

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 983 529**

51 Int. Cl.:

G10L 25/66	(2013.01) G10L 15/02	(2006.01)
G10L 25/63	(2013.01)	
A61B 5/00	(2006.01)	
A61B 5/16	(2006.01)	
G16H 10/20	(2008.01)	
G16H 10/60	(2008.01)	
G16H 15/00	(2008.01)	
G16H 40/63	(2008.01)	
G16H 50/20	(2008.01)	
G16H 50/30	(2008.01)	

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **24.10.2018** **PCT/GB2018/053062**
- 87 Fecha y número de publicación internacional: **02.05.2019** **WO19081915**
- 96 Fecha de presentación y número de la solicitud europea: **24.10.2018** **E 18795793 (1)**
- 97 Fecha y número de publicación de la concesión europea: **31.07.2024** **EP 3701529**

54 Título: **Sistema y método para evaluar el estado fisiológico**

30 Prioridad:

24.10.2017 GB 201717469

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

23.10.2024

73 Titular/es:

CAMBRIDGE COGNITION LIMITED (100.0%)
Tunbridge CourtTunbridge LaneBottisham
Cambridge, Cambridgeshire CB25 9TU, GB

72 Inventor/es:

BARNETT, JENNIFER HELEN;
CORMACK, FRANCESCA KATHLEEN;
TAPTIKLIS, NICHOLAS THEODORE y
SU, MERINA TONG

74 Agente/Representante:

PONS ARIÑO, Ángel

ES 2 983 529 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

Sistema y método para evaluar el estado fisiológico

- 5 Esta invención se refiere a sistemas y métodos para evaluar el estado fisiológico de un sujeto. La invención se refiere en particular, pero no exclusivamente, a métodos que utilizan resultados de biomarcadores de voz y evaluaciones cognitivas y/o clínicas para evaluar mejor el estado fisiológico, tal como el dolor, incluidos síntomas de afecciones neurológicas o neuropsiquiátricas.
- 10 La toma de decisiones clínicas, incluido el diagnóstico de enfermedades neurológicas y neuropsiquiátricas y la prescripción de medicamentos, se basa en pruebas precisas de función cognitiva y clasificación de síntomas. Además, las pruebas cognitivas pueden desempeñar un papel importante a la hora de abordar la salud cerebral general en personas que no están afectadas por enfermedades neurológicas. Por otra parte, las pruebas cognitivas pueden formar parte de una evaluación general de la aptitud física con implicaciones en la política de salud y seguridad. Por
- 15 ejemplo, las pruebas cognitivas proporcionan:
- información de referencia para que un sujeto evalúe el riesgo de una afección o como parte de un estudio que proporcione datos normativos para una población en particular;
 - un medio para detectar señales tempranas de la aparición de trastornos neurológicos o sus precursores;
 - una ayuda en el diagnóstico preciso de trastornos neurológicos;
- 20 • un medio para controlar la evolución de los trastornos neurológicos; y
- un método para determinar el impacto que está teniendo la evolución del tratamiento en el paciente en términos tanto de seguridad cognitiva como de eficacia.
- Las pruebas cognitivas y la evaluación clínica suelen consistir en una serie de distintas pruebas realizadas por el
- 25 paciente con la supervisión de un médico. Los procedimientos empleados durante las pruebas cognitivas suelen consistir en un conjunto estructurado de problemas o tareas estandarizados que se centran específicamente en uno o más procesos cognitivos del cerebro. Las instrucciones y 'reglas' de la prueba podrán darse por escrito, verbalmente o a través de un sistema automatizado, tal como un ordenador. Por ejemplo, se puede pedir a un paciente que aprenda una lista de palabras y recuerde tantas palabras como sea posible después de un período de tiempo determinado (o
- 30 después de realizar una tarea diferente). Esta tarea en particular pone a prueba la capacidad de memoria del paciente y es sensible a los síntomas de la demencia. Las puntuaciones calculadas en función de las respuestas del paciente pueden ser realizadas por un humano capacitado o por un ordenador, en tiempo real o basándose en un registro que puede ser escrito, oral, manual o digital. Las pruebas se califican en función de datos de pruebas anteriores del paciente o de conjuntos de datos históricos de poblaciones de interés con función cognitiva normal y aquellas con
- 35 trastornos neurológicos conocidos. La población de interés podrá seleccionarse en función de la edad, género y cuadros clínicos conocidos.
- Como ejemplos de pruebas de función cognitiva actuales para la demencia y otros trastornos neurológicos se incluyen el mini examen del estado mental, la prueba mental abreviada, la evaluación cognitiva por el médico de familia y la
- 40 prueba de aprendizaje verbal de Hopkins.
- La realización de una evaluación cognitiva completa suele consistir en varios conjuntos de interacciones entre el sujeto de la prueba y la persona que la realiza, en las que, en función del rendimiento obtenido hasta el momento, se dan instrucciones o indicaciones adicionales y se proponen otras tareas o problemas hasta que se alcanza un criterio de
- 45 rendimiento determinado. Los sistemas de pruebas cognitivas computarizados, como la batería automatizada de pruebas neuropsicológicas de Cambridge (CANTAB), automatizan todos los aspectos del procedimiento de pruebas cognitivas, incluido el orden de las tareas, todos los aspectos de la presentación de tareas, instrucciones interactivas y puntuación de las tareas.
- 50 No obstante, el estado cognitivo y fisiológico no puede predecirse perfectamente mediante puntuaciones cognitivas y/o información clínica o demográfica debido a varias limitaciones del estado actual de la técnica. El estado cognitivo se define como la capacidad del sujeto para realizar funciones cognitivas específicas del cerebro, como la memoria, atención, función ejecutiva, idioma. El estado fisiológico se define como las sensaciones de dolor, cansancio, sedación y estado de alerta del sujeto.
- 55 Una de las limitaciones de las pruebas cognitivas es la necesidad de que sean fáciles de usar y tolerables para garantizar su cumplimiento al tiempo que se superan los límites de rendimiento. Sin pruebas más allá del límite en el que el usuario puede responder fácilmente, las pruebas cognitivas no tienen forma de medir el límite superior de la capacidad cognitiva de una persona. Esta sensación puede resultar frustrante para los usuarios y hacer que las
- 60 pruebas requieran mucho tiempo, lo que resulta logísticamente problemático y costoso en los ensayos clínicos y la práctica clínica.
- Otra limitación es que para algunas pruebas cognitivas, se puede lograr el mismo rendimiento utilizando diferentes estrategias, que emplean diferentes circuitos o funciones cerebrales. Por ejemplo, en una prueba de memoria, dos
- 65 personas pueden utilizar diferentes estrategias para obtener la misma puntuación: una utilizando una estrategia aprendida, tal como una nemotécnica, y otra sin estrategia. Aunque ambas personas pueden obtener la misma

puntuación, se esperaría que la carga de los circuitos neuronales relacionados con la memoria fuera mayor en la segunda persona. Como alternativa, las dos personas pueden dar puntuaciones diferentes para el mismo nivel de capacidad de memoria debido a la diferencia en el uso de estrategias.

Otras limitaciones incluyen la experiencia subjetiva y la variación cultural, tal como "fingir" un mal rendimiento y diferencias de motivación que conducen a diferencias en el rendimiento. Por ejemplo, un paciente que siente niveles leves de dolor puede exagerar los informes de dolor a propósito para recibir fármacos opiáceos. Por lo tanto, existe la necesidad en la técnica de un sistema fácil de usar que permita medir con precisión las capacidades cognitivas y el estado fisiológico de una persona utilizando marcadores objetivos en el flujo de respuesta.

Una forma objetiva de medir el estado fisiológico es extraer características de la voz y el habla. La voz humana contiene información importante sobre nuestro procesamiento neuronal a través de lo que decimos y cómo lo decimos. En los seres humanos adultos, expresamos y evaluamos estados fisiológicos subjetivos como el dolor, la sedación, el cansancio y el estado de ánimo a través tanto del contenido del habla como de otras características del habla, incluyendo la velocidad de producción de palabras, la fluidez del habla y el tono de voz. En los niños preverbales y en los animales, estos estados pueden deducirse de vocalizaciones sin palabras.

En investigaciones anteriores (Lautenbacher *et al.*, 2017) se descubrió que los cambios en las características acústicas de la producción de vocales predecían significativamente cambios en la percepción subjetiva del dolor. Cincuenta adultos jóvenes sanos produjeron las vocales 'u', 'a', 'i' y 'schwa' (vocal neutra cuyo símbolo es ə) (a en las palabras 'alone' y u en 'circus' (en inglés)) al tiempo que sumergen sus manos en agua caliente y por debajo del valor inicial (sin inmersión en calor). Los parámetros fonéticos extraídos fueron el paso (f0 media), el rango f0 y el volumen. Se observó que el paso y el volumen de las vocales 'u' y 'schwa' (ə) aumentaban durante el dolor, y que un mayor aumento en estos parámetros fonéticos se asociaba a un mayor aumento en de las calificaciones subjetivas de la escala de dolor.

En otro estudio (Oshrat 2014) se ha demostrado que, en principio, los algoritmos de clasificación basados en aprendizaje automático pueden diferenciar entre muestras de habla tomadas de personas con o sin dolor significativo. En un pequeño estudio israelí, se utilizaron 97 grabaciones tomadas de un total de 27 adultos con lesiones traumáticas que producían dolor para generar entre 3 y 6 fragmentos de voz de un segundo de duración de cada grabación, en las que se pronunciaban cifras de su número de identificación o palabras de su nombre. Usando el aprendizaje automático, los autores pudieron seleccionar un conjunto de características que clasificaban correctamente las clasificaciones de dolor/ausencia de dolor en aproximadamente el 80 % de las muestras de hombres y el 83 % de las mujeres.

Mientras que Oshrat y colaboradores (2014) no pudieron utilizar el método para desarrollar un marcador de dolor que copiase los diversos niveles de una escala clínica (p. ej., una escala del 1 al 10, a diferencia de un criterio binario de dolor/ausencia de dolor) debido al pequeño tamaño de la muestra, Tsai et al (2016) pudieron clasificar la intensidad del dolor en clases binaria (dolor/sin dolor) y ternaria (dolor leve/moderado/grave), con una precisión del 72,3 % y 51,6 % respectivamente. Este estudio empleó una máquina de soporte vectorial que utilizó tanto características acústicas extraídas de grabaciones de voz como características faciales extraídas de grabaciones de video.

En conjunto, estos estudios proporcionan un estudio demostrativo preliminar de que el análisis basado en el aprendizaje automático de muestras de voz muy breves puede discriminar características de la voz relacionadas con el dolor.

De manera similar, se han utilizado modelos basados en inteligencia artificial (IA) para deducir el estado fisiológico a partir de la voz sola, así como junto con otras variables (p. ej., grabaciones de video, variables demográficas, variables específicas de la enfermedad y rendimiento cognitivo) en otras áreas de la neurología y la psiquiatría, incluida la depresión (Williamson *et al.*, 2014), demencia frontotemporal (Nevler *et al.*, 2017), trastorno del espectro autista (Fusaroli *et al.*, 2017), enfermedad de Parkinson (Zhang *et al.*, 2016), y el trastorno de estrés postraumático y la depresión (Place *et al.*, 2017).

Por lo tanto, es conocido en la técnica que ciertos estados fisiológicos pueden deducirse posiblemente del habla. Uno de los problemas que hay que superar en el uso del habla como marcador del estado fisiológico es la variabilidad intrínseca al habla humana.

