuadro de flujo 1100 ahora se describirá con referencia continua al sistema 100 de la Figura 1. Sin embargo, el método no se limita a esa
5 implementación.

Como se muestra en la Figura 11, el método del cuadro flujo 1100 comienza en el paso 1102 en el cual se recibe una representación digital de una pronunciación de un usuario de un asistente personal digital. La representación digital de la
10 pronunciación puede comprender, por ejemplo, la pronunciación que se genera por el servicio de reconocimiento por voz 132 o la pronunciación corregida generada a través de la interacción de usuario con el asistente personal digital 130 como se discutió anteriormente con referencia a la Figura 2. Como además se discute
15 con respecto a esa Figura, la representación digital de la pronunciación se recibe por el sistema de entendimiento de consulta y respuesta 136.

En el paso 1104, se realiza la representación digital de la pronunciación. Como se discutió anteriormente con referencia a la
20 Figura 2, este paso puede implicar el análisis realizado por el sistema de entendimiento de consulta y respuesta 136 para determinar si la pronunciación comprende una pronunciación de tipo charla.

En el paso 1106, basándose en al menos el análisis de la
25 representación digital de la pronunciación, se determina que la

pronunciación comprende una pronunciación que pretende involucrarse con un personaje del asistente personal digital. Como se discutió anteriormente con referencia a la Figura 2, este paso ocurre cuando el sistema de entendimiento de consulta y respuesta
5 136 determina que la pronunciación es una pronunciación de tipo charla. Como se observó previamente, esta determinación puede basarse en el análisis de la pronunciación realizada por el sistema de entendimiento de consulta y respuesta 136.

En el paso 1108, en respuesta a al menos la determinación
10 hecha en el paso 1106, una respuesta para la pronunciación se causa al generarse por el asistente personal digital que incluye al menos uno de un objeto multimedia asociado con una referencia de cultura popular y la voz que comprende una personificación de una voz de una persona asociada con la referencia de cultura popular.
15 Este objeto multimedia puede comprender, por ejemplo, una imagen, contenido de vídeo, o contenido de audio, y puede presentarse dentro o reproducirse después por una interfaz de usuario del asistente personal digital. La voz puede generarse o reproducirse después por el asistente personal digital. Este paso puede
20 realizarse, por ejemplo, por el sistema de entendimiento de consulta y respuesta para charla 138 el cual causa que el asistente personal digital 130 proporcione una respuesta que incluya un objeto multimedia, tal como el componente multimedia 404 descrito anteriormente con referencia a la Figura 4, o que incluye la voz
25 personificada como se indicó por el componente de personificación

por voz 406 como se describió anteriormente con referencia a la Figura 4. La forma en la cual el sistema de entendimiento de consulta y respuesta para charla 138 realiza esta función se ha descrito previamente.

5 Como además se describió previamente, la respuesta a la pronunciación discutida en el paso 1108 puede además incluir texto que se presenta dentro de la interfaz de usuario del asistente personal digital, el texto que comprende una cita asociada con la referencia de cultura popular, una representación visual del asistente personal digital (por ejemplo, una animación del asistente personal digital) que se presenta dentro de la interfaz de usuario del mismo y que invoca la referencia de cultura popular, y/o un enlace que se presenta dentro de la interfaz de usuario del asistente personal digital y que puede activarse por el usuario al acceder al contenido, tal como el contenido asociado con la pronunciación de tipo charla o con la referencia de cultura popular. Adicionalmente, la voz que se genera o reproduce después por el asistente digital personal puede comprender una cita asociada con la referencia de cultura popular.

20 En una modalidad, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende identificar una pluralidad de respuestas elegibles para la pronunciación y luego seleccionar la respuesta para la pronunciación desde entre la pluralidad de respuestas elegibles para la pronunciación.

25 En otra modalidad, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende coincidir la

pronunciación a un tipo de pronunciación particular dentro de un árbol jerárquico de tipos de pronunciación (por ejemplo, árbol jerárquico 300 como se discutió anteriormente con referencia a la Figura 3), cada tipo de pronunciación en el árbol jerárquico de tipos de pronunciación que tienen una o más respuestas asociadas del mismo. Después de coincidir, la respuesta a la pronunciación se selecciona de entre la respuesta(s) asociada con el tipo de pronunciación particular.

Incluso en otra modalidad, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende determinar que una pronunciación está asociada con una tópico de tendencia y entonces, en respuesta a determinar que la pronunciación está asociada con una tópico de tendencia, seleccionar la respuesta a la pronunciación de entre una o más respuestas asociadas con la tópico de tendencia.

