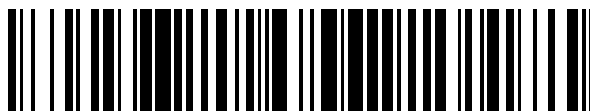


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 692 828**

51 Int. Cl.:

**H04R 25/00** (2006.01)

**G10L 15/26** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.09.2015** E 15770906 (4)

97 Fecha y número de publicación de la concesión europea: **15.08.2018** EP 3149968

54 Título: **Procedimiento de asistencia en el seguimiento de una conversación para una persona con problemas de audición**

30 Prioridad:

**29.09.2014 FR 1459203**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**05.12.2018**

73 Titular/es:

**GUEDON, CHRISTOPHE (100.0%)**  
**7 Avenue de la République**  
**92500 Rueil Malmaison, FR**

72 Inventor/es:

**GUEDON, CHRISTOPHE**

74 Agente/Representante:

**CURELL AGUILÁ, Mireia**

ES 2 692 828 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

**DESCRIPCIÓN**

Procedimiento de asistencia en el seguimiento de una conversación para una persona con problemas de audición.

5

**Campo técnico general**

La invención tiene por objeto el campo de la asistencia a las personas con problemas de audición, y se refiere de manera más precisa a un procedimiento de asistencia en el seguimiento de una conversación que mejora la selección de los interlocutores para una persona con problemas de audición.

10

**Estado de la técnica**

Las personas con problemas de audición están equipadas generalmente con prótesis auditivas que realizan una amplificación de los sonidos con el fin de compensar sus dificultades de audición.

15

Unas prótesis de este tipo están adaptadas al seguimiento de una conversación con un número restringido de interlocutores en un entorno silencioso pero una persona con problemas de audición puede continuar teniendo problemas para seguir una conversación, a pesar de la utilización de este tipo de prótesis, en un entorno en el que se pueden superponer a la conversación numerosos ruidos parásitos. Este tipo de situación puede presentarse cuando la conversación tiene lugar en un entorno ruidoso, por ejemplo en la calle o en una sala en la que numerosas personas hablan de manera simultánea, tal como una cafetería.

20

Algunos sistemas de asistencia a las personas con problemas de audición existentes como el propuesto en la solicitud de patente US 2008295040 proponen proporcionar a su usuario una transcripción textual de las palabras de los interlocutores, con el fin de que el usuario pueda leer sus expresiones. Sin embargo, este tipo de sistema es eficaz únicamente cuando el número de interlocutores es pequeño y puede resultar inaprovechable cuando el aparato capta las voces de interlocutores que participan en varias conversaciones.

25

Algunas prótesis disponen de micrófonos direccionales o de un sistema de selección que permite amplificar preferentemente la voz de algunos interlocutores que participan en la conversación que el usuario con problemas de audición desea escuchar. Unos sistemas de este tipo presentan sin embargo el fallo de no permitir que una persona con problemas de audición siga una primera conversación y continúe escuchando al mismo tiempo en un plano secundario el desarrollo de una segunda conversación. Además, en los casos de micrófonos direccionales, a una persona con problemas de audición usuaria le resulta imposible seguir las expresiones simultáneas de personas situadas en direcciones opuestas que participan en la misma conversación.

30

35

Existe por lo tanto una necesidad de un procedimiento de asistencia en el seguimiento de una conversación que permita que un usuario con problemas de audición siga eficazmente una primera conversación entre varios interlocutores incluso si éstos están dispuestos en direcciones opuestas, sin ser molestado por el ruido ambiental, y continúe conociendo al mismo tiempo el desarrollo de por lo menos una segunda conversación que se desarrolla en segundo plano.

40

**Presentación de la invención**

La presente invención se refiere así según un primer aspecto a un procedimiento de asistencia en el seguimiento de una conversación según la reivindicación 1 con una pluralidad de interlocutores para una persona con problemas de audición provista de un equipo informático, estando dicho equipo informático unido a por lo menos un módulo de emisión de voz, a por lo menos un dispositivo de visualización y a por lo menos un módulo de conversión de señales de voz en señales de texto, comprendiendo dicho procedimiento las etapas siguientes:

45

50

- adquisición de una pluralidad de señales de voz y de una pluralidad de señales de texto procedentes de una pluralidad de interlocutores, siendo dichas señales de texto obtenidas por conversión de dichas señales de voz por el módulo de conversión de señales;
- selección de por lo menos una primera señal de voz de entre la pluralidad de señales de voz adquiridas;
- selección de por lo menos una primera señal de texto de entre la pluralidad de señales de texto adquiridas, procediendo la señal de voz y la señal de texto seleccionadas de interlocutores diferentes;
- transmisión al módulo de emisión de voz de las primeras señales de voz seleccionadas para la emisión de dichas primeras señales de voz;
- orden de visualización en el dispositivo de visualización de las primeras señales de texto seleccionadas.

55

60

65

La persona con problemas de audición puede seguir así eficazmente varias conversaciones, en particular una conversación principal en forma de audio y/o unas conversaciones secundarias en forma de texto.

Según un primer modo de realización, estando por lo menos uno de dichos interlocutores equipado con un equipo informático de interlocutor que comprende por lo menos un micrófono y dicho módulo de conversión de señales de voz en señales de texto, el procedimiento según el primer aspecto comprende las etapas siguientes realizadas por dicho por lo menos un equipo informático de interlocutor:

- captura de señales de voz del interlocutor con la ayuda del micrófono de dicho equipo informático de interlocutor,
- conversión de las señales de voz capturadas en señales de texto por el módulo de conversión de dicho equipo informático de interlocutor,
- transmisión a dicho equipo informático de la persona con problemas de audición de las señales de voz capturadas y de las señales de texto obtenidas por conversión.

Como el equipo informático de la persona con problemas de audición comprende un módulo de recepción de señales y es susceptible de ser conectado a una red de comunicación de voz y a una red de comunicación de datos, dicha etapa de adquisición de señales de interlocutores puede comprender entonces la recepción de las señales de voz y de las señales de texto de la pluralidad de interlocutores por el módulo de recepción de señales por medio de las redes de comunicaciones de voz y de datos.

De esta manera, si cada interlocutor dispone de su propio equipo informático, las expresiones de cada interlocutor pueden ser capturadas localmente por su equipo, en forma de voz y de texto, y puestas a disposición de los equipos de los demás interlocutores y de la persona con problemas de audición que participa en la conversación.

Según un segundo modo de realización, como el equipo informático de la persona con problemas de audición comprende por lo menos un micrófono y dicho módulo de conversión de señales de voz en señales de texto, dicha etapa de adquisición de señales de interlocutores comprende las etapas siguientes:

- adquisición de las señales de voz de la pluralidad de interlocutores con la ayuda del micrófono de dicho equipo informático de la persona con problemas de audición,
- conversión de las señales de voz adquiridas en señales de texto por dicho módulo de conversión.

El equipo informático puede ser así perfectamente autónomo y adquirir las señales de voz y de texto del conjunto de los interlocutores sin que cualquier otra persona que no sea la persona con problemas de audición necesite estar equipado. La persona con problemas de audición puede así seguir cualquier conversación en cualquier entorno con cualquier interlocutor.

Según una primera variante, como el equipo informático de la persona con problemas de audición está unido a una base de datos de firmas vocales de interlocutores conocidos, las selecciones de las primeras señales de voz y de texto se realizan con la ayuda de un reconocimiento de las señales de voz de los interlocutores conocidos de entre las señales de voz adquiridas a partir de dicha base de datos.

El reconocimiento vocal permite que el equipo informático distinga a los diferentes interlocutores de una conversación y proponga a la persona con problemas de audición las palabras de estos interlocutores en forma de señales de voz y de texto separadas para cada interlocutor. Aunque las señales de varios interlocutores puedan ser adquiridas al mismo tiempo por el mismo micrófono, la persona con problemas de audición puede elegir así seguir las expresiones solamente de una parte de los interlocutores.

Según una segunda variante, la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende:

- una etapa previa de preselección por parte del equipo informático de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas en función de la distancia entre el interlocutor emisor de dicha señal de voz, respectivamente señal de texto, y la persona con problemas de audición,
- una etapa de selección de primeras señales de voz, respectivamente señales de texto, de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas.

La selección de los interlocutores puede resultar así más sencilla para la persona con problemas de audición preseleccionando los interlocutores en un intervalo de distancia particular, por ejemplo los más cercanos a la persona con problemas de audición.

Según una tercera variante, como el equipo como el equipo informático de la persona con problemas de audición

está unido a un dispositivo de detección de movimientos de la persona con problemas de audición, la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende:

- 5 - una etapa previa de preselección por parte del equipo informático de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas en función de un movimiento de la persona con problemas de audición detectado por el dispositivo de detección de movimientos,
- 10 - una etapa de selección de primeras señales de voz, respectivamente señales de texto, de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas.

La persona con problemas de audición puede hacer evolucionar así simplemente la preselección de los interlocutores gracias a unos movimientos sencillos, tal como unos movimientos de cabeza, sin tener que introducir una orden.