La variación en la señal del habla en humanos puede deberse a diversos factores:

- A) Variación entre hablantes debida a variaciones biológicas como el sexo, la edad, características relativas al tamaño y la forma de la anatomía de la voz, trastornos de la voz, tabaquismo
- B) Variación debida a conductas aprendidas (nivel educativo, idioma, acento regional, estilo de hablar)
- C) Las grabaciones sonoras de voces también pueden reflejar variaciones en el entorno de grabación (ruido de fondo, calidad de grabación de datos y micrófono)
- D) Variación dentro de los hablantes cuando se les somete a pruebas en distintas ocasiones debido a factores distintos del estado cognitivo o fisiológico, p.ej., hidratación, humedad, carga vocal

También hay dos fuentes de variación que se relacionan con el contexto psicológico de la tarea:

E) Variación dentro de los hablantes cuando se les somete a pruebas en distintas ocasiones debido a estados cognitivos o fisiológicos actuales, p. ej., delirio, demencia, depresión, ansiedad, dolor, cansancio (p. ej., Johnstone 2001 UWA; Vogel *et al.*, 2010)

F) Variación dentro de los hablantes cuando se les somete a pruebas en la misma ocasión pero en diferentes condiciones de tarea, p. ej., carga cognitiva alta frente a baja; multitarea (p. ej., contar mientras se controla simultáneamente la postura corporal (Andersson *et al.*, 2002), estrés relacionado con la presencia simultánea de múltiples factores estresantes, como el ruido ambiental y la carga cognitiva (Marquard *et al.*, 2017), carga emocional alta frente a baja (Johnstone 2001, UWA).

Estas fuentes de variación son particularmente desafiantes en pacientes con afecciones psiquiátricas y neurológicas.

Muchas afecciones cerebrales provocan cambios en la voz y el habla (A). Disartria es el término médico para referirse a la dificultad para hablar, causada por un trastorno cerebral congénito o adquirido o por la medicación. Puede incluir una serie de síntomas como hablar arrastrando las palabras, habla nasal o entrecortada, voz tensa y ronca, hablar excesivamente alto o bajo, problemas para hablar con un ritmo regular, con frecuentes vacilaciones, habla monótona o con sonido "gutural" y dificultad con los movimientos de la lengua y los labios

Los pacientes con cambios cerebrales temporales (p. ej., inducidos por fármacos) o permanentes (p. ej., neurodegeneración) tienen más probabilidades de tener una menor tolerancia a la carga cognitiva o emocional, u otro factor estresante, y una menor capacidad para realizar varias tareas a la vez (F).

Los pacientes con trastornos cerebrales también tienen más probabilidad de padecer síntomas comórbidos tales como depresión, dolor o cansancio (E).

Un sistema mejorado para controlar la función cerebral a través de muestras de voz debería tener en cuenta, limitar o controlar la variabilidad intrínseca en la voz individual característica de cada persona.

WO 2013/138633 AI, WO 2016/028495 AI, US 9 782 122 BI, US 2016/125748 AI, US 6 157 913 A, US 2012/116186 AI y US 2015/064669 AI, se desvela materia relevante a los antecedentes tecnológicos de la invención. La parte caracterizadora de la reivindicación 1 se basa en el documento US 2015/064669 AI.

La presente invención pretende abordar al menos parcialmente algunos de los problemas anteriores.

La invención se define en las reivindicaciones adjuntas.

La función cerebral puede deducirse del habla, pero la variación entre individuos puede suponer un desafío. Se producen otras variaciones entre las muestras de habla obtenidas del mismo individuo en diferentes momentos en función del contexto individual en el momento en que se obtienen las muestras. Hablar en condiciones que requieren un procesamiento cerebral adicional (p. ej., mientras se realiza una actividad física o comportamientos complejos como conducir) o factores estresantes (p. ej., hablar en público, cuando se está cansado, estresado, con dolor, etc.) también dejan una señal en la voz. Es probable que estas señales sean distintivas entre estados porque las funciones cognitivas individuales se ven afectadas de manera diferente por varios estados.

La invención aborda los problemas de la técnica comparando las características de la voz de un individuo en dos condiciones en la misma ocasión para anular la variación de características debida a diferencias entre individuos (A y B anteriores) y a aspectos de las ocasiones o el entorno de prueba que no están relacionados con el estado mental o fisiológico (C y D anteriores). Esto deja variaciones en las características de la voz debido a las condiciones de la tarea (F) y al estado fisiológico actual (E). Además, haciendo participar activamente a los individuos en tareas cognitivas desafiantes y moderando el nivel de dificultad de la tarea en función del desempeño, la invención minimiza el "fingimiento" del rendimiento y los síntomas de la enfermedad, maximizando así el desenmascaramiento de la "verdadera" capacidad cognitiva y el estado fisiológico. Las condiciones de la tarea (F) y el estado fisiológico actual (E) exacerban algunos aspectos de la señal de voz y minimizan otros.

Por lo tanto, en una realización de la invención, las condiciones de la tarea (F) se modificarán sistemáticamente, por ejemplo, aumentando y disminuyendo la carga cognitiva necesaria para completar la tarea en cuestión. Esto producirá señales de rendimiento y de voz relacionadas con cada carga cognitiva. La diferencia en el rendimiento de la tarea y las características de la voz registradas durante el rendimiento entre cargas cognitivas en la misma tarea, pueden representarse como puntuaciones delta para cada participante individual. Esta puntuación delta por sí sola puede utilizarse como marcador genérico del esfuerzo cerebral en cada persona individual. Adicionalmente, la puntuación delta puede compararse después con la delta obtenida en condiciones de carga alta frente a baja en pacientes con estados fisiológicos conocidos en un conjunto de entrenamiento, y una probabilidad de ese estado producida por un modelo de IA. Por lo tanto, esta señal puede añadirse a las puntuaciones de rendimiento cognitivo y a otras variables independientes conocidas del estado fisiológico (p. ej., puntuaciones clínicas, demografía) para mejorar un modelo predictivo del estado fisiológico y del trastorno neurológico o neuropsiquiátrico.

La invención se describirá con mayor detalle a continuación a modo de ejemplos no limitativos, con referencia a los

dibujos adjuntos, en los que:

La fig. 1 muestra un sistema de ejemplo de la invención;

5 La fig. 2 muestra un ejemplo de parte de un sistema de interfaz de usuario (*front-end*);

La fig. 3 muestra un ejemplo de parte de un sistema de interfaz de usuario (*front-end*);

La fig. 4 muestra un ejemplo de sistema de interfaz de administración (*back-end*);

10

La fig. 5 muestra un ejemplo de entrenamiento del sistema de IA.

La fig. 1 muestra una realización de un sistema para evaluar el estado fisiológico de un sujeto. El sistema comprende un módulo de suministro de tareas 1 configurado para comunicar a un sujeto al menos dos conjuntos de información, relacionándose cada conjunto de información con una tarea cognitiva que requiere una respuesta hablada del sujeto, y un módulo de detección de respuesta 2 configurado para registrar las respectivas respuestas habladas del sujeto como una señal de audio, comprendiendo el módulo de detección de respuesta 2 un micrófono 21. El módulo de suministro de tareas 1 y el módulo de detección de respuesta 2 forman en conjunto una 'interfaz de usuario' del sistema. La interfaz de usuario (módulo de detección de respuesta 2) puede transferir los datos de audio grabados a un módulo de análisis de 'interfaz de administración 3'. El módulo de análisis 3 puede recibir datos de señales de voz en tiempo real durante una tarea.

En la realización que se muestra en la figura 1, el sistema de la invención puede proporcionarse con un ordenador, teléfono, dispositivo portátil u otro dispositivo electrónico (es decir, un dispositivo informático móvil). El dispositivo puede suministrar interactivamente instrucciones de tareas cognitivas o clínicas (verbalmente a través de un altavoz 11 y/o a través de otros medios, por ejemplo visualmente en una pantalla) y registrar las respuestas del sujeto (verbales a través de un micrófono 21 y opcionalmente también respuestas visuales, gestuales o manuales a través de una cámara o una interfaz manual de usuario). A partir de estas respuestas, se calculan puntuaciones cognitivas y clínicas.

En otra realización, el sistema puede proporcionarse con un dispositivo informático móvil y un ordenador remoto (por ejemplo, un servidor). El dispositivo informático móvil puede incluir el módulo de suministro de tareas 1 y el módulo de detección de respuesta 2. El dispositivo informático remoto puede incluir el módulo de análisis 3. Los datos se comunican entre el dispositivo informático móvil y el ordenador remoto.

El módulo de suministro de tareas 1 puede configurarse para suministrar una batería de tareas cognitivas. Cada tarea incluirá instrucciones (por ejemplo, verbales), listas de palabras, números, asociaciones, sonidos, preguntas y otra información, así como secuencias de indicaciones, diseñadas para suscitar una respuesta verbal del usuario. El módulo de detección de respuesta 2 está configurado para realizar funciones de manejo de respuesta y puede usar sistemas de reconocimiento automático del habla (ASR, *Automatic Speech Recognition*) y de comprensión del lenguaje natural (NLU, *Natural Language Understanding*) para interpretar aspectos semánticos de la respuesta de voz.

Las tareas cognitivas pueden incluir parámetros que pueden manipularse para aumentar la dificultad (es decir, aumentar a carga cognitiva) de la tarea con respecto a diferentes aspectos de la cognición. Por ejemplo, una tarea diseñada para medir la memoria de trabajo puede parametrizarse de modo que aumente el número de elementos a recordar, o aumente la complejidad de la manipulación de los elementos recordados exigidos, o se reduzca el grado de similitud entre los elementos.

En consecuencia, los conjuntos de información comunicados pueden seleccionarse de diferentes grupos de conjuntos de información prealmacenados, estando dichos conjuntos de información prealmacenados agrupados de acuerdo con una carga cognitiva asociada a la tarea con la que se relaciona cada conjunto de información. La información relativa a tareas de diferentes dificultades puede prealmacenarse, p. ej., en una memoria asociada al módulo de suministro de tareas 1.

Las tareas suministradas por el módulo de suministro de tareas 1 pueden no diferir necesariamente en la carga cognitiva (p. ej., las mencionadas explícitamente en líneas anteriores) pero, alternativa o adicionalmente, pueden diferir en el estado mental o fisiológico inducido por la tarea. Por ejemplo, se puede indicar al sujeto que recite una lista de cuatro números mientras escucha una primera pieza musical (primera tarea) y después recite una lista de otros cuatro números mientras escucha una segunda pieza musical (segunda tarea). Aunque se puede considerar que los recitales tienen la misma carga cognitiva asociada, las diferentes condiciones (es decir, la música) pueden inducir diferentes estados mentales y fisiológicos en el sujeto, lo que puede afectar al rendimiento de los recitales.

Como ejemplos de condiciones que pueden inducir un estado físico o mental particular se incluyen: la visualización de diferentes imágenes, escuchar diferentes sonidos o música, realizar diferentes actividades físicas, evocar recuerdos emocionales, ingerir medicamentos o complementos nutritivos, centrar la atención en enfermedades o síntomas preexistentes, inducir imágenes mentales, estimulación verbal, manipular la retroalimentación sobre el rendimiento al usuario. Algunas condiciones (por ejemplo, imágenes, sonidos, música, imágenes mentales, memoria, estimulación

verbal, manipulación de retroalimentación) pueden inducir un estado mental particular, por ejemplo, si las condiciones seleccionadas son particularmente angustiosas, calmantes, agradables o desagradables, por ejemplo. Algunas condiciones pueden (por ejemplo, actividades físicas, medicamentos, complementos) inducir un estado físico o mental particular si las condiciones se seleccionan para inducir cansancio, sedación, estado de alerta mental o causar dolor al sujeto, por ejemplo.