En una modalidad adicional, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende determinar que la pronunciación es una pronunciación para la cual existen una o más respuestas predefinidas que pretenden transportar al personaje del asistente personal digital y, en respuesta a esta determinación, seleccionar la respuesta para la pronunciación de entre la una o más respuestas predefinidas.

Incluso en una modalidad adicional, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende enviar un archivo de audio que incluye la voz o información que identifica el

archivo de audio a un dispositivo de cómputo que ejecuta el asistente personal digital. Este paso puede realizarse, por ejemplo, cuando el sistema de entendimiento de consulta y respuesta para charla 138 envía un archivo de audio que incluye la voz personificada para el dispositivo de cómputo de usuario final 102 para que pueda accederse y reproducirse después por el asistente personal digital 130 o cuando el sistema de entendimiento de consulta y respuesta para charla 138 envía la información que identifica tal archivo de audio al asistente personal digital 130 para que el archivo de audio pueda obtenerse localmente por el asistente personal digital 130.

Incluso en otra modalidad, el hacer que una respuesta a la pronunciación sea generada en el paso 1108 comprende proporcionar texto para un dispositivo de cómputo que ejecuta el asistente personal digital, en donde el texto se procesa por el componente de texto a voz del asistente personal digital para generar la voz. Este paso puede realizarse, por ejemplo, cuando el sistema de entendimiento de consulta y respuesta para charla 138 envía una indicación al asistente personal digital 130 de que el asistente digital personal 130 debe aplicar un convertidor de texto a voz especial al de designar texto para hacer que el texto se convierta a voz en una forma que personifique la voz de una persona particular.

La Figura 12 describe un cuadro de flujo 1200 de un método por el cual un asistente personal digital proporciona una respuesta para una pronunciación de tipo charla que incluye una

personificación de voz de acuerdo con una modalidad. El método del cuadro de flujo 1200 puede realizarse, por ejemplo, por el asistente personal digital 130 como se discutió anteriormente con referencia a la Figura 1. Por consiguiente, el método del cuadro de flujo 1200
5 ahora se describirá con referencia continua al sistema 100 de la Figura 1. Sin embargo, el método no se limita a esa implementación.

Como se muestra en la Figura 12, el método del cuadro de flujo 1200 comienza en el paso 1202, en el cual el asistente personal digital 130 captura audio que representa una pronunciación de un
10 usuario que pretende involucrarse con un personaje del asistente personal digital 130.

En el paso 1204, el asistente personal digital 130 transmite el audio al respaldo de asistente personal digital 106.

En el paso 1206, el asistente personal digital 130 proporciona
15 una respuesta para la pronunciación basándose al menos en la información recibida del respaldo de asistente personal digital 106. Proporcionar la respuesta incluye generar o reproducir después la voz que comprende una personificación de una voz de una persona asociada con una referencia de cultura popular.

20 En una modalidad, proporcionar la respuesta en el paso 1206 incluye reproducir después un archivo de audio que incluye la voz. En una modalidad alternativa, proporcionar la respuesta en el paso 1206 incluye aplicar la conversión de texto a voz para el texto para generar la voz.

25 De acuerdo con modalidades adicionales, proporcionar la

respuesta en el paso 1206 incluye uno o más de: presentar o reproducir después un objeto multimedia por una interfaz de usuario del asistente personal digital 130, el objeto multimedia estando asociado con la referencia de cultura popular; presentar texto dentro
5 de la interfaz de usuario del asistente personal digital, el texto comprende una cita asociada con la referencia de cultura popular; presentar una representación visual del asistente personal digital que invoca la referencia de cultura popular dentro de la interfaz de usuario del asistente personal digital; y presentar un enlace dentro
10 de la interfaz de usuario del asistente personal digital que puede activarse por el usuario al acceder contenido, tal como contenido asociado con la pronunciación o con la referencia de cultura popular.

La Figura 13 describe un cuadro de flujo 1300 de un método por el cual un asistente personal digital proporciona una respuesta
15 para una pronunciación de tipo charla que incluye un objeto multimedia de acuerdo con una modalidad. El método del cuadro flujo 1300 puede realizarse, por ejemplo, por el asistente personal digital 130 como se discutió anteriormente con referencia a la Figura 1. Por consiguiente, el método del cuadro de flujo 1300 ahora se describirá
20 con referencia continua al sistema 100 de la Figura 1. Sin embargo, el método no se limita a esa implementación.