15 Según una cuarta variante, como el equipo informático de la persona con problemas de audición está unido a por lo menos una cámara que captura una imagen de los interlocutores en el campo de visión de la persona con problemas de audición y comprende un módulo de reconocimiento de imagen que permite identificar a dichos interlocutores en dicha imagen capturada por dicha cámara, la etapa de selección de primeras señales de voz, respectivamente de señales de texto, comprende:

- 25 - una etapa previa de preselección por parte del equipo informático de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de interlocutores, denominados interlocutores preseleccionados, presentes en el campo de visión de la persona con problemas de audición de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas,
- una etapa de identificación, por parte del módulo de reconocimiento, de dichos interlocutores preseleccionados,
- 30 - una etapa de visualización en el dispositivo de visualización de la imagen de dichos interlocutores preseleccionados en el campo de visión de la persona con problemas de audición, asociados a su identificación,
- 35 - una etapa de selección por parte de la persona con problemas de audición de primeras señales de voz, respectivamente de primeras señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, de los interlocutores preseleccionados visualizados e identificados.

La persona con problemas de audición puede así conocer inmediatamente la identidad de los interlocutores seleccionables y realizar su selección en función de su interés por las expresiones de un interlocutor particular al que ya conoce.

Según una quinta variante, el procedimiento según el primer aspecto comprende:

- 45 - una etapa de recepción de una orden vocal de la persona con problemas de audición relativa a un interlocutor y a un tipo de señales a seleccionar;
- una etapa de selección de primeras señales de voz o de texto del interlocutor en función de dicha orden vocal, comprendiendo el equipo informático de la persona con problemas de audición un módulo de reconocimiento vocal configurado para analizar dicha orden vocal.

50 La persona con problemas de audición puede ordenar así la selección de interlocutores mediante una sencilla orden vocal, sin tener que manipular el equipo informático o cualquier interfaz de introducción de datos manual.

55 Según una sexta variante, la selección de por lo menos una primera señal de voz y/o de una primera señal de texto se realiza en función de preferencias y/o usos de la persona con problemas de audición.

Una selección se puede realizar así automáticamente de manera que se seleccionen los interlocutores que la persona con problemas de audición ha seleccionado con más frecuencia en el pasado. Esto permite proponer automáticamente a la persona con problemas de audición que siga las expresiones de los interlocutores que presentan *a priori* el mayor interés para ella.

60 Como el equipo informático de la persona con problemas de audición está unido a por lo menos una cámara que captura una imagen de por lo menos un interlocutor de la persona con problemas de audición, tras la selección de primeras señales de voz, respectivamente de primeras señales de texto, de dicho interlocutor, dicho procedimiento puede comprender además una etapa de visualización de la imagen de dicho interlocutor situado fuera del campo de visión de la persona con problemas de audición en una viñeta insertada en la imagen visualizada en dicho dispositivo de visualización.

5 La persona con problemas de audición puede tener así permanentemente en su campo de visión una imagen de cada uno de los interlocutores de una conversación en la que toma parte, incluso si estos interlocutores están situados en direcciones opuestas. La comprensión de las expresiones de estos interlocutores mejora así gracias a la posibilidad de tener en cuenta las expresiones corporales, en particular faciales, de estos interlocutores.

10 Según una séptima variante, la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende la determinación de un primer número predeterminado, respectivamente de un segundo número predeterminado, y la selección de dicho primer número predeterminado de primeras señales de voz, respectivamente de dicho segundo número predeterminado de primeras señales de texto, por orden de recepción.

15 Esto permitirá que una persona con problemas de audición siga una conversación entre numerosos interlocutores en la que el interlocutor que tiene la palabra cambia con mucha frecuencia, sin que la persona con problemas de audición deba corregir manualmente en cada momento su selección de interlocutores.

20 Por otro lado, la etapa de emisión de las primeras señales de voz seleccionadas puede comprender previamente una etapa de filtrado del ruido y/o de las señales de voz no seleccionadas.

25 La calidad de las señales de voz suministradas a la persona con problemas de audición puede aumentar así de manera que se mejore la comprensión de la conversación.

30 Como el equipo informático de la persona con problemas de audición comprende un módulo de reconocimiento vocal, la etapa de filtrado puede comprender una etapa de filtrado de las señales de voz seleccionadas con la ayuda del resultado de un reconocimiento vocal por parte del módulo de reconocimiento vocal.

35 Este tipo de filtrado permite tratar de manera diferente las señales de voz de interlocutores deseados y de interlocutores no deseados de manera que se mejore la comprensión de las expresiones de los interlocutores deseados.

40 Según un segundo aspecto, la invención se refiere a un producto de programa de ordenador según la reivindicación 15 que comprende unas instrucciones de código para la realización de un procedimiento según el primer aspecto cuando este programa es ejecutado por un procesador.

45 Según un tercer aspecto, la invención se refiere a un equipo informático, según la reivindicación 16, para la asistencia en el seguimiento de una conversación con una pluralidad de interlocutores, en posesión de una persona con problemas de audición y unido a por lo menos un módulo de emisión de voz, a por lo menos un dispositivo de visualización y a por lo menos un módulo de conversión de señales de voz en señales de texto, estando dicho equipo informático configurado para:

- 40 - adquirir una pluralidad de señales de voz y una pluralidad de señales de texto procedentes de una pluralidad de interlocutores, siendo dichas señales de texto obtenidas por conversión de dichas señales de voz por el módulo de conversión de señales;
- 45 - seleccionar por lo menos una primera señal de voz de entre la pluralidad de señales de voz adquiridas;
- seleccionar por lo menos una primera señal de texto de entre la pluralidad de señales de texto adquiridas, procediendo la señal de voz y la señal de texto de interlocutores diferentes;
- 50 - transmitir al módulo de emisión de voz unas primeras señales de voz seleccionadas para la emisión de dichas primeras señales de voz;
- ordenar la visualización en el dispositivo de visualización de las primeras señales de texto seleccionadas.

55 Según un cuarto aspecto, la invención se refiere a unas gafas de realidad aumentada que comprenden un equipo informático según el tercer aspecto.

60 Dichos producto de programa de ordenador, equipo informático, y gafas, presentan las mismas ventajas que las mencionadas para el procedimiento según el primer aspecto.

### Presentación de las figuras

65 Otras características y ventajas aparecerán con la lectura de la descripción siguiente de un modo de realización. Esta descripción se dará haciendo referencia a los dibujos adjuntos, en los que:

- la figura 1 ilustra un ejemplo de arquitectura para la realización del procedimiento según un primer modo de realización;

- la figura 2 ilustra un ejemplo de arquitectura para la realización del procedimiento según un segundo modo de realización;
- 5        - la figura 3 ilustra un ejemplo de arquitectura para la realización de una etapa de filtrado según un modo de realización de la invención.

**Descripción detallada**

10        Con referencia a la figura 1, una forma de realización de la invención se refiere a un procedimiento de asistencia en el seguimiento de una conversación con una pluralidad de interlocutores 1 para una persona con problemas de audición 2 provista de un equipo informático 3.

15        Este procedimiento tiene por objetivo permitir que la persona con problemas de audición seleccione de entre la pluralidad de interlocutores aquéllos cuya voz quiere escuchar y aquéllos cuyas palabras transcritas quiere leer en forma de texto, de manera que pueda seguir una primera conversación en forma oral y una o varias conversaciones diferentes en forma de texto.

20        Para ello, este equipo informático 3 es susceptible de estar unido a por lo menos un módulo de emisión de voz 4, tal como un altavoz o un casco de audio, a por lo menos un dispositivo de visualización 5, a por lo menos un módulo de conversión de señales de voz en señales de texto 6, y a por lo menos un micrófono 7. Se entenderá por micrófono un módulo de adquisición de señales sonoras, cualquiera que sea la tecnología empleada para realizar este tipo de adquisición. Un micrófono de este tipo puede ser direccional para reducir la captación de las voces de interlocutores no deseados o ruido, o bien bidireccional u omnidireccional con el fin de capturar el conjunto de un ambiente sonoro. El equipo informático 3 puede comprender asimismo un módulo de tratamiento 8 tal como un microprocesador, y una interfaz de introducción de datos 9 que permite que la persona con problemas de audición 2 realice una selección o introduzca una orden. Según un modo de realización, la interfaz de introducción de datos está incluida en el dispositivo de visualización 5 en forma de una pantalla táctil.

25        El equipo informático 3 adquiere, en primer lugar, una pluralidad de señales de voz y una pluralidad de señales de texto procedentes de la pluralidad de interlocutores 1, siendo dichas señales de texto obtenidas por conversión de dichas señales de voz por el módulo de conversión de señales 6.

30        A continuación, la persona con problemas de audición 2 selecciona con la ayuda de su equipo informático 3 por lo menos una primera señal de voz de entre la pluralidad de señales de voz adquiridas. La persona con problemas de audición selecciona así los interlocutores cuyas expresiones desea oír en forma vocal con el fin de seguir su conversación.

35        La persona con problemas de audición 2 selecciona además con la ayuda de su equipo informático 3 por lo menos una primera señal de texto de entre la pluralidad de señales de texto adquiridas. La persona con problemas de audición selecciona así los interlocutores de los cuales desea obtener una transcripción de las palabras en forma de texto, con el fin de poder seguir sus conversaciones en segundo plano sin obstaculizar su comprensión de la conversación seguida en forma oral.

40        Por último, el equipo informático 3 transmite a dicho por lo menos un módulo de emisión de voz 4 las primeras señales de voz seleccionadas para que éstas sean emitidas, y ordena la visualización en el dispositivo de visualización 5 de las primeras señales de texto seleccionadas.