En consecuencia, los conjuntos de información comunicados pueden, alternativa o adicionalmente, seleccionarse entre diferentes grupos de conjuntos de información prealmacenados, estando dichos conjuntos de información prealmacenados agrupados de acuerdo con un estado físico o mental inducido por la tarea con la que se relaciona cada conjunto de información. Las tareas pueden tener las mismas o diferentes cargas cognitivas asociadas. La información relativa a tareas de diferentes dificultades puede prealmacenarse, p. ej., en una memoria asociada al módulo de suministro de tareas 1.

El módulo de suministro de tareas 1 también puede incluir lógica interna que modifique de forma adaptativa el flujo y la dificultad de las tareas en respuesta tanto al contenido semántico de las respuestas del sujeto como, opcionalmente, a la entrada en tiempo real procedente del análisis de voz. La lógica interna de la tarea puede adaptar continuamente la dificultad de los ensayos durante el flujo de la tarea hasta que el sujeto cometa una determinada cantidad de errores, o hasta que el sistema detecte un cambio específico en la señal de voz, como un indicador del nivel de estrés requerido. El módulo de suministro de tareas 1 también puede incluir lógica interna configurada para utilizar el rendimiento semántico y/o la información de características de voz de una tarea anterior para adaptar el flujo y los parámetros de una tarea posterior.

En consecuencia, el módulo de detección de respuesta 2 puede comprender un módulo de reconocimiento del habla 20 configurado para analizar cada respuesta hablada, el módulo de detección de respuesta 2 puede configurarse para comparar el resultado desde el módulo de detección del habla con una respuesta esperada prealmacenada, y asignar una puntuación a cada respuesta hablada basándose en dicha comparación. El módulo de suministro de tarea 1 puede configurarse para seleccionar un conjunto de información relacionada con una siguiente tarea en función de la puntuación asociada a una respuesta a una tarea anterior. Si la puntuación de la tarea anterior es inferior a una puntuación umbral predeterminada, la siguiente tarea puede seleccionarse de manera que tenga una carga cognitiva asociada inferior a la carga cognitiva asociada a la tarea anterior. Si la puntuación de la tarea anterior es superior a una puntuación umbral predeterminada, la siguiente tarea puede seleccionarse de manera que tenga una carga cognitiva asociada superior a la carga cognitiva asociada a la tarea anterior. Si se determina que la puntuación de la tarea anterior es un valor atípico, la siguiente tarea puede seleccionarse de manera que tenga una carga cognitiva asociada igual o similar a la carga cognitiva asociada a la tarea anterior.

Como alternativa, o además, el sistema puede determinar la carga cognitiva de una tarea en función de la respuesta detectada por el módulo de detección de respuesta 2. Por ejemplo, el módulo de análisis 3 puede determinar la carga cognitiva de una tarea basándose en las características de la respuesta hablada. Las características pueden incluir las mismas características que las utilizadas para determinar el estado fisiológico del sujeto, p. ej., tono, intensidad, frecuencias formánticas, flujo glotal, duración del habla, velocidad del habla y calidad de la voz. Sin embargo, estas pueden parametrizarse de forma diferente en cada caso.

En un estudio dirigido a esta característica, sesenta participantes de entre 21 y 78 años completaron una prueba verbal automatizada de memoria de trabajo. La amplitud de la memoria de trabajo varió entre 3 y 8 elementos. Las respuestas se clasificaron como "carga alta" si eran $> 0,6$ de la amplitud máxima de ese participante. Las características de audio extraídas de cada respuesta dada por cada uno de los participantes se normalizaron, expresando diferencias dentro de los sujetos en cuanto a las características vocales en todos los ensayos de carga variable. Los datos se dividieron en conjuntos de datos de entrenamiento (70 %) y de prueba (30 %), y se analizaron utilizando un clasificador de soporte vectorial (SVC, *Support-Vector Classifier*) que predice la condición de la carga cognitiva. La precisión de la clasificación de los datos de prueba fue del 88 % al distinguir entre alta y baja carga cognitiva basándose únicamente en las características vocales, a partir de grabaciones tomadas en entornos variables, condiciones de grabación y ruido de fondo.

Opcionalmente, el sistema puede proporcionar retroalimentación al sujeto, p. ej., como respuesta de voz, como 'correcto', 'incorrecto' o tonos de audio. La lógica interna del módulo de suministro de tarea 1 puede manipular la retroalimentación (proporcionando retroalimentación falsa positiva o negativa) para inducir cambios en el estado mental actual del usuario y, por lo tanto, provocar cambios en la señal de voz.

En consecuencia, el módulo de detección de respuesta 2 de información, puede configurarse para comunicar al sujeto información basándose en dicha puntuación para una tarea presente, antes de comunicar un conjunto de información relacionada con una siguiente tarea.

La figura 2 muestra un ejemplo de un sistema de interfaz de usuario en el que el módulo de detección de respuesta 2 escucha una respuesta. Si no se detecta respuesta, se puede comunicar una indicación verbal y/u otra tarea. Si se detecta una respuesta, los datos de audio se almacenan en una memoria y se transfieren al módulo de reconocimiento del habla 20 para ser puntuados y simultáneamente se transfieren al módulo de interfaz de administración de IA.

Después, se almacena la puntuación, asociada a la respuesta. Si se determina que la puntuación es un valor atípico (p. ej., si se detecta una respuesta pero no pertenece a la tarea), se puede comunicar una indicación verbal para que el usuario responda nuevamente.

- 5 La figura 3 muestra un ejemplo de un sistema de interfaz de usuario en el que la carga cognitiva de las tareas comunicadas por el módulo de suministro de tarea 1 se modula en función del resultado del módulo de reconocimiento del habla 20.

En la presente invención se pueden utilizar los siguientes ejemplos de tareas cognitivas. Sin embargo, esta no es una lista exhaustiva. Algunos tipos de tareas se diseñarán con ensayos cuya carga cognitiva varía de forma gradual/paramétrica, permitiendo establecer comparaciones entre condiciones (es decir, puntuación delta), en cuanto a rendimiento cognitivo y las características de voz asociadas. Otras pruebas están diseñadas específicamente para maximizar u optimizar la extracción de características de voz en diferentes estados mentales y físicos inducidos por tareas.

15 1) Intervalo de dígitos verbales - directo

Se pide al participante que escuche y repita una secuencia de dígitos. La carga cognitiva varía entre condiciones manipulando el número de dígitos dentro de la secuencia de cada ensayo. Por ejemplo, el ensayo 1 (carga cognitiva baja) presentará una secuencia corta de dígitos (p. ej., '7...1...3...4'), mientras que el ensayo 2 (carga cognitiva alta) presentará una secuencia larga de dígitos (p. ej., '9...6...5...8...7...1...3...4'). Ambas condiciones requieren mantener en la memoria a corto plazo una secuencia de dígitos. Sin embargo, el contraste entre las dos condiciones está en la carga de la memoria auditiva a corto plazo.

25 2) Intervalo de dígitos verbales- inverso

Es similar a la tarea de intervalo de dígitos verbales directo, donde se le pide al participante que escuche una secuencia de dígitos. Sin embargo, en esta tarea, al participante se le pide que repita los dígitos en orden inverso. Por ejemplo, cuando el participante escuche "4... 3... 1...7", la respuesta correcta sería "7314". La carga cognitiva puede aumentar nuevamente aumentando el número de dígitos en una secuencia.

Además de la participación de la memoria auditiva a corto plazo, el intervalo de dígitos inverso requiere que los dígitos se manipulen utilizando la memoria de trabajo. Por lo tanto, al comparar el rendimiento del intervalo de dígitos directo e inverso se obtendrá una medida de la carga de la memoria de trabajo.

35 3) Aprendizaje verbal por parejas (PAL verbal)

Al participante se le indica que escuche pares de palabras, se le pregunta por una de ellas y se le pide que responda con la segunda. La carga cognitiva se manipula a través del nivel de similitud semántica entre los pares. Un ejemplo de carga cognitiva baja (similitud semántica alta) puede ser "hierba - verde", mientras que un ejemplo de carga cognitiva alta (similitud semántica baja) puede ser "hierba - ruidoso". El contrato entre los ensayos de carga cognitiva refleja la capacidad de aprendizaje asociativo.

45 4) Aprendizaje verbal sin palabras por parejas (PAL sin palabras)

Esta tarea es una variación del PAL verbal, donde las palabras se emparejan con sin palabras. Se trata de palabras que se pueden pronunciar como en el mundo real pero que no tienen una asociación/significado semántico, por ejemplo "narav". Esta tarea es más difícil que la PAL verbal ya que los participantes no podrán confiar en estrategias semánticas para aprender pares, exigiendo así una mayor capacidad de aprendizaje asociativo.

50 5) Aprendizaje de listas verbales

En esta tarea, se pide a los participantes a escuchar una lista de palabras y recordar tantos elementos de la lista como sea posible. Las mediciones de interés en esta tarea son el número de respuestas correctas, lo que refleja el rendimiento de la memoria, así como características de voz como pausas, respiración, tartamudeo, muletillas ('ehm', 'err', 'pff').

60 6) Repetición de oraciones

Se pide a los participantes que escuchen una oración y la repitan exactamente con las mismas palabras. La carga cognitiva puede manipularse aumentando el número de palabras dentro de cada oración y/o variando la complejidad sintáctica de su estructura. Por ejemplo, "el gato se sienta encima de la mesa" es menos difícil de procesar y recordar (carga cognitiva baja) que "la pelota que está delante del bolígrafo que se ha roto sale rodando" (carga cognitiva alta).

7) Fluidez verbal - categorías semánticas

Se pide a los participantes que nombren tantas palabras como puedan que encajen en una categoría semántica. Por ejemplo, una categoría puede ser "animales" y las respuestas pueden incluir 'gato, perro, rana, elefante, rinoceronte, pájaro' y etcétera. Las mediciones de interés en esta tarea son el número total de respuestas y el número de respuestas correctas que reflejan fluidez verbal, así como características de voz como pausas, respiración, tartamudeo, muletillas.

5

8) Fluidez verbal - fonológica

Se trata de una variación de la tarea de fluidez verbal semántica, en esta tarea se le pide al participante que nombre tantas palabras como sea posible que comiencen con una letra o sonido en particular. Por ejemplo, como palabras que comienzan con la letra 't' se incluyen 'task, test, timid, thin, tree, tame'. Como palabras que comienzan con el sonido 'ff' se incluyen 'fish, phone, fire, fist, pharaoh' y etcétera. Las mediciones de interés en esta tarea son el número total de respuestas y el número de respuestas correctas que reflejan fluidez verbal, así como características de voz como pausas, respiración, tartamudeo, muletillas.

10

15 9) Similitudes

A los participantes se les muestra un par de palabras y se les pide que expliquen en qué se parecen o son similares. La carga cognitiva puede variar manipulando la abstracción de la relación entre las palabras. Por ejemplo, "¿en qué se parecen el 'verde' y el 'azul'?" puede provocar la respuesta "los dos son colores" (carga cognitiva baja). En cambio, "¿en qué se parecen la 'guerra' y la 'paz'?" requiere una respuesta tal como "pueden ser situaciones políticas de un país" (carga cognitiva alta). Tenga en cuenta que "son opuestas" no es una respuesta correcta en este ensayo, ya que la pregunta se refiere a cómo dos palabras son similares, no diferentes. El contrato entre condiciones refleja la capacidad de razonamiento verbal.