Como se muestra en la Figura 13, el método del cuadro de flujo 1300 comienza en el paso 1302 en el cual el asistente personal digital 130 captura audio que representa una pronunciación de un
25 usuario que pretende involucrarse como un personaje del asistente

personal digital 130.

En el paso 1304, el asistente personal digital 130 transmite el audio al respaldo del asistente personal digital 106.

En el paso 1306, el asistente personal digital 130 proporciona
5 una respuesta a la pronunciación basándose al menos en la información recibida del respaldo de asistente personal digital 106. Proporcionar la respuesta incluye presentar o reproducir después un objeto multimedia asociado con la referencia de cultura popular por una interfaz de usuario del asistente personal digital 130.

10 En una modalidad, presentar o reproducir después el objeto multimedia en el paso 1306 comprende presentar una imagen o contenido de video o reproducir después contenido de audio por la interfaz de usuario del asistente personal digital 130.

De acuerdo con modalidades adicionales, proporcionar la
15 respuesta en el paso 1306 incluye uno o más de: generar o reproducir después la voz que comprende una personificación de una voz una persona asociada con la referencia de cultura popular; presentar el texto dentro de la interfaz de usuario del asistente personal digital, el texto comprende una cita asociada con la
20 referencia de cultura popular; presentar una representación visual del asistente personal digital que invoca la referencia de cultura popular dentro de la interfaz de usuario del asistente personal digital; y presentar un enlace dentro de la interfaz de usuario del asistente personal digital que puede activarse por el usuario para
25 acceder al contenido, tal como el contenido asociado con la

pronunciación o con la referencia de cultura popular.

IV. Implementación de dispositivo móvil ilustrativa

La Figura 14 es un diagrama de bloques de un dispositivo móvil
5 ilustrativo 1402 que puede utilizarse para implementar el dispositivo
de cómputo de usuario final 102 como se describió anteriormente con
referencia a la Figura 1. Como se muestra en la Figura 14, el
dispositivo móvil 1402 incluye una variedad de componentes de
hardware y *software* opcionales. Cualquier componente en el
10 dispositivo móvil 1402 puede comunicarse con cualquier otro
componente, aunque no todas las conexiones se muestran para
facilitar la ilustración. El dispositivo móvil 1402 puede ser cualquiera
de una variedad de dispositivos de cómputo (por ejemplo, teléfono
celular, teléfono inteligente, computadora portátil, asistente digital
15 personal (PDA), etc.) y puede permitir comunicaciones
bidireccionales inalámbricas con una o más redes de comunicaciones
móviles 1404, tal como una red celular o satelital, o con una red de
local o de área amplia.

El dispositivo móvil 1402 ilustrado puede incluir un controlador
20 o procesador 1410 (por ejemplo, procesador de señal,
microprocesador, ASIC, u otro control y circuito de lógica de
procesamiento) para realizar varias tareas tales como codificación de
señal, procesamiento de datos, procesamiento de entrada/salida,
control de energía, y/u otras funciones. Un sistema operativo 1412
25 puede controlar la asignación y uso de los componentes del

dispositivo móvil 1402 y soportar uno o más programas de aplicación 1414 (también denominados como "aplicaciones" o "apps"). Los programas de aplicación 1414 pueden incluir aplicaciones de cómputo móvil comunes (por ejemplo, aplicaciones de correo electrónico, calendarios, administradores de contacto, navegadores web, aplicaciones de mensajería) y cualquiera de las otras aplicaciones de cómputo (por ejemplo, aplicaciones de procesamiento de palabra, aplicaciones de asignación, aplicaciones de reproductor de medios). En una modalidad, los programas de aplicación 1414 incluyen el asistente personal digital 130.

El dispositivo móvil 1402 ilustrado puede incluir memoria 1420. La memoria 1420 puede incluir memoria no removible 1422 y/o memoria removible 1424. La memoria no removible 1422 puede incluir RAM, ROM, memoria flash, un disco duro, u otros dispositivos de memoria bien conocidos o tecnologías. La memoria removible 1424 puede incluir memoria flash o una tarjeta de módulo de identificación de suscriptor (SIM), la cual es bien conocida en los sistemas de comunicación GSM, u otros dispositivos de memoria bien conocidos o tecnologías, tales como "tarjetas inteligentes". La memoria 1420 puede utilizarse para almacenar datos y/o codificar para ejecutar el sistema operativo 1412 y aplicaciones 1414. Datos ilustrativos pueden incluir páginas web, texto, imágenes, archivos de sonido, datos de video, u otros datos para enviarse a y/o recibirse desde uno más servidores de red u otros dispositivos a través de una o más redes por cablea o inalámbricos. La memoria 1420 puede

utilizarse para almacenar un identificador de suscriptor, tal como una identidad de suscriptor móvil internacional (IMSI), y un identificador de equipo, tal como un identificador de equipo móvil internacional (IMEI). Tales identificadores pueden transmitirse a un servidor de red para identificar a usuarios y equipo.