45        A continuación de la adquisición de las señales de voz y de texto de la pluralidad de interlocutores, éstos pueden ser presentados a la persona con problemas de audición 2, por ejemplo por medio de la pantalla de visualización, en forma de lista o de imagen, para permitir que haga sus selecciones de primeras señales de voz y de primeras señales de texto.

50        Según un primer modo de realización representado en la figura 1, cada interlocutor, incluida la persona con problemas de audición 2, dispone de su propio equipo informático 3 dedicado a la captura y a la transcripción de sus palabras. Cada equipo informático 3 comprende entonces su módulo de conversión de señales de voz en señales de texto 6 y utiliza una etapa de captura de señales de voz de su usuario con la ayuda de por lo menos un micrófono 7 del equipo informático 3, una etapa de conversión de las señales de voz capturadas en señales de texto por el módulo de conversión 6, y una etapa de transmisión a los demás equipos informáticos, incluyendo el equipo informático de la persona con problemas de audición, de las señales de voz capturadas y de las señales de texto obtenidas.

55        En este modo de realización, el equipo informático 3 de la persona con problemas de audición 2 comprende un módulo de recepción de señales 10 y está conectado a una red de comunicación de voz y a una red de comunicación de datos. La etapa de adquisición de señales de interlocutores puede entonces comprender una etapa de recepción de señales durante la cual el módulo de recepción de señales 10 recibe las señales de voz y de texto de la pluralidad de interlocutores por medio respectivamente de la red de comunicación de voz y de la

red de comunicación de datos.

5 Las señales de voz de cada interlocutor son así capturadas y transcritas en texto por su equipo informático 3, y después las señales de voz capturadas y las señales de texto correspondientes son transmitidas al equipo de la persona con problemas de audición 2 por medio de las redes de voz y de datos. En el caso de que varias personas con problemas de audición participen en una conversación, cada una puede disponer de un equipo que realiza al mismo tiempo por un lado la captura, la transcripción y la emisión de sus expresiones, y por otra parte la recepción de las señales que proceden de los otros equipos informáticos.

10 Las redes de comunicación de voz y de datos pueden consistir en unas redes alámbricas tales como unas redes Ethernet, o unas redes inalámbricas tales como unas redes Wifi, Bluetooth 3G o 4G o también en una combinación de las dos.

15 La persona con problemas de audición 2 puede proponer a la selección las señales de voz y de texto adquiridas en función de un identificador del equipo que ha realizado la adquisición de estas señales. Un identificador de este tipo puede ser arbitrario o corresponder a un parámetro tal como una dirección IP. En el caso de una instalación fija, que equipa por ejemplo una sala de reuniones o una mesa de comedor, la persona con problemas de audición 2 puede seleccionar así las señales de voz y de texto procedentes de un sitio particular en esta sala o en esta mesa, independientemente de la identidad de la persona que ocupa este sitio.

20 Según un segundo modo de realización representado en la figura 2, el equipo informático 3 de la persona con problemas de audición 2 es autónomo y realiza él mismo la adquisición y el tratamiento de las palabras de los interlocutores 1. El equipo informático 3 de la persona con problemas de audición 2 comprende entonces el módulo de conversión de señales de voz en señales de texto 6. La etapa de adquisición de las señales de interlocutores comprende entonces una etapa de adquisición de las señales de voz de la pluralidad de interlocutores con la ayuda del por lo menos un micrófono 7 del equipo informático de la persona con problemas de audición 3, y una etapa de conversión de las señales de voz adquiridas en señales de texto por dicho módulo de conversión 6.

30 En este modo de realización, con el fin de distinguir las palabras de los diferentes interlocutores de entre el conjunto de los sonidos capturados por el o los micrófonos 7, el equipo informático 3 puede estar unido a una base de datos de firmas vocales de interlocutores conocidos y el equipo informático de la persona con problemas de audición 3 puede comprender un módulo de reconocimiento vocal 11. Las selecciones de las primeras señales de voz y de texto se pueden realizar entonces con la ayuda de un reconocimiento de las señales de voz de los interlocutores conocidos de entre las señales de voz adquiridas a partir de dicha base de datos, realizado por el módulo de reconocimiento vocal 11. Cada señal de voz capturada por el micrófono 7 en un instante determinado puede ser comparada con las firmas de interlocutores conocidos almacenadas en la base de datos de firmas con el fin de presentar separadamente a la persona con problemas de audición 2 las palabras de cada interlocutor 1 así como su identidad, si ésta estuviera memorizada en la base de datos de firmas.

40 Es posible combinar los dos modos de realización descritos anteriormente con el fin de que las señales de voz de cada interlocutor sean capturadas al mismo tiempo a nivel del interlocutor y a nivel de la persona con problemas de audición. Las señales de voz capturadas de cada interlocutor, o las señales de texto correspondientes, pueden ser comparadas entonces con las señales de voz o de texto adquiridas a nivel de la persona con problemas de audición y distinguidas unas de otras por reconocimiento vocal. Una comparación de este tipo permite determinar si la voz de un interlocutor puede ser escuchada de manera inteligible a nivel de la persona con problemas de audición. Se puede emplear este tipo de determinación para detectar una toma de palabra por parte de un interlocutor remoto destinada a toda una asamblea, por ejemplo en el marco de una conferencia o de una asamblea. Cuando este interlocutor toma la palabra hablando con una voz fuerte, su voz resulta inteligible a nivel de la persona con problemas de audición mientras que este interlocutor no formaba parte antes de los interlocutores cercanos a la persona con problemas de audición y que su voz no llegaba hasta ésta. En caso de detección de este tipo de toma de palabra, un mensaje puede ser visualizado por el equipo informático en el dispositivo de visualización con el fin de prevenir a la persona con problemas de audición y se puede visualizar una transcripción de las palabras de este nuevo interlocutor.

55 Independientemente del modo de realización empleado de entre los modos descritos anteriormente, el equipo informático 3 puede consistir en un dispositivo portátil tal como un teléfono inteligente ("smartphone") o una tableta. Como alternativa, el equipo informático 3 puede estar integrado en unas gafas en forma de gafas de visualización de realidad aumentada. En este caso, las informaciones suministradas a la persona con problemas de audición 2 pueden ser visualizadas en los cristales de las gafas utilizados como dispositivo de visualización 5. La persona con problemas de audición puede entonces ver al mismo tiempo las informaciones visualizadas y su entorno en transparencia. Según otra variante, el equipo informático 3 está compartido entre unas gafas utilizadas para la visualización y que llevan algunos sensores, y un dispositivo portátil que comprende la unidad de tratamiento, la interfaz de introducción de datos y el resto de los elementos del equipo informático 3.

65 Según un modo de realización alternativo, las señales de texto no se obtienen por conversión de las señales de voz sino por reconocimiento labial. Las señales de texto que corresponden a las palabras de un interlocutor se

pueden obtener entonces mediante el módulo de conversión de señales a partir de una señal de vídeo de la cara del interlocutor capturada por lo menos por una cámara de vídeo unida a dicho módulo de conversión. Este tipo de conversión permite que la determinación de las señales de texto sea independiente de las perturbaciones acústicas que pueden afectar a la señal de voz correspondiente.

5

Cuando la cantidad de señales de texto a visualizar en el dispositivo de visualización es importante, el equipo informático puede sustituir la totalidad o parte del texto a visualizar por un resumen de éste. El texto puede ser resumido mediante unas frases sencillas o incluso mediante palabras aisladas que expresen conceptos. Este tipo de modificación del texto puede ser iniciada manualmente por el usuario del equipo informático, para uno o varios de sus interlocutores, o bien de manera automática cuando la cantidad de texto a visualizar sobrepasa un umbral determinado. La visualización de un texto resumido puede permitir disminuir el cansancio de la persona con problemas de audición en conversaciones rápidas o con numerosos interlocutores durante las cuales puede resultar difícil leer la totalidad de las expresiones durante largo tiempo.

10

15

Según un modo de realización, el equipo informático de la persona con problemas de audición 3 puede comprender un módulo de reconocimiento vocal 11. Se puede realizar entonces una etapa de recepción de una orden vocal de la persona con problemas de audición 2 relativa a un interlocutor y a un tipo de señales a seleccionar y una etapa de selección de primeras señales de voz o de texto del interlocutor en función de dicha orden vocal, estando entonces el módulo de reconocimiento vocal 11 configurado para analizar dicha orden vocal. La persona con problemas de audición puede interactuar así vocalmente con el equipo informático 3 para ordenar la selección de señales de voz o de texto relativas a un interlocutor particular. El reconocimiento por parte del equipo informático 3 de una orden vocal emitida por la persona con problemas de audición puede ser efectiva únicamente si la persona con problemas de audición ha efectuado previamente una acción particular, como realizar un movimiento particular o presionar un botón de interfaz, o mientras la persona con problemas de audición efectúa una acción de este tipo.

20

25

Como alternativa, la persona con problemas de audición puede interactuar con el equipo informático 3 mediante el pensamiento. El equipo informático puede estar unido por ejemplo a un casco neuronal colocado en el cráneo de la persona con problemas de audición y configurado para realizar un electroencefalograma de ésta. El usuario puede así someter al equipo informático unas órdenes mediante el pensamiento, por ejemplo unas órdenes sencillas tales como la selección de un interlocutor particular. La persona con problemas de audición puede así interactuar con su equipo informático sin utilizar las manos y sin molestar a las personas de su entorno como podría hacerlo una orden vocal.