20

25 10) Reconocimiento de emociones verbales

En esta tarea, a los participantes se les muestra fragmentos de audio cortos (p. ej., tonos, música, discurso) que tienen una etiqueta emocional (posiblemente adquirida a través de externalización abierta de tareas (*crowdsourcing*) durante el desarrollo de la tarea). Se les pide que seleccionen la categoría emocional que creen que corresponde al estímulo mostrado (p. ej., feliz, triste, enfadado, sorpresa, miedo, repugnancia). Las mediciones de interés en esta tarea incluyen el número de respuestas correctas, así como el sesgo en la clasificación de emociones neutras.

30

11) Fonación mantenida

Se pide a los participantes que produzcan un sonido estable durante un período de tiempo determinado. Por ejemplo, decir 'aaaa' durante 2 segundos. La capacidad de producir y mantener un sonido continuo permite caracterizar rasgos de la calidad de la voz como la fluctuación y la vibración. Se espera que las características de la voz extraídas de esta tarea cambien bajo diferentes condiciones físicas y mentales (p. ej., cansancio, estrés, dolor, sedación, feliz, triste).

35

40 12) Tarea de diadococinesia (pa-ta-ka)

En esta tarea, se pide a los participantes a repetir una sílaba o una combinación de sílabas rápidamente durante un período de tiempo determinado, por ejemplo, repetir la sílaba 'papapa' rápidamente durante 2 segundos (carga cognitiva baja) o repetir 'pataka' rápidamente durante 2 segundos (carga cognitiva alta). Esta tarea evalúa las habilidades motoras orales y es sensible a trastornos del habla como la disartria.

45

13) Tarea de adición en serie auditiva con ritmo (PASAT)

Se muestran dígitos individuales cada 3 segundos y se le indica al participante que añada cada dígito nuevo al inmediatamente anterior. La carga cognitiva puede manipularse variando el intervalo de tiempo entre dígitos (intervalo entre estímulos - IEE). Por ejemplo, un IEE más corto aumenta la dificultad de la tarea, por tanto, mayor carga cognitiva. Los contratos entre condiciones reflejan atención, velocidad de procesamiento de información auditiva y flexibilidad mantenidas.

50

55 14) Resta en serie

Se pide a los participantes que cuenten hacia atrás desde 100 restando un número determinado. La carga cognitiva puede manipularse variando este número determinado que se va a restar. Por ejemplo, contar hacia atrás desde 100 en pasos de 1 o 2 (carga cognitiva baja) es más fácil que contar hacia atrás desde 100 en pasos de 7 o 9 (carga cognitiva alta).

60

15) Secuencias conocidas

En esta tarea se pide a los participantes recordar rápidamente secuencias conocidas de memoria. Por ejemplo, 'nombra los días de la semana lo más rápido que puedas, empezando por el lunes', 'cuenta del 1 al 20 tan rápido como puedas'. La carga cognitiva puede manipularse pidiendo a los participantes que recuerden secuencias conocidas en

65

orden inverso, por ejemplo, 'cuenta hacia atrás de 20 a 1 lo más rápido que puedas', 'nombra los días de la semana al revés, empezando por el domingo, lo más rápido que puedas'. Los contrastos entre condiciones reflejarán dificultades en la velocidad de procesamiento y la capacidad de la memoria de trabajo.

5 16) Administración del cuestionario verbal

Los cuestionarios estandarizados se adaptarán para su uso en la plataforma de voz de esta invención. La mayoría de los cuestionarios existentes en la técnica dependen de que los participantes completen manualmente preguntas abiertas y escalas de calificación (ya sea respondiendo usando las teclas de un ordenador o lápiz y papel). Este método requiere tiempo, los usuarios lo encuentran aburrido y no siempre es apropiado, por ejemplo, para su uso en pacientes con discapacidad visual o personas con problemas de aprendizaje. Esta invención pretende mejorar la técnica existente adaptando dichos cuestionarios a una plataforma de audio utilizando métodos conversacionales e IA para mejorar la experiencia del usuario, manteniendo la validez y fiabilidad clínicas. Las preguntas abiertas de los cuestionarios estandarizados también generarán datos de audio de gran riqueza de los que se extraerán características de voz.

En consecuencia, las tareas cognitivas suministradas por el módulo de suministro de tarea 1 pueden incluir: un intervalo de dígitos verbales directo, un intervalo de dígitos verbales inverso, aprendizaje verbal por parejas, aprendizaje verbal sin palabras por parejas, aprendizaje de listas verbales, repetición de oraciones, fluidez verbal por categoría semántica, fluidez verbal fonológica, reconocimiento de similitud, reconocimiento de emociones verbales, fonación mantenida, diadococinesia, adición en serie auditiva con ritmo, resta en serie, secuencias conocidas y un cuestionario verbal.

El módulo de análisis 3 está configurado para analizar las señales de audio correspondientes a las respectivas respuestas habladas grabadas por el módulo de detección de respuestas 2 para determinar a partir de las respectivas respuestas habladas una o más características indicativas del estado fisiológico del sujeto, comparar dichas características a partir de las respectivas respuestas habladas, y determinar el estado fisiológico del sujeto basándose en dicha comparación. Opcionalmente, el módulo de análisis 3 puede configurarse para determinar el estado fisiológico del sujeto basándose adicionalmente en la puntuación determinada por el módulo de detección de respuesta 2.

El módulo de análisis 3 puede extraer señales relevantes del estado cognitivo y emocional de un sujeto y cambios en dicho estado. Estas señales pueden indexarse en el tiempo para poder asignar la señal de estado al flujo de tareas y, por lo tanto, deducir la relación causal entre la tarea y el estado fisiológico. Como alternativa, la señal puede ser una característica agregada de toda la respuesta de audio del sujeto.

Los datos de audio correspondientes a las respuestas pueden analizarse con respecto a tres tipos generales de características: 1) características paralingüísticas, 2) características prosódicas relacionadas con el tono, 3) características de calidad de la voz. Pueden utilizarse métodos establecidos en la técnica para extraer estas señales, tales como el análisis de coeficientes cepstrales en frecuencias de Mel (MFCC, *mel-frequency cepstral coefficients*), predicción lineal perceptual (PLP) y codificación predictiva lineal (CPL) (Huang, Acero y Hon, 2001). Más específicamente, las siguientes características, así como otras disponibles en programas informáticos de código abierto como openSMILE, se extraerán, normalizarán y utilizarán con el módulo de análisis 3:

- **Paso** es la percepción psicológica de los cambios de frecuencia. Por ejemplo, el aumento de la frecuencia se percibe como un aumento del tono. El paso de un tono complejo (habla) corresponde a la frecuencia fundamental (f_0). El paso también refleja la frecuencia de las vibraciones de las cuerdas vocales durante la producción del habla. Por ejemplo, una pregunta tiene un paso ascendente, mientras que una afirmación o declaración tiene un paso decreciente. También se medirán diversas estadísticas de paso, que corresponden a diferentes características de la señal del habla:

- frecuencia fundamental (f_0)
- f_0 media, DE, intervalo, mediana
- f_0 pendiente (p. ej., ascendente, descendente, plana)

- **Intensidad** es la presión sonora y es una propiedad física de la señal acústica. Es la medida de energía transportada por un sonido, eso lleva a la percepción del volumen. Las estadísticas relacionadas con la intensidad que se extraerán incluyen:

- media, DE
- pendiente
- curvatura

- **Frecuencias formánticas** se determinan a través de la medición de la codificación predictiva lineal (CPL). Estas caracterizan la forma del tracto vocal, que a su vez está determinada por la posición de los articuladores (lengua, labios, mandíbula, paladar blando/velo del paladar).

- **Flujo glotal** es la velocidad del volumen que fluye a través de la glotis y, como tal, es la fuente de excitación del habla sonora. El flujo glotal puede combinarse con un modelo de radiación labial (filtro de paso alto en el dominio de la frecuencia) para formar la derivada del flujo glotal. Los cambios en la calidad de la voz se reflejan en el flujo glotal.

- **Duración del discurso** se mide como longitud en segundos. Esto puede aplicarse a enunciaciones completas, oraciones, palabras o sílabas (a menudo se distinguen entre sílabas acentuadas y átonas).

- **Velocidad del habla:**

- palabras por segundo
- sílabas por segundo
- número de pausas
- duración de las pausas

- 5 • **Calidad de la voz** abarca una amplia variedad de características que se atribuyen a un control imperfecto de las vibraciones de las cuerdas vocales que producen el habla. El efecto percibido produce afonía, respiración entrecortada, crujidos, etc.:
 - fluctuación (irregularidades en el paso)
 - vibración (irregularidades en la intensidad)
 - 10 • relación armónico-ruido (RAR)
 - análisis cepstrales (frecuencia de cambio en la señal de frecuencia)

En consecuencia, las características indicativas del estado fisiológico determinado por el módulo de análisis 3 pueden incluir: tono, intensidad, frecuencias formánticas, flujo glotal, duración del habla, velocidad del habla y calidad de la voz. Se puede determinar un valor que esté asociado a una o a más de estas características. El módulo de análisis 3 puede configurarse para comparar dichas características (p. ej., sus valores) y determinar un cambio en una cualquiera de dichas características entre las respuestas habladas.

Como se muestra en la figura 4, los datos de audio pueden introducirse en un módulo analizador del habla 31. Este módulo incorpora un algoritmo para detectar partes con determinadas duraciones o características dentro del audio completo (p. ej., enunciación completa, oración, palabras, sílabas) y etiquetarlas o segmentarlas. Además, las señales del habla se etiquetarán con metadatos que indican la carga cognitiva y el estado físico o mental.

En consecuencia, el módulo de análisis 3 puede comprender un módulo analizador del habla 31 configurado para detectar partes de una respuesta correspondientes a características del habla y etiquetar y/o segmentar las partes detectadas, incluyendo dichas características del habla enunciaciones, oraciones, palabras y sílabas completas.

Como se muestra en la figura 4, el módulo analizador del habla 31 puede entonces introducir estos segmentos en un selector de características 32, que extrae características acústicas. Estas características se extraen y analizan segundo a segundo y a nivel de expresión completa. Por lo tanto, se obtendrán características tanto locales (segundo a segundo u otra submuestra definida de audio completo) como globales (obtenidas de la enunciación total). Los análisis de estas características incluyen el cálculo de los valores delta para las características procedentes de ensayos con carga cognitiva alta en comparación con los de carga cognitiva baja.

Opcionalmente, como se muestra en la figura 4, las características extraídas por el selector de características 32 pueden pasar a través de un módulo de reducción de características 33. Dependiendo de la naturaleza de la tarea cognitiva realizada y de la manera de extraer las características de voz, determinadas combinaciones de características pueden proporcionar resultados más precisos que otras. Por ejemplo, una tarea de memoria que requiere que una persona recuerde una lista de números producirá a su vez enunciaciones cortas y discontinuas, mientras que una tarea de habla libre produce un flujo de habla largo y continuo. Las características de interés en la primera tarea pueden estar relacionadas con las propiedades acústicas de bajo nivel de cada palabra discontinua enunciada, mientras que las características de interés en la segunda tarea pueden incluir características paralingüísticas tales como velocidad del habla, número de pausas, respiración, tartamudeo, etc. Además, un análisis segundo a segundo de la señal del habla será menos informativo en la primera tarea en comparación con la segunda. Por lo tanto, en lugar de clasificar el habla basándose en todas las características obtenidas, puede ser deseable utilizar un subconjunto de características. Reducir la cantidad total de características que se introducirán en el módulo final de IA también puede aumentar la velocidad del proceso de clasificación y predicción.