El dispositivo móvil 1402 puede soportar uno o más dispositivos de entrada 1430, tal como una pantalla táctil 1432, un micrófono 1434, una cámara 1436, un teclado físico 1438 y/o una seguidora 1440 y uno o más dispositivos de salida 1450, tal como una bocina 1452 y una presentación 1454. Las pantallas táctiles tal como la pantalla táctil 1432, pueden detectar entrada en diferentes formas. Por ejemplo, las pantallas táctiles capacitivas detectan entrada táctil cuando un objeto (por ejemplo, una punta del dedo) distorsiona o interrumpe una corriente eléctrica que corre a través de la superficie. Como otro ejemplo, las pantallas táctiles pueden utilizar sensores ópticos para detectar entrada táctil cuando los haces de los sensores ópticos se interrumpen. El contacto físico con la superficie de la pantalla no es necesario para que la entrada sea detectada por algunas pantallas táctiles.

Otros dispositivos de salida posibles (no mostrados) pueden incluir dispositivos piezoeléctricos u otros de salida hápticos. Algunos dispositivos pueden servir más que una función de entrada/salida. Por ejemplo, la pantalla táctil 1432 y la presentación 1454 pueden combinarse en un dispositivo de entrada/salida individual. Los dispositivos de entrada 1430 pueden incluir una

interfaz de usuario natural (NUI).

El módem(s) inalámbrico 1460 puede acoplarse a la antena(s) (no mostrada) y puede soportar comunicaciones bidireccionales entre el procesador 1410 y los dispositivos externos, como se entiende bien en la técnica. El módem(s) 1460 se muestra genéricamente y puede incluir un módem celular 1466 para comunicarse con la red de comunicación móvil 1404 y/u otros módems basados en radio (por ejemplo, Bluetooth 1464 y/o Wi-Fi 1462). Al menos uno de los módem(s) inalámbricos 1460 típicamente están configurados para comunicación con una o más redes celulares, tal como una red GSM para los datos y comunicaciones por voz dentro de una red celular individual, entre las redes celulares, o entre el dispositivo móvil y una red de teléfono conmutada pública (PSTN).

El dispositivo móvil 1402 puede además incluir al menos un puerto de entrada/salida 1480, un suministro de energía 1482, un receptor de sistema de navegación de satélite 1484, tal como un receptor de sistema de posicionamiento global (GPS), un acelerómetro 1486, y/o un conector físico 1490, el cual puede ser un puerto USB, puerto IEEE 1394 (FireWire), y/o puerto RS-232. Los componentes ilustrados del dispositivo móvil 1402 no requieren o todos son inclusivos, ya que cualquiera de los componentes puede eliminarse y otros componentes pueden agregarse como podría reconocerse por un experto en la técnica.

En una modalidad, ciertos componentes del dispositivo móvil 1402 están configurados para realizar las operaciones atribuidas al

asistente personal digital 130 como se describió en las secciones precedentes. La lógica de programa de computadora para realizar las operaciones atribuidas al asistente personal digital 130 como se describió anteriormente pueden almacenarse en la memoria 1420 y
5 ejecutarse por el procesador 1410. Al ejecutar tal lógica de programa de computadora, el procesador 1410 puede implementar cualquiera de las características del asistente personal digital 130 como se describió anteriormente con referencia a la Figura 1. También, al ejecutar tal lógica de programa de computadora, el procesador 1410
10 puede realizar cualquiera o todos de los pasos de cualquiera o todos de los cuadros de flujo descritos en las Figuras 12 y 13.

V. Implementación de sistema de computadora ilustrativa

La Figura 15 describe un sistema de computadora basado en un
15 procesador ilustrativo 1500 que puede utilizarse para implementar varias modalidades aquí descritas. Por ejemplo, el sistema 1500 puede utilizarse para implementar el dispositivo de cómputo de usuario final 102 o cualquiera de las computadoras utilizadas para implementar el respaldo de asistente personal digital 106 como se
20 describió anteriormente con referencia a la Figura 1. El sistema 1500 también puede utilizarse para implementar cualquiera o todos de los pasos de cualquiera o todos de los cuadros de flujo descritos en las Figuras 11-13. La descripción del sistema 1500 aquí proporcionada se proporciona para propósitos de ilustración, y no pretende ser
25 limitante. Las modalidades pueden implementarse en tipos

adicionales de sistemas de computadora, como podría conocerse por aquellos expertos en la técnica(s) relevante.