30

35

Con el fin de facilitar la selección de las señales de voz y de texto de interés para la persona con problemas de audición 2 de entre el conjunto de las señales de voz y de texto adquiridas, se puede realizar una etapa de preselección durante la cual se seleccionan una o varias señales de voz o de texto de interlocutores automáticamente o en función de acciones de la persona con problemas de audición. Si esta preselección le conviene, la persona con problemas de audición puede ordenar a continuación simplemente la selección de las señales de voz y/o de texto de los interlocutores preseleccionados con una simple orden, por ejemplo una orden vocal tal como se ha descrito anteriormente o por medio de la interfaz de introducción de datos del equipo informático 3. En los párrafos siguientes un interlocutor, respectivamente un equipo remoto, cuyas señales de voz o de texto están preseleccionadas, se denominará interlocutor, respectivamente equipo, preseleccionado. La preselección de uno o varios interlocutores o equipos puede ser comunicada a la persona con problemas de audición 2 por medio del dispositivo de visualización 5. Por ejemplo, en el caso de una visualización en forma de lista se pueden poner en alta intensidad unos elementos de la lista, en el caso de una visualización en forma de mapa los interlocutores preseleccionados pueden estar representados por un icono particular, en el caso de una visualización de realidad aumentada los interlocutores preseleccionados pueden estar rodeados por un contorno.

40

45

50

Unas señales de voz, respectivamente de texto, se pueden preseleccionar de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas en función de la distancia entre el interlocutor emisor de dicha señal de voz, respectivamente señal de texto, y la persona con problemas de audición. Unas primeras señales de voz, respectivamente de texto, pueden ser seleccionadas a continuación de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas. La persona con problemas de audición puede así proponer para la selección de señales de voz o de texto los interlocutores más próximos o bien, en un segundo tiempo por ejemplo, los interlocutores situados en un intervalo de distancia predeterminado. Con el fin de proponer este tipo de preselección en función de la distancia, el equipo informático 3 debe tener conocimiento de la distancia que separa cada interlocutor o equipo remoto de la persona con problemas de audición. Para ello, la posición de cada interlocutor o equipo remoto puede ser determinada con respecto a la persona con problemas de audición por geolocalización.

55

60

Se puede realizar este tipo de geolocalización de manera absoluta con la ayuda de dispositivos de geolocalización externos tales como unos receptores GPS. Una geolocalización de este tipo se puede realizar asimismo con la ayuda de una instalación local. Por ejemplo, diferentes equipos pueden ser posicionados por triangulación con respecto a los puntos de acceso de una red local inalámbrica a la que están conectados, tal como una red Wifi. La geolocalización de un equipo también se puede realizar de manera relativa con respecto a otros objetos cuya posición absoluta es conocida. La posición de estos objetos se puede registrar por ejemplo en

65



una base de datos de coordenadas de objetos y el equipo a geolocalizar puede determinar su posición con respecto a estos objetos con la ayuda de un detector de tipo radar. Estos objetos pueden ser unos objetos conectados que declinan su identidad por medio de una red inalámbrica. Según otra variante, cada elemento a geolocalizar está equipado con un dispositivo de detección de movimiento que puede comprender por ejemplo un acelerómetro de tres ejes y/o un girómetro de tres ejes y/o un giroscopio. Cuando tiene lugar una inicialización, se puede registrar la posición de cada elemento a geolocalizar con respecto a un punto de referencia. La evolución de cada elemento con respecto a este punto de referencia puede ser seguida a continuación gracias a la detección de los movimientos de estos elementos por su dispositivo de detección de movimiento. La geolocalización a partir de una base de coordenadas de objetos o con la ayuda de un dispositivo de detección de movimiento puede permitir obtener una precisión de localización más importante que la obtenida por una localización GPS o Wifi. Esta precisión puede mejorar la eficacia de los tratamientos utilizados en función de esta localización.

Esta geolocalización se puede emplear asimismo con el fin de visualizar automáticamente en el dispositivo de visualización 5 las señales de texto correspondientes a los interlocutores más cercanos a la persona con problemas de audición. Este tipo de visualización automática se puede utilizar por ejemplo mientras la persona con problemas de audición no haya seleccionado ninguna señal de texto. Al llegar a una estancia en la que se desarrolla una conversación puede tener así una visión general de las palabras de los interlocutores en su entorno cercano sin tener que realizar una selección. Esta visualización se actualiza de manera dinámica y permite así que la persona con problemas de audición conozca las expresiones de personas en movimiento que se encuentran en su entorno cercano únicamente durante un breve periodo.

Se pueden preseleccionar señales de voz, respectivamente de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas en función de un movimiento de la persona con problemas de audición. Se pueden seleccionar a continuación unas primeras señales de voz, respectivamente señales de texto de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas. Para ello, el equipo informático 3 de la persona con problemas de audición puede estar unido a un dispositivo de detección de movimientos de la persona con problemas de audición, tal como el dispositivo de detección de movimiento descrito anteriormente. A título de ejemplo, dicho dispositivo puede detectar los movimientos de la cabeza o de cualquier otra parte del cuerpo de la persona con problemas de audición.

Se pueden combinar los dos modos de preselección descritos anteriormente. Por ejemplo, los movimientos verticales de la cabeza pueden ser detectados para seleccionar un intervalo de distancia con respecto a la persona con problemas de audición, en función del ángulo de la cabeza de la persona con problemas de audición con respecto a la horizontal. Los interlocutores situados en este intervalo de distancia pueden ser identificados en el dispositivo de visualización 5, por ejemplo visualizando un plano de la zona situada en el intervalo de distancia seleccionado que indica la posición de los interlocutores. Los movimientos horizontales de la cabeza pueden ser detectados asimismo con el fin de preseleccionar el interlocutor situado enfrente de la persona con problemas de audición en el intervalo de distancia seleccionado. La persona con problemas de audición puede modificar el interlocutor preseleccionado girando la cabeza hacia un lado o inclinándola hacia arriba o abajo. El usuario puede así preseleccionar fácilmente un interlocutor, incluso lejano, sin tener que introducir ninguna orden, sin necesitar incluso ver al interlocutor preseleccionado.

Con el fin de facilitar la selección de interlocutores, el equipo informático 3 puede proceder por otra parte a una identificación de los interlocutores presentes cerca del usuario. Para ello, el equipo informático de la persona con problemas de audición 3 puede estar unido a por lo menos una cámara de vídeo que captura una imagen de los interlocutores en el campo de visión de la persona con problemas de audición 2 y puede comprender un módulo de reconocimiento de imagen que permite identificar a dichos interlocutores en dicha imagen capturada por dicha cámara. La selección de primeras señales de voz, respectivamente de primeras señales de texto, puede comprender entonces una etapa previa de preselección por parte del equipo informático 3 de señales de voz, respectivamente de señales de texto, de interlocutores, denominados interlocutores preseleccionados, presentes en el campo de visión de la persona con problemas de audición de entre la pluralidad de señales de voz, respectivamente señales de texto, adquiridas, una etapa de identificación, por parte del módulo de reconocimiento, de dichos interlocutores preseleccionados, una etapa de visualización en el dispositivo de visualización 5 de la imagen de dichos interlocutores preseleccionados en el campo de visión de la persona con problemas de audición, asociados a su identificación, y una etapa de selección por parte de la persona con problemas de audición de primeras señales de voz, respectivamente de primeras señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, de los interlocutores preseleccionados visualizados e identificados.

Para ello, el equipo informático 3 puede estar unido a una base de datos de huellas faciales que almacena unos retratos de interlocutores conocidos asociados a su identidad. El equipo informático 3 puede utilizar las imágenes procedentes de su o sus cámaras de vídeo para intentar reconocer interlocutores registrados en la base de datos de huellas faciales aplicando un algoritmo de reconocimiento de imagen. Cuando cada interlocutor de la persona con problemas de audición 2 dispone de su propio equipo informático 3 remoto, el equipo informático 3 puede utilizar asimismo los datos de la base de datos de firmas vocales para intentar identificar al interlocutor usuario de cada equipo remoto mediante un algoritmo de reconocimiento vocal. La identidad de cada interlocutor

identificado puede ser visualizada a continuación en el dispositivo de visualización 5 para ayudar a la persona con problemas de audición a realizar su selección.

5 Según una variante, se captura una imagen de cada interlocutor presente cerca del usuario y se visualiza en el dispositivo de visualización del equipo informático del usuario, ya estén estos interlocutores presentes o no en el campo de visión del usuario. Se puede visualizar una pluralidad de viñetas, correspondientes cada una a una imagen de la cara de un interlocutor, conjuntamente en forma de mosaico. En el caso de que cada interlocutor esté equipado con su propio equipo informático, tal como un teléfono móvil, dichas imágenes pueden ser capturadas por ejemplo por los captadores de fotografías instalados en estos equipos y enviadas por estos equipos al equipo informático del usuario.