En consecuencia, el módulo de análisis 3 puede configurarse para determinar las características indicativas del estado fisiológico a partir de un subconjunto de las características del habla detectadas por el módulo analizador del habla 31, seleccionándose dicho subconjunto de características basándose en las tareas a las que se refiere la información comunicada. Como alternativa o adicionalmente, el módulo de análisis 3 puede configurarse para determinar el estado fisiológico del sujeto basándose en un subconjunto de una o más características, seleccionándose dicho subconjunto basándose en las tareas a las que se refiere la información comunicada. Como alternativa, el módulo de análisis 3 puede configurarse para determinar el estado fisiológico del sujeto basándose en los datos de audio sin procesar completos de la respuesta del sujeto.

Cabe señalar que aunque el analizador del habla, el selector de características y el proceso de reducción de características opcional se describen como elementos distintos, en la práctica, estos elementos pueden implementarse y ejecutarse en el mismo sistema físico (p. ej., servidor, nube).

El módulo de análisis 3 puede entonces combinar puntuaciones cognitivas y/o clínicas generadas por el sistema, características demográficas y cualquier otra información conocida externamente (p. ej., diagnóstico, datos biométricos, datos de imágenes cerebrales), características técnicas de entrada de los dispositivos utilizados para grabar, y la señal de audio sin procesar o la salida del selector de características (es decir, características de audio analizadas). Esto puede suceder en tiempo real o más tarde; en el dispositivo o en la nube.

Dependiendo de los tipos de características extraídas, así como de la cantidad de características extraídas y retenidas, puede utilizarse un clasificador de IA 34. Pueden utilizarse diferentes clasificadores de IA 34, por ejemplo, un clasificador gaussiano, un clasificador del vecino más cercano, puede utilizarse una red neuronal o un modelo disperso de mínimos cuadrados parciales para diferentes conjuntos de características. Como alternativa, si se prefiere un modelo de IA en particular, esto puede guiar al selector de características y a los procesos de reducción de características.

Después, opcionalmente, el resultado del algoritmo se compara con normas o puntuaciones anteriores. En consecuencia, el módulo de análisis 3 está configurado para determinar el estado fisiológico del sujeto basándose adicionalmente en puntuaciones anteriores almacenadas. Como alternativa o adicionalmente, el estado fisiológico determinado del sujeto se puede comparar con un valor de referencia predeterminado.

Opcionalmente, el módulo de análisis 3 puede emitir su determinación como un informe. El resultado de la invención puede informar de cualquiera o de todos los siguientes aspectos:

- El estado fisiológico del participante en ese momento particular. Por ejemplo, en el caso del dolor crónico, si el participante tiene un dolor leve, moderado o intenso.
- El riesgo de enfermedad del participante, cuando el módulo de análisis 3 compara puntuaciones y características con normas.

- Cambio del estado fisiológico o enfermedad (es decir, avance de la enfermedad), cuando el módulo de análisis 3 compara puntuaciones y características con puntuaciones y características anteriores del mismo participante.

- Efecto de un fármaco/dispositivo/intervención, cuando el módulo de análisis 3 compara puntuaciones y características en condiciones externas conocidas.

Los datos de audio pueden procesarse previamente utilizando métodos de normalización para excluir variaciones relacionadas con la edad, género y extraer características de bajo nivel como indicadores de energía, intensidad, paso, formantes, flujo glotal, duración y velocidad del habla, calidad de la voz y forma espectral. Pueden utilizarse técnicas estándar de aprendizaje automático supervisado para entrenar al sistema a reconocer el estado cognitivo o emocional específico a partir de las características de bajo nivel. Un ejemplo de sistema que utiliza este enfoque es el extractor de características de audio openSMILE (el kit de herramientas de código abierto de Munich para la interpretación del habla y la música mediante extracción en grandes espacios) (Eyben *et al.* 2013).

Como alternativa, los datos de audio sin procesar de cada respuesta del sujeto pueden introducirse en un sistema de aprendizaje profundo sin procesamiento previo. Un ejemplo de sistema que utiliza este enfoque es la red neuronal profunda WaveNet de DeepMind, un modelo generativo profundo de formas de onda de audio sin procesar. El aprendizaje profundo normalmente proporciona modelos informáticos que se componen de múltiples capas de procesamiento para aprender representaciones de datos con múltiples niveles de abstracción. El aprendizaje profundo normalmente determina la estructura intrincada en grandes conjuntos de datos (incluidos los datos de audio) utilizando algoritmos de retropropagación para indicar cómo una máquina debe cambiar sus parámetros internos que se utilizan para calcular la representación de los datos en cada capa a partir de la representación de la capa anterior.

Como alternativa, en una red neuronal pueden introducirse datos de audio sin procesar. Un ejemplo de este enfoque es utilizar características de espectrograma de magnitud no procesada introducidas en una red neuronal de convolución profunda. Otro enfoque es introducir la forma de onda de audio sin procesar en una red profunda, en cuyo caso es posible que no sea necesaria una etapa de extracción de características.

Independientemente del enfoque de aprendizaje automático utilizado, el método que se utiliza para obtener los datos de entrenamiento es importante. Los datos de entrenamiento pueden comprender un conjunto de muestras de audio y las correspondientes etiquetas de estado fisiológico.

En un ejemplo, los datos de entrenamiento pueden obtenerse mediante pruebas de los participantes utilizando una variante especial del sistema de interfaz de usuario que se ha configurado para optimizar la calidad de los datos de entrenamiento. Las optimizaciones pueden incluir aumentar la duración de la sesión de prueba, el intervalo dinámico de los parámetros de la tarea, la manipulación de la retroalimentación proporcionada a los sujetos y la selección específica de la batería de pruebas optimizada para una submuestra de la población caracterizada por edad, sexo, estudios, profesión, estado fisiológico y/o patológico.

Las etiquetas del estado fisiológico pueden obtenerse a través de una combinación de cualquiera o todos los siguientes aspectos:

- Inducción del estado fisiológico a través de la valoración de la dificultad de las tareas de la interfaz de usuario
- Inducción del estado fisiológico a través de medios externos tales como provocar dolor, aumentar los distractores (p. ej., ruido), realizar una doble tarea para aumentar la carga cognitiva (p. ej., una tarea postural/de equilibrio) y manipular la retroalimentación que se presenta al participante
- Imputación del estado fisiológico a partir del rendimiento de la tarea del participante

- Estimación del estado fisiológico a través de mediciones fisiológicas tales como el reconocimiento de emociones faciales, la conductancia cutánea y la frecuencia cardíaca, que se sabe que son indicadores indirectos de estrés

- Medición del estado fisiológico a través de datos de imágenes cerebrales (por ejemplo, EEG, RM) de activación del circuito cerebral
- Determinación del estado fisiológico basado en el autoinforme del paciente
- Determinación del estado fisiológico basado en la evaluación o el diagnóstico del médico - Determinación del estado fisiológico basado en registros médicos, en cuestionarios estandarizados y en el autoinforme del paciente

En otro ejemplo, los datos de entrenamiento pueden obtenerse extrayendo datos de corpus de habla existentes de individuos sanos que realizan tareas cognitivas de carga cognitiva variable, así como corpus de habla de muestras de habla de pacientes con cuadros clínicos conocidos. Estas muestras de discurso de pacientes pueden encontrarse en el dominio público (tal como YouTube®) o adquirirse a través de colaboraciones con instituciones académicas y organizaciones sin fines de lucro tales como organizaciones benéficas de investigación o de apoyo al paciente. Las etiquetas de estados fisiológicos en estos conjuntos de datos pueden obtenerse a través de una combinación de cualquiera o de todos los siguientes aspectos:

- Cálculo de características delta en cada individuo
- Etiquetas conocidas determinadas durante la recopilación de datos por los propietarios de las bases de datos
- Etiquetado de muestras de datos a través de una plataforma de externalización abierta de tareas (*crowd-sourcing*)

Como se ilustra en la figura 5, para entrenar un sistema de interfaz de administración de IA, el sistema de IA recibe información de recursos externos, es decir, datos no generados por el módulo de interfaz de usuario. Como ejemplos de dichos datos externos se incluyen, pero sin limitación, muestras de habla de corpus de habla existentes de individuos sanos que realizan tareas cognitivas de carga cognitiva variable, o corpus de habla de muestras de habla de pacientes con cuadros clínicos conocidos.

Los conjuntos de datos de entrenamiento se habrán etiquetado de acuerdo con clases de interés predefinidas. Por ejemplo, en el caso del dolor crónico, estas etiquetas pueden ser 'leve', 'moderado' e 'intenso'. También pueden tener etiquetas adicionales que caractericen la carga cognitiva o emocional, p. ej., 'elevada' o 'alta'.

Dependiendo del tipo de conjuntos de datos de entrenamiento, pueden utilizarse uno o ambos de los siguientes métodos de entrenamiento para ajustar el sistema de clasificación de IA.

En un ejemplo, cuando el conjunto de datos comprende varios puntos de datos dentro de los mismos individuos (paradigma de mediciones repetidas dentro del sujeto), una plataforma digital (ordenador, nube, servidor) recibe muestras de audio y se produce el procesamiento de la señal de habla, que incluyen la extracción de señales de audio en diferentes condiciones/etiquetas en la misma persona (p. ej., diferentes estados de ánimo, diferentes cargas cognitivas, estado de medicación, etc.). Después, un algoritmo de IA combina características demográficas y cualquier otra información conocida externamente, tales como las características técnicas de entrada de los dispositivos utilizados para grabar.

En otro ejemplo, cuando el conjunto de datos se compone de uno o varios puntos de datos en una variedad de personas, algunas de las cuales pueden tener un diagnóstico médico conocido (paradigma entre sujetos), una plataforma digital (ordenador, nube, servidor) recibe muestras de audio y se produce el procesamiento de la señal de habla, que incluye la extracción de señales de audio en diferentes condiciones que son similares en todos los participantes (p. ej., estado de ánimo, cognitivo) y asociación del grupo participante (p. ej., paciente o control). Después, un algoritmo de IA combina características demográficas y cualquier otra información conocida externamente, tales como las características técnicas de entrada de los dispositivos utilizados para grabar.

La presente invención puede proporcionar un sistema y un método para clasificar el estado fisiológico, incluyendo síntomas de trastorno neurológico o trastorno neuropsiquiátrico en un sujeto. Los estados fisiológicos clasificados son preferentemente aquellos que generalmente se obtienen a través de medios subjetivos del sujeto, incluidos, entre otros, dolor, estrés, ansiedad o sedación.

En consecuencia, el estado fisiológico determinado por el módulo de análisis puede estar relacionado con uno o más de: dolor, mareo, estrés, ansiedad, estado de alerta, cansancio o sedación. Por ejemplo, el estado fisiológico puede ser un nivel de dolor que siente el sujeto, un nivel de alerta, cansancio o sedación del sujeto, un nivel de ansiedad o estrés que siente el sujeto.

Como alternativa, o además, el estado fisiológico determinado por el módulo de análisis puede estar relacionado con un trastorno neurológico o neuropsiquiátrico. Por ejemplo, el estado fisiológico puede ser una probabilidad de que el sujeto padezca un trastorno neurológico o neuropsiquiátrico particular.

Como ejemplos no limitativos de trastornos, afecciones o enfermedades neurológicas o neuropsiquiátricas a los que se hace referencia en el presente documento se incluyen, sin limitación, cánceres de cerebro, demencia, deterioro cognitivo leve, epilepsia, enfermedad de Alzheimer, enfermedad de Parkinson, esclerosis múltiple, depresión, esquizofrenia, TDAH, TEPT, trastorno bipolar, trastornos por tics (incluido el síndrome de Tourette), TOC, trastornos de ansiedad (incluidas fobias y trastorno de ansiedad social), trastorno del espectro autista, adicción, trastornos de la

conducta alimentaria, neuropatía, afasia.