Como se muestra en la Figura 15, el sistema 1500 incluye una unidad de procesamiento 1502, una memoria de sistema 1504, y un conductor común 1506 que acopla varios componentes de sistema que incluye la memoria de sistema 1504 a la unidad de procesamiento 1502. La unidad de procesamiento 1502 puede comprender uno o más microprocesadores o núcleos de microprocesador. El conductor común 1506 representa uno o más de cualquiera de los varios tipos de estructuras de conductor común, que incluyen conductor común de memoria o controlador de memoria, un conductor común periférico, un puerto de gráficos acelerado, y un conductor común de procesador local que utiliza cualquiera de una variedad de arquitecturas de conductor común. La memoria de sistema 1504 incluye memoria de sólo lectura (ROM) 1508 y memoria de acceso aleatorio (RAM) 1510. Un sistema de entrada/salida básica 1512 (BIOS) se almacena en la ROM 1508.

El sistema 1500 también tiene una o más de las siguientes unidades: una unidad de disco duro 1514 para leer de y escribir a un disco duro, una unidad de disco magnético 1516 para leer de o escribir a un disco magnético removible 1518, y una unidad de disco óptico 1520 para leer de o escribir a un disco óptico removible 1522 tal como un CD ROM, DVD ROM, disco BLU-RAY™ u otro medio óptico. La unidad de disco duro 1514, unidad de disco magnético 1516, y unidad de disco óptico 1520 están conectadas al conductor

común 1506 por una interfaz de unidad de disco duro 1524, una interfaz de unidad de disco magnético 1526 y una interfaz de unidad óptica 1528, respectivamente. Las unidades y sus medios legibles por computadora asociados proporcionan almacenamiento no volátil de las instrucciones legibles por computadora, estructuras de datos, módulos de programa y otros datos para la computadora. Aunque un disco duro, un disco magnético removible y un disco óptico removible se describen, otros tipos de dispositivos de memoria legible por computadora y estructuras de almacenamiento pueden utilizarse para almacenar datos, tales como tarjetas de memoria flash, discos de video digital, memorias de acceso aleatorio (RAM), memoria de sólo lectura (ROM), y similares.

Un número de módulos de programa pueden almacenarse en el disco duro, disco magnético, disco óptico, ROM, o RAM. Estos módulos de programa incluyen un sistema operativo 1530, uno o más programas de aplicación 1532, otros módulos de programa 1534, y datos de programa 1536. De acuerdo con varias modalidades, los módulos de programa pueden incluir lógica de programa de computadora que se ejecutable por la unidad de procesamiento 1502 para realizar cualquiera de todas de las funciones y características del dispositivo de cómputo usuario final 102 o cualquiera de las computadoras utilizadas para implementar respaldo de asistente personal digital 106 como se describió anteriormente con referencia a la Figura 1. Los módulos de programa también pueden incluir lógica de programa de computadora que, cuando se ejecuta por la

unidad de procesamiento 1502, realiza cualquiera de los pasos u operaciones mostradas o descritas con referencia a los cuadros de flujo de las Figuras 11-13.

Un usuario puede ingresar comandos e información en el sistema 1500 a través de los dispositivos de entrada tales como un teclado 1538 y un dispositivo de señalamiento 1540. Otros dispositivos de entrada (no mostrados) pueden incluir un micrófono, palanca de mandos, controlador de juego, escáner, o similares. En una modalidad, se proporciona una pantalla táctil en conjunto con una presentación 1544 para permitir a un usuario proporcionar la entrada al usuario a través de la aplicación de un tacto (como por un dedo o estilete, por ejemplo) a uno o más puntos de la pantalla táctil. Estos y otros dispositivos de entrada frecuentemente se conectan a la unidad de procesamiento 1502 a través de una interfaz de puerto en serie 1542 que está acoplada al conductor común 1506, pero puede conectarse mediante otras interfaces, tal como un puerto paralelo, puerto de juegos, o un conductor común en serie universal (USB). Tales interfaces pueden ser por cable o interfaces inalámbricas.