10 El equipo informático de la persona con problemas de audición 3 puede estar unido a una base de datos de conversación que almacena unos datos relativos a conversaciones anteriores seguidas por la persona con problemas de audición de manera vocal o textual. Este tipo de base de datos puede memorizar por ejemplo para estas conversaciones la identidad de los interlocutores identificados, las elecciones de selección realizadas por la persona con problemas de audición, datos de localización de la conversación, etc. A partir de estos datos, en una situación determinada en la que la persona con problemas de audición está rodeada de diversos interlocutores, el equipo informático 3 puede preseleccionar automáticamente los interlocutores seleccionados con más frecuencia en el pasado por la persona con problemas de audición. El equipo informático 3 puede iniciar asimismo automáticamente la reproducción vocal, respectivamente la visualización en modo texto, palabras de los interlocutores presentes cuyas señales de voz, respectivamente de texto, han sido seleccionadas con más frecuencia en el pasado por la persona con problemas de audición. Esto permite proponer automáticamente a la persona con problemas de audición los interlocutores más interesantes para ella, incluso si éstos no son los más cercanos. La selección de por lo menos una primera señal de voz y/o de una primera señal de texto se realiza así en función de preferencias y/o usos de la persona con problemas de audición. La actualización al final de cada conversación de la base de datos de conversación permite que el equipo informático 3 reaccione de manera personalizada mediante un aprendizaje basado en un análisis del comportamiento de la persona con problemas de audición.

30 En ciertas situaciones, la persona con problemas de audición puede estar implicada en una conversación con varios interlocutores que no pueden encontrarse todos simultáneamente en su campo de visión. Es el caso por ejemplo de la persona con problemas de audición que está conversando en la mesa con unos interlocutores situados a su derecha o su izquierda. Como las personas con problemas de audición compensan a menudo sus problemas de audición mediante informaciones visuales, tales como las expresiones faciales, es primordial para una persona con problemas de audición poder ver a sus interlocutores. El equipo informático 3 puede estar unido a por lo menos una cámara que captura una imagen de uno o varios interlocutores de la persona con problemas de audición. Tras la selección de primeras señales de voz, respectivamente de primeras señales de texto, de este o estos interlocutores, el equipo informático 3 puede visualizar una imagen de aquéllos interlocutores filmados por lo menos por una cámara y situados fuera del campo de visión de la persona con problemas de audición en una viñeta insertada en la imagen visualizada en dicho dispositivo de visualización 5. La persona con problemas de audición puede percibir así las informaciones visuales que proceden de estos interlocutores incluso si no se encuentran en su campo de visión. Con el fin de visualizar únicamente la imagen de la cara de estos interlocutores, el equipo informático 3 puede comprender un sistema de reconocimiento de imagen o de tratamiento de imagen aislando la cara de los interlocutores en las imágenes capturadas por la por lo menos una cámara. Este tipo de visualización puede ser actualizado de manera dinámica con el fin de detectar a los interlocutores que se incorporan o que salen del campo de visión de la persona con problemas de audición en función de los movimientos de éstos y de ésta. Las imágenes de los interlocutores visualizadas en el dispositivo de visualización 5 pueden ser actualizadas en función de esta detección para que la persona con problemas de audición pueda visualizar permanentemente una imagen del conjunto de estos interlocutores.

50 Algunas conversaciones pueden hacer intervenir un gran número de locutores, que se expresan en un tiempo corto, a unas distancias variables de la persona con problemas de audición 2. La selección manual en cada momento por la persona con problemas de audición de los interlocutores a seguir puede resultar entonces fastidiosa. La selección de los interlocutores se puede realizar entonces no en base a sus características individuales (identidad, posición, distancia, etc.), sino en función del orden de recepción de las señales de voz y de texto. El equipo informático 3 puede estar configurado por ejemplo para seleccionar automáticamente en cada momento un número predeterminado de interlocutores más recientes de entre el conjunto de los interlocutores. De esta manera, la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, puede comprender la determinación de un primer número predeterminado, respectivamente de un segundo número predeterminado, y la selección de dicho primer número predeterminado de primeras señales de voz, respectivamente de dicho segundo número predeterminado de primeras señales de texto, por orden de recepción. En el caso de la selección de las señales de texto, la persona con problemas de audición puede así leer en cada momento las expresiones de los N últimos interlocutores que han tomado la palabra, siendo N el segundo número descrito anteriormente.

65 Las dos últimas funciones descritas anteriormente pueden ser combinadas de manera que se visualice en el dispositivo de visualización 5 la cara de los interlocutores situados fuera del campo de visión de la persona con

problemas de audición y seleccionados de entre los interlocutores que han tomado la palabra recientemente.

El equipo informático 3 puede estar provisto de una función de grabación de las conversaciones y puede proponer a la persona con problemas de audición una visualización del historial de las conversaciones grabadas. La persona con problemas de audición puede interactuar con el equipo informático 3 para recorrer estos datos del historial con la ayuda de botones de interfaz o por medio de una pantalla táctil. En el caso de una visualización 3D o en realidad aumentada, las señales de voz y de texto adquiridas pueden estar escalonadas en el eje de la profundidad de la visualización en función de su fecha de adquisición. En el caso de un número importante de interlocutores, esta visualización, que combina una visualización según la profundidad para visualizar el historial de las expresiones y una visualización con diferentes ángulos de visualización para diferenciar a los interlocutores, asegura una mejor visibilidad de las conversaciones reproducidas.

La emisión de las señales de voz seleccionadas por la persona con problemas de audición puede estar precedida de un tratamiento o de un filtrado con el fin de suprimir los ruidos o las señales de voz no seleccionadas por la persona con problemas de audición.

Este filtrado se puede realizar con la ayuda del resultado de un reconocimiento vocal mediante un módulo de reconocimiento vocal 11 unido al equipo informático 3 o comprendido en éste. El reconocimiento vocal permite disociar a partir de una señal de audio capturada las señales de voz correspondientes a diversos interlocutores. Las expresiones de un interlocutor deseado pueden ser extraídas así a partir del resultado del reconocimiento vocal. Estas expresiones se pueden utilizar a continuación para sintetizar una señal de voz reconstituida por síntesis vocal. Esta síntesis se puede realizar con la ayuda de una firma vocal del locutor previamente grabada con el fin de reproducir una señal de voz del interlocutor más cercano. Esta señal de voz reconstituida puede ser sustraída de la señal de audio capturada con el fin de deducir de la misma la componente de la señal correspondiente a los otros interlocutores y al ruido. Una señal de voz filtrada puede ser reconstituida por el interlocutor, por ejemplo atenuando, de manera más o menos fuerte, en la señal capturada la componente de la señal capturada correspondiente a los otros interlocutores y al ruido, o también mezclando esta componente con la señal obtenida por síntesis vocal. Esto permite realizar una corrección fina de la señal, independientemente de la acústica, pudiendo conservar al mismo tiempo a voluntad una parte del ambiente sonoro del entorno.

Se puede realizar asimismo un filtrado para limitar en una señal de audio las contribuciones debidas a un interlocutor no deseado lejano. Como se ha representado en la figura 3, el equipo informático 3 puede estar unido a dos micrófonos M1 y M2, respectivamente cercanos a un interlocutor deseado I1 y a un interlocutor no deseado I2. La señal de voz correspondiente al interlocutor no deseado I2 capturada por el micrófono M1 puede ser estimada por correlación con la señal de voz correspondiente al interlocutor no deseado I2 capturada por el micrófono M2. Como alternativa, al conocer la distancia D que separa los micrófonos M1 y M2, el equipo informático 3 puede determinar en una fase de calibrado la función de transformación que une la señal de voz del interlocutor no deseado I2 adquirida por el micrófono M1 y la señal de voz de este interlocutor adquirida por el micrófono M2. La distancia D puede ser determinada por geolocalización. La función de transformación puede ser determinada de manera separada para diferentes frecuencias. Esta fase de calibrado se puede realizar cuando el interlocutor I2 es el único que habla o cuando varias personas hablan al mismo tiempo pero de manera constante. Cuando tiene lugar una adquisición posterior de la señal de voz del locutor deseado I1 a nivel del micrófono M1, la señal adquirida puede ser corregida sustrayendo la contribución de los interlocutores no deseados tales como el interlocutor I2. A la señal capturada por el micrófono M1 en un instante t se le puede sustraer la señal de voz del interlocutor no deseado I2 capturada en un instante  $(t - \Delta t)$  por el micrófono M2 multiplicada por su función de transformación, correspondiendo  $\Delta t$  al tiempo de propagación del sonido del micrófono M1 al micrófono M2. Esta función de transformación permite tener en cuenta la atenuación y el desfase de la voz del interlocutor I2 durante su propagación entre M2 y M1, así como las diferencias de ganancia entre el micrófono M1 y el micrófono M2.

Este filtrado se puede mejorar asociando a cada interlocutor un micrófono direccional y un micrófono de otro tipo tal como un micrófono omnidireccional o bidireccional. El micrófono direccional puede ser utilizado para determinar los rangos de frecuencia cubiertos por la voz de cada interlocutor. Las funciones de transformación de cada interlocutor pueden ser definidas frecuencia por frecuencia en función de los rangos de frecuencias cubiertos. La contribución de un interlocutor a las señales de voz de los otros interlocutores puede así ser suprimida en cada rango de frecuencia que solo él cubre. Esto permite una mejor robustez del calibrado así como una mejor tolerancia al ruido ambiental.

La fase de calibrado se puede realizar cuando varios interlocutores están hablando utilizando en las señales capturadas una etapa de reconocimiento vocal tal como la descrita anteriormente. El calibrado converge así sobre la marcha, a medida que se tienen en cuenta a los diferentes interlocutores. La función de transformación que une la señal de voz del interlocutor no deseado I2 adquirida por el micrófono M2 y la señal de voz de este interlocutor adquirida por el micrófono M1 puede ser determinada entonces comparando las señales de voz que corresponden al locutor no deseado I2 capturadas por el micrófono M1 y por el micrófono M2.