Sin embargo, el experto en la materia apreciará que las expresiones enfermedad neurológica y/o trastorno neurológico abarcan más de mil afecciones médicamente reconocidas y, además, que la frontera entre las enfermedades neurológicas y neuropsiquiátricas puede superponerse. La clasificación estadística internacional de enfermedades y problemas relacionados con la salud de la Organización Mundial de la Salud, décima revisión (CIE-10), versión de la OMS; 2016, Capítulos V (trastornos mentales y conductuales) y VI (enfermedades del sistema nervioso) proporcionan respectivamente una lista de dichos trastornos neuropsiquiátricos (Capítulo V del mismo) y trastornos neurológicos (Capítulo VI del mismo).

Como se utiliza en el presente documento, el dolor puede ser un síntoma de otras afecciones subyacentes, ya sean neurológicas, neuropsiquiátrico u otras, o puede ser una afección de dolor crónico. La afección de dolor crónico puede deberse a causas conocidas o sospechosas, tales como artritis, fibromialgia, lumbalgia, migrañas, otros problemas musculoesqueléticos, diabetes, daños en nervio, enfermedad de Crohn, síndrome de cansancio crónico, síndrome del intestino irritable o cáncer. Los resultados de la invención pueden adoptar la forma de una puntuación o un informe de riesgo.

A continuación se describirán varias realizaciones de la presente invención para ilustrar ejemplos de usos de los sistemas y métodos descritos anteriormente. Estos ejemplos pretenden demostrar la gama de posibles usos que pueden hacerse y no deben considerarse en modo alguno limitativos de los usos que podrían hacerse de la presente invención. Los usos expuestos en los ejemplos siguientes pueden modificarse para adaptarse a las necesidades específicas de un usuario o afección neurológica en particular.

Ejemplo 1 - monitorización remota por teléfono, móvil o web

En una realización de la invención, la invención está integrada en un sistema remoto utilizado para monitorizar a un paciente con una afección neurológica o neuropsiquiátrica. Esto beneficiará a los pacientes que viven lejos de un puesto de control médico, que pueden estar en lista de espera para recibir tratamiento, que no necesiten evaluaciones frecuentes en persona, o que no tengan la capacidad física o mental para desplazarse a efectos de realizar evaluaciones en persona. Por ejemplo, los pacientes con depresión y escasa alfabetización faltan con frecuencia a las citas médicas (Miller-Matero *et al.*, 2016). Se ha comprobado que los sistemas de recordatorio, tales como las llamadas telefónicas previas a las citas, el correo electrónico y los historiales médicos electrónicos basados en la web, aumentan eficazmente el cumplimiento de las citas médicas, el diagnóstico y el tratamiento (p. ej., Liu *et al.*, 2014; Gurol-Urganci *et al.*, 2013).

El paciente recibirá una llamada telefónica, un correo electrónico o una notificación en un dispositivo móvil a intervalos de tiempo regulares, determinados y preestablecidos en función de los historiales médicos o solicitados por el equipo asistencial del paciente. La llamada, el correo electrónico y la notificación en el dispositivo móvil se relacionarán con una realización de la invención alojada en un servidor basado en la nube. El módulo de interfaz de usuario presentará al paciente un conjunto de pruebas cognitivas, que pueden seleccionarse específicamente para minimizar las limitaciones de tiempo y maximizar las características de voz clínicamente informativas. Las respuestas del paciente se registrarán y procesarán a través del módulo de habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA envía una respuesta al paciente en la que reconoce que ha finalizado la monitorización, le notifica la siguiente sesión y le proporciona información opcional sobre su rendimiento. La IA envía simultáneamente a un equipo asistencial designado, o a los expedientes clínicos, un resumen del rendimiento, del estado fisiológico y de la puntuación de la enfermedad según los cálculos de la IA.

Esta realización preferida mejorará significativamente la monitorización del avance de la enfermedad; reducirá el tiempo que el médico dedica a realizar pruebas de monitorización rutinarias, permitiéndole disponer de más tiempo para realizar intervenciones; mejorará el cumplimiento de las citas médicas/hospitalarias; y proporcionará un modelo de atención médica centrado en el paciente. El valor añadido de la presente invención sobre la técnica anterior es que la modulación activa de la carga cognitiva en el módulo de suministro de tareas en la interfaz de usuario impedirá que los pacientes manipulen sus informes sobre síntomas o enfermedades.

Ejemplo 2 - sistema de repetición de la prescripción

En otra realización preferida de la invención, un sistema telefónico o basado en la web incorpora la presente invención y se utiliza para decidir si un paciente reúne los requisitos para una repetición de la prescripción. Por ejemplo, los pacientes con dolor crónico a menudo requieren fármacos opiáceos, donde la dependencia física y la adicción son efectos secundarios comunes. Estos casos requieren una estrecha monitorización de la prescripción y de los síntomas del dolor. En la actualidad, no existe ninguna medición objetiva del dolor en la técnica anterior y las evaluaciones de repetición de la prescripción se basan en los autoinformes del paciente sobre su dolor. Este autoinforme subjetivo puede exagerarse para sustentar una adicción.

Cuando un paciente solicita por teléfono o internet (p. ej., a través de su historial médico electrónico) una repetición de la prescripción, el módulo de interfaz de usuario presentará al paciente un conjunto de pruebas cognitivas,

específicamente seleccionadas, que son sensibles al dolor. Al variar la carga cognitiva de la tarea y, por tanto, aumentar el esfuerzo mental y el compromiso necesarios para realizar con éxito la tarea, el sistema hace que sea difícil o imposible que un paciente finja señales de dolor con la voz. Las respuestas del paciente se registrarán y procesarán a través del módulo del habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA envía una respuesta al paciente en la que reconoce que ha finalizado la evaluación. La IA genera simultáneamente un resultado de puntuación de dolor de ese paciente para el farmacéutico y/o equipo asistencial designado y/o expedientes clínicos. La decisión final de repetir la prescripción recae en el equipo asistencial.

10 **Ejemplo 3: Alta postoperatoria de atención hospitalaria**

En otra realización preferida, la presente invención puede servir como una herramienta de cabecera de paciente hospitalizado para determinar cuándo un paciente está preparado para ser dado de alta del hospital después de una cirugía. Por ejemplo, es posible que los pacientes deseen volver a casa lo antes posible tras la intervención, pero es posible que no estén preparados para hacerlo si padecen efectos tales como sedación, mareos y dolor, que podrían ocultar a propósito para ser dados de alta. Como alternativa, puede haber situaciones en las que un paciente no quiera que le den el alta y pueda exagerar los síntomas a propósito.

La invención puede presentarse a través de un pequeño dispositivo en la mesilla de noche del paciente, tal como un teléfono inteligente, Amazon Echo® o Google Home®. El paciente será sometido a pruebas antes de la intervención quirúrgica y en varios intervalos después de la misma. El módulo de presentación de tareas de interfaz de usuario presentará un conjunto de pruebas cognitivas, seleccionadas específicamente por su sensibilidad a los efectos de la sedación, mareos y dolor. La carga cognitiva de las pruebas variará hasta alcanzar un umbral en el que el paciente realice la tarea con éxito. Las respuestas del paciente se registrarán y procesarán a través del módulo del habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA envía una respuesta al paciente en la que reconoce que ha finalizado la evaluación y si está preparado para recibir el alta. La IA genera simultáneamente un resultado sobre el resumen del rendimiento al equipo asistencial designado, incluyendo la comparación del umbral del paciente con otros y con su propio umbral preoperatorio.

En consecuencia, el módulo de análisis 3 está configurado para determinar si el sujeto es apto o no para recibir el alta hospitalaria después del tratamiento de un cuadro clínico.

Ejemplo 4: Monitorización de los efectos de la intervención

Otra realización de la invención incorpora el sistema descrito en la misma como una evaluación objetiva de la eficacia de intervenciones no basadas en medicamentos, tales como fisioterapia, psicoterapia o aplicación de salud digital. Actualmente, la eficacia de este tipo de intervenciones se evalúa utilizando mediciones autoinformadas de dolor, estado de ánimo y calidad de vida subjetivas. Estas mediciones son sensibles a las características específicas de cada persona, como la motivación, afecto personal hacia el terapeuta y el efecto placebo. La presente invención proporciona un método objetivo para medir la eficacia de la intervención comparando el rendimiento y las características de la voz del participante antes y después de la intervención (y los puntos de ajuste intermedios).

El módulo de suministro de tarea de la interfaz de usuario presentará un conjunto de pruebas cognitivas, específicamente seleccionadas por su sensibilidad a los efectos del dolor. La carga cognitiva de las pruebas variará y, por tanto, el esfuerzo mental necesario para completar las tareas con éxito. Este método maximiza las características de la voz asociadas al dolor y al esfuerzo mental. Las respuestas del paciente se registrarán y procesarán a través del módulo del habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA envía una respuesta al paciente en la que reconoce que ha finalizado la evaluación y si está preparado para recibir el alta. La IA genera simultáneamente un resultado al terapeuta sobre un resumen del rendimiento, incluyendo cómo se compara el umbral del paciente con sesiones anteriores. Los terapeutas pueden encontrar útil esta información para personalizar mejor sus estrategias de intervención.

Ejemplo 5 - control de seguridad

En una realización preferida adicional de la presente invención, el sistema se utiliza como sistema de control de seguridad para su uso en relación con personas que trabajan en profesiones de alto riesgo en particular, tales como control del tráfico aéreo, pilotos, cirujanos, operarios de maquinaria pesada, o en relación con dispositivos y vehículos que requieren un nivel de estado de alerta para conducir, como un auto, tranvía, tren.

Antes de llevar a cabo procedimientos de alto riesgo y maquinaria pesada, se pide a la persona que realice una serie de tareas cognitivas breves, específicamente seleccionadas, que son sensibles para detectar cambios en el estado de alerta y la sedación. Estas tareas se presentarán variando la carga cognitiva hasta que se alcance un umbral en el que la persona realice la tarea con éxito. Las respuestas del individuo se grabarán y procesarán a través del módulo del habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA realiza cálculos que comparan el umbral cognitivo de la persona y las características de la voz relacionadas con las puntuaciones delta de rendimiento cognitivo con umbrales predeterminados que se consideran

aceptables para realizar un procedimiento o profesión de alto riesgo. La IA informará a la persona si se la considera apta para continuar. La IA genera simultáneamente un resultado a la empresa o al supervisor si el umbral de la persona está por debajo del umbral predeterminado.

- 5 En consecuencia, el módulo de análisis 3 puede configurarse para determinar si el sujeto es o no apto para realizar una actividad de alto riesgo, en donde la actividad de alto riesgo se selecciona de un grupo que comprende: control de tráfico aéreo, pilotar un avión, realizar cirugía, trabajar con maquinaria pesada, conducir un coche, un tranvía o un tren.

10 **Ejemplo 6 - uso general del consumidor**

Otra realización prevista de la presente invención es la integración de esta invención en un sistema de autoayuda centrado en la persona. Por ejemplo, una persona utiliza una aplicación en un dispositivo portátil (teléfono inteligente, tableta, reloj electrónico, o incluso el teléfono) que utiliza la presente invención para monitorizar el estado cognitivo, el avance de una enfermedad o la medicación a lo largo del tiempo. El resultado de la interfaz de administración de la IA (p. ej., puntuación de riesgo, cambio del cuadro clínico, efecto de la medicación) podría configurarse para relacionarlo con un sistema de 'alerta' o "realimentación" que envíe comentarios al usuario (con fines de auto-monitorización) así como a otras partes especificadas por el usuario. Estos pueden ser su equipo asistencial, personas responsables de su cuidado, o relacionadas con otras aplicaciones en la técnica que proporcionen intervención y asesoramiento. Por ejemplo, si la presente invención establece que el usuario está sintiendo altos niveles de dolor que impactan negativamente en su función cognitiva, esto podría 1) alertar al usuario para que considere tomar medicamentos adicionales (si están dentro del alcance de su plan de atención), 2) ofrecer asesoramiento sobre estrategias de autogestión como la meditación, afrontamiento cognitivo y/o 3) registrar esto como una entrada temporal en un 'diario del dolor' para el equipo de atención médica (cuando corresponda).