Una presentación 1544 también se conecta al conductor común 1506 a través de una interfaz, tal como un adaptador de video 1546. Además de la presentación 1544, el sistema 1500 puede incluir otros dispositivos de salida periféricos (no mostrados) tales como bocinas e impresoras.

El sistema 1500 está conectado una red 1548 (por ejemplo, una

red de área local o red de área ancha tal como el Internet) a través de una interfaz o adaptador de red 1550, un módem 1552 u otros medios adecuados para establecer comunicaciones a través de la red. El módem 1552, el cual puede ser interno o externo se conecta
5 al conductor común 1506 a través de la interfaz de puerto en serie 1542.

Como se utiliza aquí, los términos "medio de programa de computadora", "medio legible por computadora", y "medio de almacenamiento legible por computadora" se utilizan para
10 generalmente referirse a los dispositivos de memoria o estructuras de almacenamiento tales como el disco duro asociado con la unidad de disco duro 1514, disco magnético removible 1518, disco óptico removible 1522, además de otros dispositivos de memoria o estructuras de almacenamiento tales como tarjetas de memoria flash,
15 discos de video digital, memorias de acceso aleatorio (RAM), memorias de sólo lectura (ROM), y similares. Tales medios de almacenamiento legibles por computadora se distinguen de y no sobrepasan con el medio de comunicación (no incluyen medios de comunicación). Los medios de comunicación típicamente incorporan
20 instrucciones legibles por computadora, estructuras de datos, módulos de programa u otros datos en una señal de datos modulada tal como una onda portadora. El término "señal de datos modulada" se refiere a una señal que tiene una o más de sus características configuradas o cambiadas de tal forma que la información se codifica
25 en la señal. A manera de ejemplo, y no de limitación, los medios de

comunicación incluyen medios inalámbricos tales como acústicos, RF, infrarrojos y otros medios inalámbricos. Las modalidades también se dirigen a tales medios de comunicación.

Como se observó anteriormente, los programas de computadora y módulos (que incluyen programas de aplicación 1532 y otros módulos de programa 1534) pueden almacenarse en el disco duro, disco magnético, disco óptico, ROM, o RAM. Tales programas de computadora también pueden recibirse a través de una interfaz de red 1550, interfaz de puerto en serie 1542, o cualquier otro tipo de interfaz. Tales programas de computadora, cuando se ejecutan o cargan por una aplicación, permiten a la computadora 1500 implementar características de las modalidades de la presente invención aquí discutidas. Por consiguiente, tales programas de computadora representan controladores del sistema 1500.

Las modalidades también están dirigidas a los productos de programa de computadora que comprenden *software* almacenado en cualquier medio utilizable de computadora. Tal *software*, cuando se ejecuta en uno o más dispositivos de procesamiento de datos, causa que un dispositivo(s) de procesamiento de datos opere como se describe aquí. Las modalidades de la presente invención emplean cualquier medio utilizable por computadora legible por computadora, conocido ahora o en el futuro. Ejemplos de medios legibles por computadora incluyen, pero no se limitan a dispositivos de memoria y estructuras de almacenamiento tales como RAM, unidades duras, discos flexibles, CD-ROM, DVD-ROM, discos Zip, cintas, dispositivos

de almacenamiento magnético, dispositivos de almacenamiento óptico, MEM, dispositivos de almacenamiento basados en nanotecnología, y similares.

En implementaciones alternativas, el sistema 1500 puede
5 implementarse como circuitería de lógica/eléctrica de *hardware* o *firmware*. De acuerdo con las modalidades adicionales, uno o más de estos componentes pueden implementarse en un sistema en chip (SoC). El SoC puede incluir un chip de circuito integrado que incluye uno o más de un procesador (por ejemplo, un microcontrolador,
10 microprocesador, procesador de señal digital (DSP), etc.), memoria, una o más interfaces de comunicación, y/o circuitos adicionales y/o *firmware*, incorporado para realizar sus funciones.

VI. Conclusión

15 Aunque varias modalidades se han descrito anteriormente, se entenderá que pueden presentarse a manera únicamente de ejemplo, y no de limitación. Se apreciará por aquellos expertos en la técnica(s) relevante que varios cambios en la forma y detalles pueden hacerse aquí sin apartarse del espíritu y alcance de la
20 invención. De esa forma, la amplitud y alcance de la presente invención podría no limitarse por cualquiera de las modalidades ilustrativas descritas anteriormente, pero podrían definirse únicamente de acuerdo con las siguientes reivindicaciones y sus equivalentes.