Se pueden realizar asimismo unos tratamientos con el fin de filtrar en las señales de audio capturadas los sonidos indirectos procedentes de una o varias reflexiones, por ejemplo en las paredes de una estancia.

Una señal de audio capturada por un micrófono puede ser analizada con el fin de diferenciar la componente de la señal procedente de la propagación directa del sonido entre el interlocutor y el micrófono y las otras componentes procedentes de reflexiones del sonido. Dicho análisis puede ser realizado de manera separada para diferentes frecuencias.

Este análisis puede buscar una subida brutal de amplitud seguida por un período con una señal de amplitud constante. La llegada de los ecos sucesivos de la señal directa produce a continuación unas variaciones sucesivas de la amplitud de la señal capturada, por etapas, que produce una señal cuya primera etapa corresponde a la señal directa a aislar.

Las características de los ecos generados en un lugar pueden ser determinadas previamente en una fase de calibrado, con el fin de sustraer estos ecos de una señal de voz capturada posteriormente. En esta fase de calibrado, el ambiente sonoro habitual de este lugar puede ser utilizado para esta determinación. Como alternativa, se puede emplear por lo menos un generador de sonido con el fin de generar unos sonidos utilizados específicamente para el calibrado. Las señales utilizadas para el calibrado, por ejemplo las señales generadas por un generador de sonidos, pueden ser de diferentes tipos, por ejemplo rectangulares o por impulsos, o también una sucesión de variaciones transitorias y de rangos de señal de amplitud constante. En el caso de una señal por impulsos, los ecos del lugar corresponden a la respuesta impulsional capturada como respuesta a la emisión de una señal por impulsos.

Otro análisis puede consistir en buscar los extremos de la derivada de una señal de voz capturada. Estos extremos pueden corresponder a unos aumentos brutales de la amplitud de la señal, cuando llega un eco. En la detección de un extremo, la variación de la señal puede ser comparada con la señal en propagación directa medida anteriormente. El análisis de este eco puede dar asimismo informaciones sobre las características de eco de la sala tales como el tiempo de retardo y la función de transformación de la señal debida al eco.

Si un eco está bien detectado y es analizado, podrá ser anulado en una conversación posterior. Otro análisis puede ser realizado con la ayuda de por lo menos dos micrófonos direccionales alejados entre sí en una distancia  $L$  y posicionados en fila frente a un interlocutor. La señal capturada por el micrófono más alejado puede ser sustraída de la señal capturada por el micrófono más cercano de manera que se minimicen los sonidos indirectos procedentes de interlocutores no deseados. Esta corrección se puede realizar separadamente para cada frecuencia. Se puede realizar un tratamiento de manera que se corrija la diferencia de fase entre las dos señales capturadas por los dos micrófonos.

Las señales de voz correspondientes a unos interlocutores no deseados capturadas por un micrófono de un equipo informático pueden ser determinadas a partir de las señales de voz de estos interlocutores capturadas por los micrófonos de los equipos informáticos de estos interlocutores mediante un cálculo de simulación acústica que estima la propagación, la atenuación, las reflexiones, etc., de estas señales en una estancia. Con el fin de realizar dicho cálculo, el equipo informático puede estar unido a una base de datos que almacena informaciones referentes a las características acústicas de esta estancia, tales como su geometría o los coeficientes de reflexión o de absorción de los materiales que componen sus diferentes muros. Dichos coeficientes pueden ser proporcionados en función de la frecuencia. Esta base de datos puede comprender asimismo informaciones referentes a los principales objetos o locutores presentes en la estancia, tales como su posición y su geometría. Las señales de voz parásitas generadas por las palabras de los usuarios no deseados a nivel del micrófono del equipo informático pueden ser determinadas así por cálculo y sustracciones de la señal de audio capturada por el equipo informático. Esta simulación puede aprovechar asimismo informaciones obtenidas a partir de pruebas de audio y/o acústicas que se pueden realizar en una sala o interlocutores dedicados. De esta manera se pueden tomar unas mediciones en una sala vacía con el fin de determinar las características acústicas de la sala. También se pueden realizar unas mediciones relativas a un interlocutor dedicado en una sala dedicada para determinar las características sonoras de la señal de voz de este interlocutor en dicha sala dedicada tales como la reflexión de la señal de voz del interlocutor en esta sala, o también las alteraciones de sonidos en la estancia, tales como reflexiones o atenuaciones, debidas a la presencia de este interlocutor en la estancia. Estas características sonoras serán específicas del interlocutor ya que dependen de la geometría del interlocutor. Estas características pueden ser añadidas a continuación al módulo de cálculo de simulación acústica con el fin de afinar más el tratamiento del ruido.

Los sonidos no deseados pueden ser anulados en la fuente por interferencia destructiva siendo capturados y reemitidos en oposición de fase. Esta anulación puede ser realizada por la persona con problemas de audición utilizando unos altavoces de manera que se anulen las voces de los interlocutores no deseados a nivel del micrófono unido al equipo informático de la persona con problemas de audición. A título de ejemplo, en el caso de una sala muy ruidosa o con mucha reverberación, la activación de dicha anulación de sonidos para todos los interlocutores permitiría restringir la transmisión del sonido de cualquier interlocutor hacia los micrófonos de los otros interlocutores. Cada micrófono captaría entonces únicamente la voz del interlocutor que lleve el micrófono y cada interlocutor equipado con un equipo informático escucharía entonces únicamente las señales de voz de los interlocutores seleccionados.

El equipo informático 3 puede estar unido a un primer micrófono que capta la totalidad del entorno sonoro, es decir el ruido ambiental complejo (las señales sonoras de los multi-locutores, los ecos, etc.) y la señal útil, y a un segundo micrófono que capta únicamente el ruido ambiental complejo. Para que el segundo micrófono capte únicamente el ruido ambiental complejo sin la señal útil, se realiza un montaje particular. Un tercer micrófono direccional que capta únicamente la señal de voz del locutor está unido a un módulo de supresión activa de esta señal de voz que reemite una señal en oposición de fase a esta señal de voz de manera que la señal útil, es decir la señal de voz del locutor, se anule a nivel del segundo micrófono. El segundo micrófono capta entonces únicamente el ruido ambiental complejo. La señal útil del locutor puede entonces ser obtenida sustrayendo de la señal relativa a la totalidad del entorno sonoro captado por el primer micrófono, la señal relativa al ruido ambiental complejo captada por el segundo micrófono. Los primer y segundo micrófonos pueden ser unos micrófonos omnidireccionales de manera que se obtenga una mejor calidad de toma de este sonido. Dichos micrófonos son efectivamente menos sensibles a las perturbaciones que un micrófono direccional aprovechando al mismo tiempo la poca sensibilidad al ruido de un micrófono direccional. Este modo de realización es independiente del tipo de ruido y permite así el tratamiento de ruidos complejos aprovechando al mismo tiempo una mejor calidad de toma de sonido que con un micrófono direccional. Como variante, el módulo de supresión activa puede ser omitido y la señal de voz del locutor captada por el tercer micrófono direccional puede ser sustraída de la señal captada por el segundo micrófono con el fin de determinar el ruido ambiental complejo.

Los ruidos emitidos de manera recurrente en el entorno de la persona con problemas de audición pueden ser detectados en unas grabaciones con el fin de suprimirlos. Para ello, los sonidos no reconocidos por reconocimiento vocal por el equipo informático 3 pueden ser analizados con el fin de determinar si se han emitido de manera repetida unos sonidos cercanos. Se puede realizar un cálculo de autocorrelación de la señal para determinar una emisión repetida de sonidos cercanos. Si es el caso, la señal grabada puede ser filtrada con el fin de hacerlos desaparecer, por ejemplo filtrando las frecuencias a las que se ha emitido este sonido.

Unos ruidos característicos, por ejemplo de una sala, pueden ser grabados asimismo en una base de datos de ruidos. Un sonido de una señal capturada no reconocido por reconocimiento vocal pero identificado como un ruido en la base de datos de ruido puede ser filtrado con el fin de corregir la señal capturada.

Se puede realizar asimismo un tratamiento con el fin de filtrar las resonancias que aparecen cuando un sonido es emitido a una frecuencia de resonancia de una sala. Las frecuencias de resonancia de las salas frecuentadas por la persona con problemas de audición pueden ser grabadas en una base de datos de resonancia. Cuando el equipo informático 3 detecta por geolocalización que la persona con problemas de audición se encuentra en una estancia en la que las frecuencias de resonancia están registradas en la base de datos de resonancia, estas frecuencias pueden ser filtradas con el fin de suprimir las resonancias de las señales de voz adquiridas.

Cuando unos equipos presentes cerca de la persona con problemas de audición pueden emitir unos ruidos molestos, se pueden posicionar unos micrófonos cerca de estos equipos de manera que detecten la emisión de tales ruidos. Cuando se detecta un ruido de este tipo, las frecuencias en las que este ruido es emitido pueden ser filtradas por el equipo informático 3 de manera que se suprima este ruido de las señales de voz adquiridas. Este tipo de sistema dedicado permite tratar los ruidos emitidos por un equipo conocido de manera irregular. Según una variante, dichos equipos pueden estar equipados con altavoces destinados a emitir un sonido en oposición de fase a estos ruidos, permanentemente o en el momento de la detección de la emisión de un ruido de este tipo.