Adicionalmente, la IA integrada en la presente invención se puede desarrollar aún más para aprender 'activamente' a seleccionar qué tareas posteriores presentar al usuario en función de las respuestas dadas en los cuestionarios de selección, datos demográficos y/o tareas anteriores. Por ejemplo, un usuario de un teléfono inteligente puede acceder al módulo de interfaz de usuario a través de una aplicación (app) sanitaria. Después, el módulo de interfaz de usuario solicita al usuario determinada información demográfica (u obtiene esta información de la API (*application programming interface*, interfaz de programación de aplicaciones) de la app sanitaria). Esta información se registra y analiza en la IA, que calcula el riesgo de determinados síntomas comparando la información del usuario con una base de datos de otros datos demográficos del usuario. Después, la IA selecciona un conjunto de tareas que son las más sensibles para detectar síntomas de enfermedades que el usuario tiene un alto riesgo de padecer. Esto se presenta al usuario a través del módulo de interfaz de usuario. Las respuestas de los usuarios se grabarán y procesarán a través del módulo del habla (análisis, extracción de características, normalización de características) e interfaz de administración de IA. Después, la IA realiza cálculos que comparan las características de la voz de la persona relacionadas con las puntuaciones delta de rendimiento cognitivo con su rendimiento anterior. La IA alertará al usuario si se deben considerar otras medidas (como medicación o estrategias de autoayuda). La IA genera simultáneamente un resultado a un cuidador(a) designado(a).

Ejemplo 7 - uso de ensayos clínicos para evaluar la seguridad o eficacia de un tratamiento

Otra aplicación de la invención son los ensayos clínicos para evaluar la seguridad o eficacia de un tratamiento. Los ensayos clínicos están diseñados en torno a criterios de valoración objetivos para determinar la seguridad y/o eficacia de un tratamiento. En la actualidad, la evaluación de la función cognitiva puede verse influenciada por resultados subjetivos (por ejemplo, dolor, calidad de vida), motivación del sujeto o efectos del entrenamiento. La invención se puede aplicar para proporcionar una determinación objetiva mejorada del impacto de un tratamiento sobre la función cognitiva para evaluar la toxicidad o eficacia de los tratamientos en ensayos clínicos. La invención proporciona un medio para mejorar criterios de valoración objetivos relacionados con síntomas autodiagnosticados (como dolor o ansiedad) o efectos sobre la función cognitiva (como la memoria o la capacidad de aprendizaje).

En una realización preferida, la invención se aplica en un ensayo clínico para determinar la eficacia de un tratamiento para el dolor.

En otra realización preferida, la invención se aplica para determinar si un tratamiento produce reacciones adversas con respecto a la función cognitiva de los pacientes. La función cognitiva a evaluar en los ensayos clínicos puede ser la memoria auditiva, memoria de trabajo, capacidad de aprendizaje asociativo, fluidez verbal, razonamiento verbal, reconocimiento de emociones, habilidades motoras orales, la velocidad de procesamiento y los cambios asociados en estos dominios de la función cognitiva se obtienen de las características de la voz extraídas de las sesiones de prueba de los pacientes.

En determinadas realizaciones, se recogen muestras de voz de pacientes utilizando un micrófono en un sistema telefónico, que puede ser un teléfono inteligente, permitiendo así monitorizar a los pacientes fuera de la clínica.

El módulo de análisis 3 puede comparar el estado fisiológico del sujeto con un valor de referencia que incluya datos

de población específicos para un trastorno neurológico, en donde el trastorno se selecciona de un grupo de trastornos que comprende: dolor, cáncer cerebral, demencia, deterioro cognitivo leve, epilepsia, enfermedad de Alzheimer, enfermedad de Parkinson, esclerosis múltiple, depresión, esquizofrenia, TDAH, TEPT, trastorno bipolar, trastornos por tics (incluido el síndrome de Tourette), TOC, trastornos de ansiedad (incluidas fobias y trastorno de ansiedad social), trastorno del espectro autista, adicción, trastornos de la conducta alimentaria, neuropatía, afasia. El módulo de análisis 3 puede configurarse para determinar la seguridad o eficacia de un tratamiento para el dolor, cáncer cerebral, demencia, deterioro cognitivo leve, epilepsia, enfermedad de Alzheimer, enfermedad de Parkinson, esclerosis múltiple, depresión, esquizofrenia, TDAH, TEPT, trastorno bipolar, trastornos por tics (incluido el síndrome de Tourette), TOC, trastornos de ansiedad (incluidas fobias y trastorno de ansiedad social), trastorno del espectro autista, adicción, trastornos de la conducta alimentaria, neuropatía, afasia.

La presente invención puede proporcionar una medida del éxito de una intervención clínica. Después de realizar una primera evaluación del estado fisiológico del sujeto utilizando los sistemas descritos anteriormente, un médico puede realizar una intervención (por ejemplo, cambiar la dosis del medicamento o proporcionar fisioterapia o cualquier otra intervención descrita anteriormente). Después de un período de tiempo apropiado (dependiendo de la intervención) se realiza una segunda evaluación del estado fisiológico del sujeto utilizando los sistemas descritos anteriormente. En consecuencia, la invención permite determinar un cambio en el estado fisiológico. Por tanto, la invención puede proporcionar una medida del éxito de una intervención clínica.

Referencias

Andersson, G. *et al.* (2002) 'Effect of cognitive load on postural control', Brain Research Bulletin. Elsevier, 58(1), págs. 135-139. doi: 10.1016/S0361-9230(02)00770-0.

Eyben, F., Weninger, F., Gross, F., y Schuller, B. (2013) Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor, In Proc. ACM Multimedia (MM), Barcelona, España, ACM, ISBN 978-1-4503-2404-5, págs. 835-838. doi: 10.1145/2502081.2502224

Fusaroli, R. *et al.* (2017) Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis, Autism Research, 10(3), págs. 384-407. doi: 10.1002/aur.1678.

Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. (2013) Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database of Systematic Reviews 2013, Issue 12. Art. No.: CD007458.DOI: 10.1002/14651858.CD007458.pub3 Huang, X., Acero, A. y Hon, H. (2001) Spoken Language Processing: A guide to theory, algorithm, and system development. Prentice Hall.

Johnstone T (2001) Tesis doctoral. Universidad de Australia Occidental. Con acceso a través de <http://brainimaging.waisman.wise.edu/~tjohnstone/Thesis.pdf>

Lautenbacher, S. *et al.* (2017) 'Phonetic Characteristics of Vocalizations during Pain', Pain Reports, 2(e597), págs. 1-5.

Marquard, C. *et al.* (2017) 'Speak, Think, Act: A phonetic analysis of the combinatorial effects of respiratory mask, physical and cognitive stress on phonation and articulation', MSc Thesis.

Con acceso a través de https://www.researchgate.net/publication/315892441_Speak_Think_Act_A_phonetic_analysis_of_the_combinatorial_effects_of_respiratory_mask_physical_and_cognitive_stress_on_phonation_and_articulation

Miller-Matero, LR, Clark, KB, Brescacin, C, Dubaybo, H y Willens, DE (2016) Depression and literacy are important factors for missed appointments, Psychology, Health & Medicine, 21:6, 686-695, DOI: 10.1080/13548506.2015.1120329

Nevler, N. *et al.* (2017) 'Automatic measurement of prosody in behavioral variant FTD', págs. 1-8. Oshrat, Y. (2014) The Fingerprints of Pain in Human Voice (Thesis, The Open University of Israel Computer Science Division)

Place et al (2017) Behavioral Indicators on a Mobile Sensing Platform Predict Clinically Validated Psychiatric Symptoms of Mood and Anxiety Disorders. J Med Internet Res. 19(3):e75 doi: 10.2196/jmir.6678

Tsai, F. S. *et al.* (2016) 'Toward development and evaluation of pain level-rating scale for emergency triage based on vocal characteristics and facial expressions', Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 8-12-NaN-2016, págs. 92-96. doi: 10.21437/Interspeech.2016-408.

Vogel, A. P., Fletcher, J. y Maruff, P. (2010) 'Acoustic analysis of the effects of sustained wakefulness on speech', The Journal of the Acoustical Society of America. Acoustical Society of America, 128(6), págs. 3747-3756. doi: 10.1121/1.3506349.

Zhang, H.-H. *et al.* (2016) 'Classification of Parkinson's disease utilizing multi-edit nearest-neighbor and ensemble learning algorithms with speech samples.', Biomedical engineering online. BioMed Central, 15(1), pág. 122. doi: 10.1186/s12938-016-0242-6.

REIVINDICACIONES

1. Un sistema para evaluar el estado fisiológico de un sujeto, que comprende:
un módulo de suministro de tarea (1) configurado para comunicar a un sujeto al menos dos conjuntos de información,
5 relacionándose cada conjunto de información con una tarea cognitiva que requiere una respuesta hablada del sujeto;
un módulo de detección de respuesta (2) configurado para grabar las respectivas respuestas habladas del sujeto como
una señal de audio, comprendiendo el módulo de detección de respuesta (2) un micrófono;
un módulo de análisis (3) configurado para analizar las señales de audio correspondientes a las respectivas respuestas
10 respuestas habladas grabadas por el módulo de detección de respuestas (2) para determinar, a partir de las respectivas
características a partir de las respectivas respuestas habladas, y determinar el estado fisiológico del sujeto basándose
en dicha comparación;
caracterizado por que los al menos dos conjuntos de información son conjuntos de información diferentes,
seleccionándose los al menos dos conjuntos de información diferentes de manera que cada tarea cognitiva relacionada
15 con cada conjunto de información diferente tenga una carga cognitiva asociada diferente.
2. El sistema de la reivindicación 1, en donde el estado fisiológico del sujeto comprende: un nivel de dolor que siente
el sujeto, y/o un nivel de estado de alerta, cansancio o sedación del sujeto, y/o un nivel de estrés o ansiedad que siente
20 el sujeto, y/o una probabilidad de que el sujeto padezca un trastorno neurológico o neuropsicológico.
3. El sistema de la reivindicación 1, en donde los conjuntos de información comunicados se seleccionan de diferentes
grupos de conjuntos de información prealmacenados, estando dichos conjuntos de información prealmacenados
agrupados de acuerdo con una carga cognitiva asociada a la tarea con la que se relaciona cada conjunto de
información.
25
4. El sistema de cualquier reivindicación anterior, en donde los conjuntos de información comunicados se seleccionan
de diferentes grupos de conjuntos de información prealmacenados, estando dichos conjuntos de información
prealmacenados agrupados de acuerdo con un estado físico o mental inducido por la tarea con la que se relaciona
cada conjunto de información.
30
5. El sistema de cualquier reivindicación anterior, en donde el módulo de detección de respuesta (2) comprende un
módulo de reconocimiento del habla (20) configurado para analizar cada respuesta hablada, el módulo de detección
de respuesta (2) está configurado para comparar el resultado del módulo de detección del habla (20) con una respuesta
esperada prealmacenada, y asignar una puntuación a cada respuesta hablada basándose en dicha comparación,
35 opcionalmente en donde el módulo de análisis está configurado para determinar el estado fisiológico del sujeto
basándose adicionalmente en dicha puntuación, opcionalmente en donde el módulo de análisis (3) está configurado
para determinar el estado fisiológico del sujeto basándose adicionalmente en puntuaciones anteriores almacenadas.
6. El sistema de una cualquiera de las reivindicaciones 5, en donde el módulo de suministro de tareas (1) selecciona
40 un conjunto de información relacionada con una siguiente tarea basándose en la puntuación asociada a una respuesta
a una tarea anterior.
7. El sistema de la reivindicación 6, en donde, si la puntuación de la tarea anterior es inferior a una puntuación umbral
predeterminada, la siguiente tarea se selecciona de manera que tenga una carga cognitiva asociada inferior a la carga
45 cognitiva asociada a la tarea anterior.
8. El sistema de la reivindicación 6 o 7, en donde, si la puntuación de la tarea anterior es superior a una puntuación
umbral predeterminada, la siguiente tarea se selecciona de manera que tenga una carga cognitiva asociada superior
a la carga cognitiva asociada a la tarea anterior.
50
9. El sistema de la reivindicación 6 o 7 u 8, en donde, si se determina que la puntuación de la tarea anterior es un valor
atípico, la siguiente tarea se selecciona de manera que tenga una carga cognitiva asociada igual o similar a la carga
cognitiva asociada a la tarea anterior.
- 55 10. El sistema de una cualquiera de las reivindicaciones 5 a 9, en donde el módulo de detección de respuesta (2) está
configurado para comunicar al sujeto información basándose en dicha puntuación para una tarea presente, antes de
comunicar un conjunto de información relacionada con una siguiente tarea.
11. El sistema de cualquier reivindicación anterior, en donde las características indicativas del estado fisiológico
60 incluyen: paso, intensidad, frecuencias formánticas, flujo glotal, duración del habla, velocidad del habla y calidad de la
voz, opcionalmente en donde el módulo de análisis está configurado para comparar dichas características y determinar
un cambio en una cualquiera de dichas características entre las respuestas habladas.
12. El sistema de la reivindicación 11, en donde el módulo de análisis (3) está configurado para determinar el estado
65 fisiológico del sujeto basándose en un subconjunto de la una o más características, seleccionándose dicho
subconjunto basándose en las tareas a las que se refiere la información comunicada.