Los sonidos reemitidos por los altavoces unidos al equipo informático 3 pueden ser emitidos de manera que recreen para la persona con problemas de audición una impresión de espacialización. Para ello, la amplitud de los sonidos reemitidos con destino a la oreja izquierda de la persona con problemas de audición puede ser diferente de la amplitud de los sonidos reemitidos con destino a la oreja derecha. La diferencia de amplitud puede ser determinada en función del ángulo entre la dirección situada frente a la persona con problemas de audición y la recta que va de la persona con problemas de audición hacia el micro que ha capturado los sonidos reemitidos. Los sonidos capturados por unos micrófonos situados a la izquierda de la persona con problemas de audición pueden ser así reemitidos con una mayor amplitud con destino a la oreja izquierda que a la oreja derecha de manera que den a la persona con problemas de audición la impresión de que el sonido reemitido procede del emplazamiento del micrófono que ha capturado el sonido. Una espacialización de este tipo se puede realizar independientemente para diferentes frecuencias. Dicha espacialización puede ser adaptada de forma dinámica de manera que siga los movimientos de la cabeza de la persona con problemas de audición.

De la misma manera, la fase de las señales reemitidas puede ser modulada de manera que reproduzca las diferencias de fase entre las señales que llegan a las dos orejas de la persona con problemas de audición. Para ello, por lo menos dos micrófonos pueden estar colocados en fila con respecto a un interlocutor de manera que determinen la diferencia de fase correspondiente a la diferencia de recorrido de un sonido entre la oreja derecha y la oreja izquierda de la persona con problemas de audición.

En el caso de la utilización de un casco para reproducir los sonidos, éste puede estar equipado con un sistema de supresión activa de ruido por oposición de fase con el fin de reducir la amplitud de los ruidos que pasan a través del casco.

La persona con problemas de audición puede así seguir eficazmente una primera conversación entre varios interlocutores incluso si éstos están dispuestos en unas direcciones opuestas, sin ser molestada por los ruidos ambientales, mientras continúa conociendo al mismo tiempo el desarrollo de por lo menos un segunda conversación que se desarrolla en segundo plano.

5

**REIVINDICACIONES**

1. Procedimiento de asistencia en el seguimiento de una conversación con una pluralidad de interlocutores (1) para una persona con problemas de audición (2) provista de un equipo informático (3),

5

estando dicho equipo informático unido a por lo menos un módulo de emisión de voz (4), a por lo menos un dispositivo de visualización (5) y a por lo menos un módulo de conversión (6) de señales de voz en señales de texto,

10

comprendiendo dicho procedimiento las etapas siguientes:

- adquisición de una pluralidad de señales de voz y de una pluralidad de señales de texto procedentes de una pluralidad de interlocutores, siendo dichas señales de texto obtenidas por conversión de dichas señales de voz por el módulo de conversión (6) de señales;
- selección de por lo menos una primera señal de voz de entre la pluralidad de señales de voz adquiridas;
- selección de por lo menos una primera señal de texto de entre la pluralidad de señales de texto adquiridas, procediendo la señal de voz y la señal de texto seleccionadas de interlocutores diferentes;
- transmisión al módulo de emisión de voz (4) de las primeras señales de voz seleccionadas para la emisión de dichas primeras señales de voz;
- orden de la visualización en el dispositivo de visualización (5) de las primeras señales de texto seleccionadas.

15

20

25

2. Procedimiento según la reivindicación anterior, estando por lo menos uno de dichos interlocutores equipado con un equipo informático de interlocutor (3) que comprende por lo menos un micrófono (7) y dicho módulo de conversión (6) de señales de voz en señales de texto, que comprende las etapas siguientes realizadas por dicho por lo menos un equipo informático de interlocutor:

30

- captura de señales de voz del interlocutor con la ayuda del micrófono (7) de dicho equipo informático de interlocutor,
- conversión de las señales de voz capturadas en señales de texto por el módulo de conversión de dicho equipo informático de interlocutor (6),
- transmisión a dicho equipo informático de la persona con problemas de audición de las señales de voz capturadas y de las señales de texto obtenidas por conversión.

35

40

3. Procedimiento según la reivindicación anterior, en el que, comprendiendo el equipo informático de la persona con problemas de audición (3) un módulo de recepción de señales (10) y siendo susceptible de ser conectado a una red de comunicación de voz y a una red de comunicación de datos, dicha etapa de adquisición de señales de interlocutores comprende la recepción de las señales de voz y de las señales de texto de la pluralidad de interlocutores por el módulo de recepción de señales (10) por medio de las redes de comunicaciones de voz y de datos.

45

4. Procedimiento según la reivindicación 1, en el que, comprendiendo el equipo informático de la persona con problemas de audición (3) por lo menos un micrófono (7) y dicho módulo de conversión (6) de señales de voz en señales de texto, dicha etapa de adquisición de señales de interlocutores comprende las etapas siguientes:

50

- adquisición de las señales de voz de la pluralidad de interlocutores con la ayuda del micrófono (7) de dicho equipo informático de la persona con problemas de audición,
- conversión de las señales de voz adquiridas en señales de texto por dicho módulo de conversión (6).

55

5. Procedimiento según una de las reivindicaciones anteriores, en el que, estando el equipo informático de la persona con problemas de audición (3) unido a una base de datos de firmas vocales de interlocutores conocidos, las selecciones de las primeras señales de voz y de texto se realizan con la ayuda de un reconocimiento de las señales de voz de los interlocutores conocidos de entre las señales de voz adquiridas a partir de dicha base de datos.

60

6. Procedimiento según una de las reivindicaciones anteriores, en el que la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende:

65

- una etapa previa de preselección por parte del equipo informático de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de entre la pluralidad de señales de

voz, respectivamente de señales de texto, adquiridas en función de la distancia entre el interlocutor emisor de dicha señal de voz, respectivamente señal de texto, y la persona con problemas de audición,

- 5 - una etapa de selección de primeras señales de voz, respectivamente señales de texto, de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas.

7. Procedimiento según una de las reivindicaciones anteriores, en el que, estando el equipo informático (3) de la persona con problemas de audición unido a un dispositivo de detección de movimientos de la persona con problemas de audición (2), la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende:

- 10 - una etapa previa de preselección por parte del equipo informático (3) de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas en función de un movimiento de la persona con problemas de audición detectado por el dispositivo de detección de movimientos,
- 15 - una etapa de selección de primeras señales de voz, respectivamente señales de texto, de entre la pluralidad de señales de voz, respectivamente señales de texto, preseleccionadas.

8. Procedimiento según una de las reivindicaciones anteriores, en el que, estando el equipo informático (3) de la persona con problemas de audición unido a por lo menos una cámara que captura una imagen de los interlocutores en el campo de visión de la persona con problemas de audición (2) y que comprende un módulo de reconocimiento de imagen que permite identificar a dichos interlocutores en dicha imagen capturada por dicha cámara, la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende:

- 20 - una etapa previa de preselección por parte del equipo informático (3) de la persona con problemas de audición de señales de voz, respectivamente de señales de texto, de interlocutores, denominados interlocutores preseleccionados, presentes en el campo de visión de la persona con problemas de audición de entre la pluralidad de señales de voz, respectivamente de señales de texto, adquiridas,
- 25 - una etapa de identificación, por medio del módulo de reconocimiento, de dichos interlocutores preseleccionados,
- 30 - una etapa de visualización en el dispositivo de visualización (5) de la imagen de dichos interlocutores preseleccionados en el campo de visión de la persona con problemas de audición, asociados a su identificación,
- 35 - una etapa de selección por parte de la persona con problemas de audición de primeras señales de voz, respectivamente de primeras señales de texto, de entre la pluralidad de señales de voz, respectivamente de señales de texto, de los interlocutores preseleccionados visualizados e identificados.
- 40

9. Procedimiento según una de las reivindicaciones anteriores, que comprende:

- 45 - una etapa de recepción de una orden vocal de la persona con problemas de audición (2) relativa a un interlocutor y a un tipo de señales a seleccionar;
- 50 - una etapa de selección de primeras señales de voz o de texto del interlocutor en función de dicha orden vocal, comprendiendo el equipo informático de la persona con problemas de audición (3) un módulo de reconocimiento vocal (11) configurado para analizar dicha orden vocal.

10. Procedimiento según una de las reivindicaciones anteriores, en el que la selección de por lo menos una primera señal de voz y/o de una primera señal de texto se realiza en función de preferencias y/o usos de la persona con problemas de audición (2).

11. Procedimiento según una de las reivindicaciones anteriores, en el que, estando el equipo informático (3) de la persona con problemas de audición unido a por lo menos una cámara que captura una imagen de por lo menos un interlocutor (1) de la persona con problemas de audición (2), tras la selección de primeras señales de voz, respectivamente de primeras señales de texto, de dicho interlocutor, dicho procedimiento comprende una etapa de visualización de la imagen de dicho interlocutor situado fuera del campo de visión de la persona con problemas de audición (2) en una viñeta insertada en la imagen visualizada en dicho dispositivo de visualización (5).

12. Procedimiento según una de las reivindicaciones anteriores, en el que la etapa de selección de primeras señales de voz, respectivamente de primeras señales de texto, comprende la determinación de un primer número predeterminado, respectivamente de un segundo número predeterminado, y la selección de dicho primer número predeterminado de primeras señales de voz, respectivamente de dicho segundo número predeterminado



de primeras señales de texto, por orden de recepción.

5 13. Procedimiento según una de las reivindicaciones anteriores, en el que la etapa de emisión de las primeras señales de voz seleccionadas comprende previamente una etapa de filtrado del ruido y/o de las señales de voz no seleccionadas.

10 14. Procedimiento según la reivindicación anterior, en el que, comprendiendo el equipo informático (3) de la persona con problemas de audición un módulo de reconocimiento vocal, la etapa de filtrado comprende una etapa de filtrado de las señales de voz seleccionadas con la ayuda del resultado de un reconocimiento vocal por parte del módulo de reconocimiento vocal (11).

15 15. Producto de programa de ordenador que comprende unas instrucciones de código para la realización de un procedimiento según cualquiera de las reivindicaciones anteriores cuando este programa es ejecutado por un procesador.

20 16. Equipo informático (3), para la asistencia en el seguimiento de una conversación con una pluralidad de interlocutores (1), en posesión de una persona con problemas de audición (2) y unido a por lo menos un módulo de emisión de voz (4), a por lo menos un dispositivo de visualización (5) y a por lo menos un módulo de conversión de señales de voz en señales de texto (6),

estando dicho equipo (3) configurado para:

25 - adquirir una pluralidad de señales de voz y una pluralidad de señales de texto procedentes de una pluralidad de interlocutores, siendo dichas señales de texto obtenidas por conversión de dichas señales de voz por el módulo de conversión de señales (6);

- seleccionar por lo menos una primera señal de voz de entre la pluralidad de señales de voz adquiridas;

30 - seleccionar por lo menos una primera señal de texto de entre la pluralidad de señales de texto adquiridas, procediendo la señal de voz y la señal de texto de interlocutores diferentes;

- transmitir al módulo de emisión de voz (4) unas primeras señales de voz seleccionadas para la emisión de dichas primeras señales de voz;

35 - ordenar la visualización en el dispositivo de visualización (5) de las primeras señales de texto seleccionadas.

17. Gafas de realidad aumentada que comprenden un equipo informático (3) según la reivindicación 16.

FIG. 1

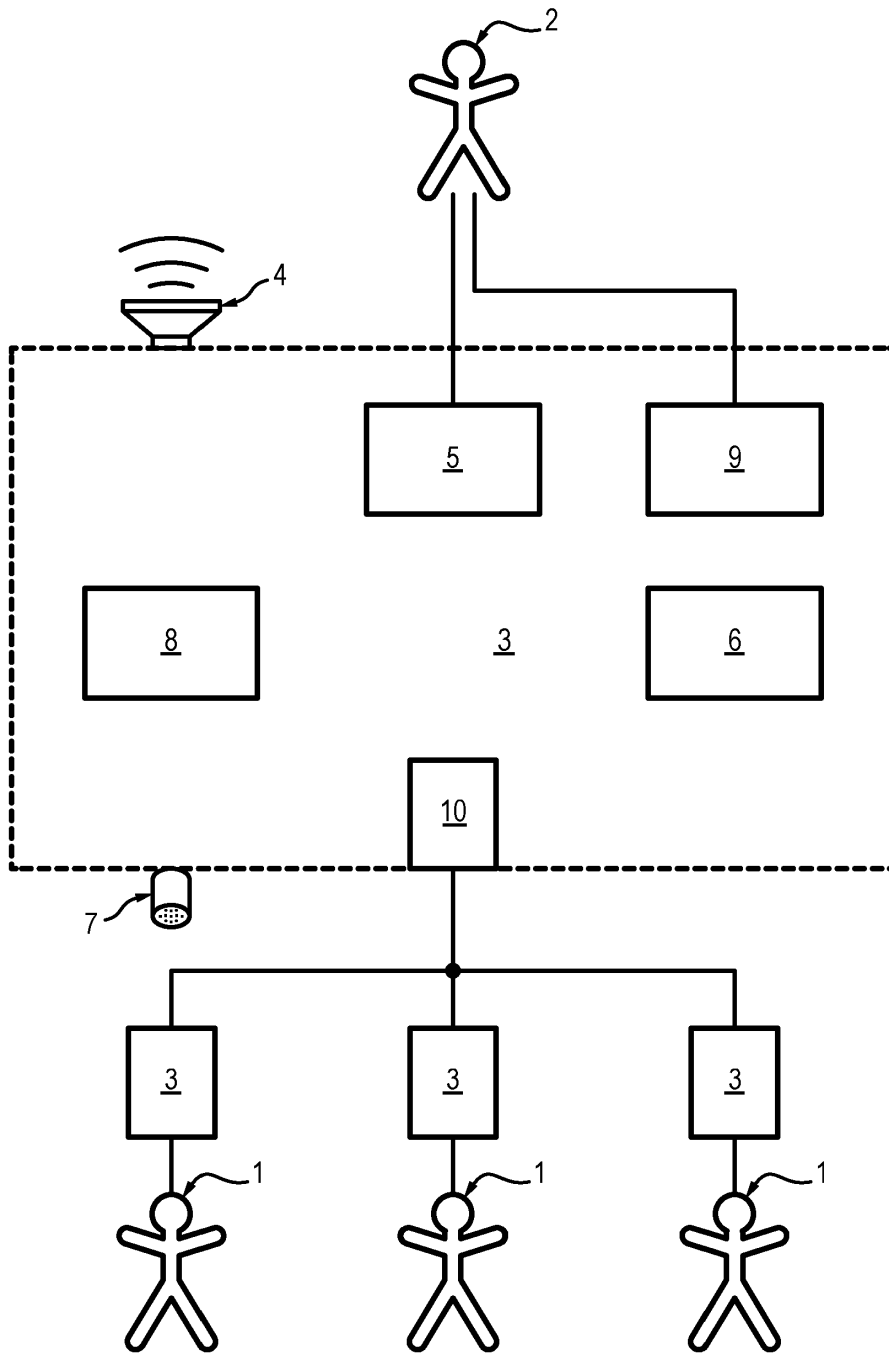


FIG. 2

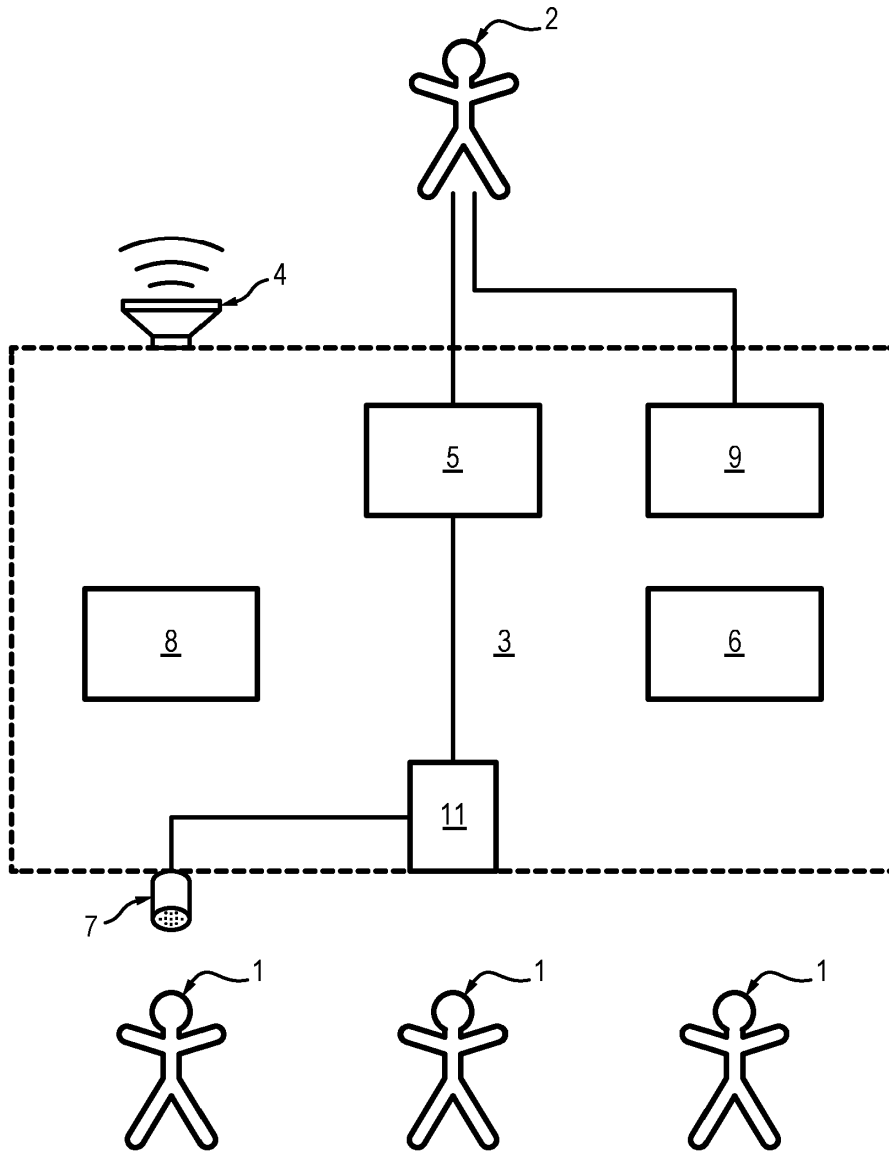


FIG. 3