- 5 13. El sistema de cualquier reivindicación anterior, en donde el módulo de análisis (3) comprende un módulo analizador del habla (31) configurado para detectar partes de una respuesta correspondientes a características del habla y etiquetar y/o segmentar las partes detectadas, incluyendo dichas características del habla, enunciaciones, oraciones, palabras y sílabas completas, opcionalmente en donde el módulo de análisis (3) está configurado para determinar las características indicativas del estado fisiológico a partir de un subconjunto de las características del habla detectadas por el módulo analizador del habla (31), seleccionándose dicho subconjunto de características basándose en las tareas a las que se refiere la información comunicada.
- 10 14. El sistema de cualquier reivindicación anterior, en donde las tareas cognitivas incluyen: un intervalo de dígitos verbales directo, un intervalo de dígitos verbales inverso, aprendizaje verbal por parejas, aprendizaje verbal sin palabras por parejas, aprendizaje de listas verbales, repetición de oraciones, fluidez verbal por categoría semántica, fluidez verbal fonológica, reconocimiento de similitud, reconocimiento de emociones verbales, fonación mantenida, diadococinesia, adición en serie auditiva con ritmo, resta en serie, secuencias familiares o un cuestionario verbal.
- 15 15. Un método para evaluar el estado fisiológico de un sujeto, que comprende: comunicar a un sujeto al menos dos conjuntos de información, relacionándose cada conjunto de información con una tarea que requiere una respuesta hablada del sujeto;
- 20 grabar las respectivas respuestas habladas del sujeto como una señal de audio usando un micrófono;
- analizar las señales de audio correspondientes a las respectivas respuestas habladas grabadas para determinar, a partir de las respectivas respuestas habladas, una o más características indicativas del estado fisiológico del sujeto, comparar dichas características a partir de las respectivas respuestas habladas, y determinar el estado fisiológico del sujeto basándose en dicha comparación;
- 25 caracterizado por que los al menos dos conjuntos de información son conjuntos de información diferentes, seleccionándose los al menos dos conjuntos de información diferentes de manera que cada tarea cognitiva relacionada con cada conjunto de información diferente tenga una carga cognitiva asociada diferente.

Fig. 1

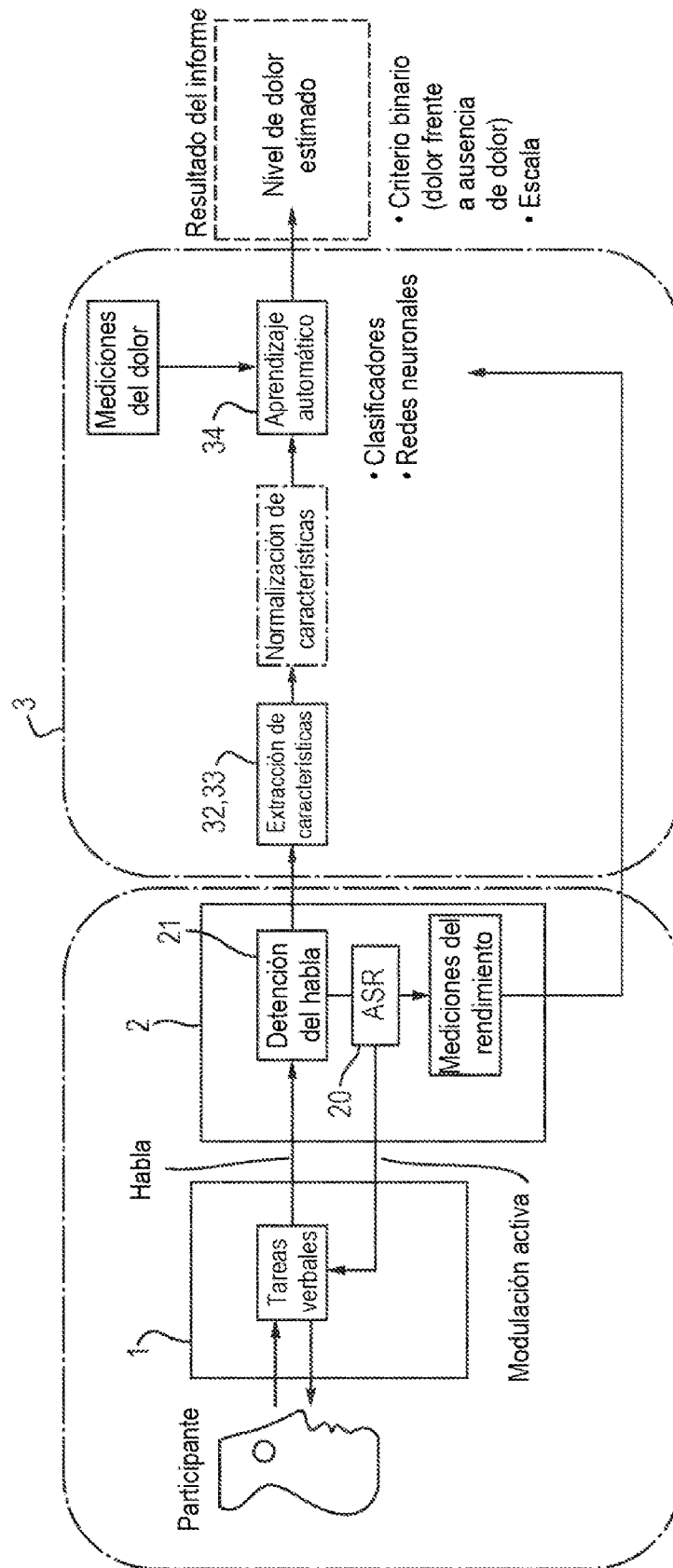


Fig. 2

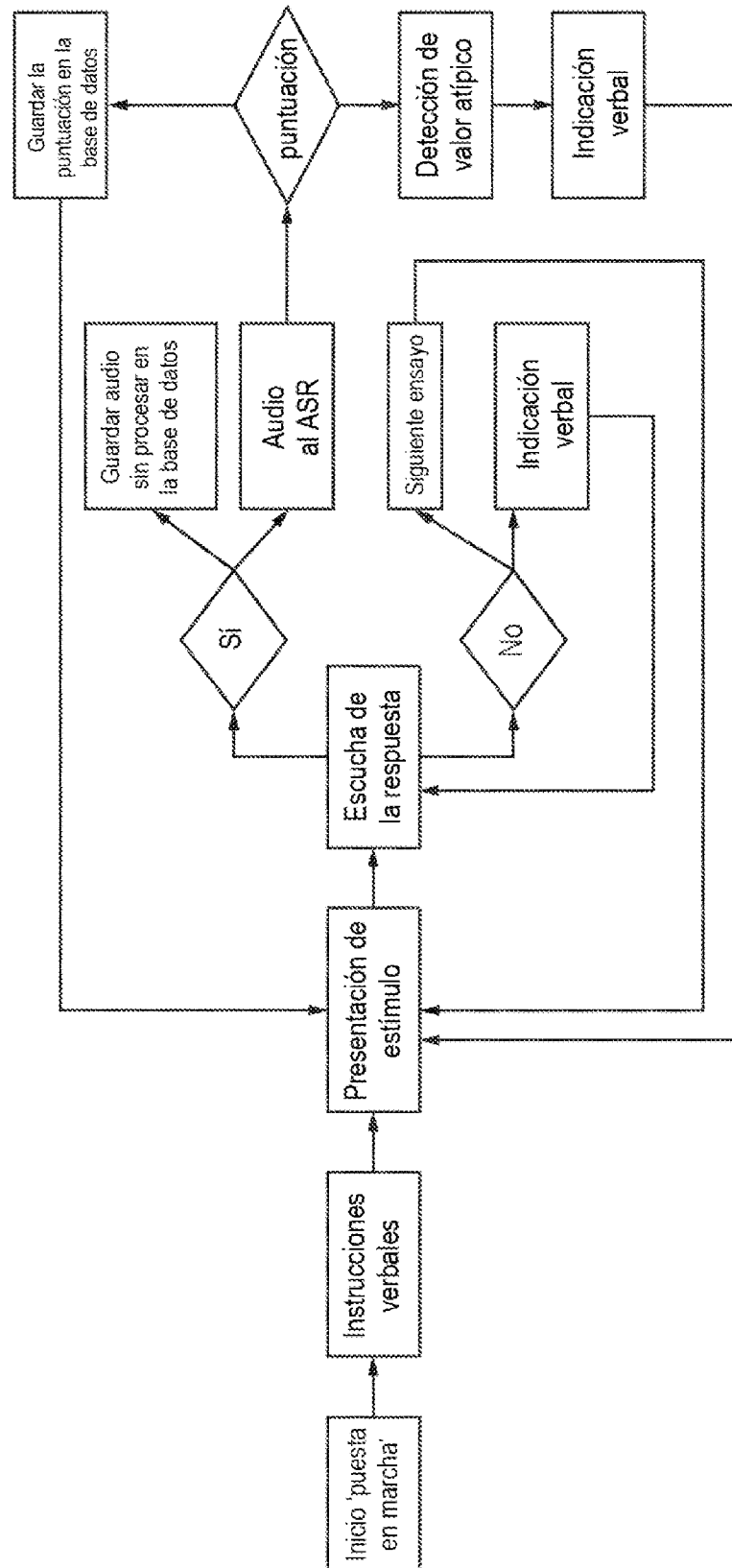


Fig. 3

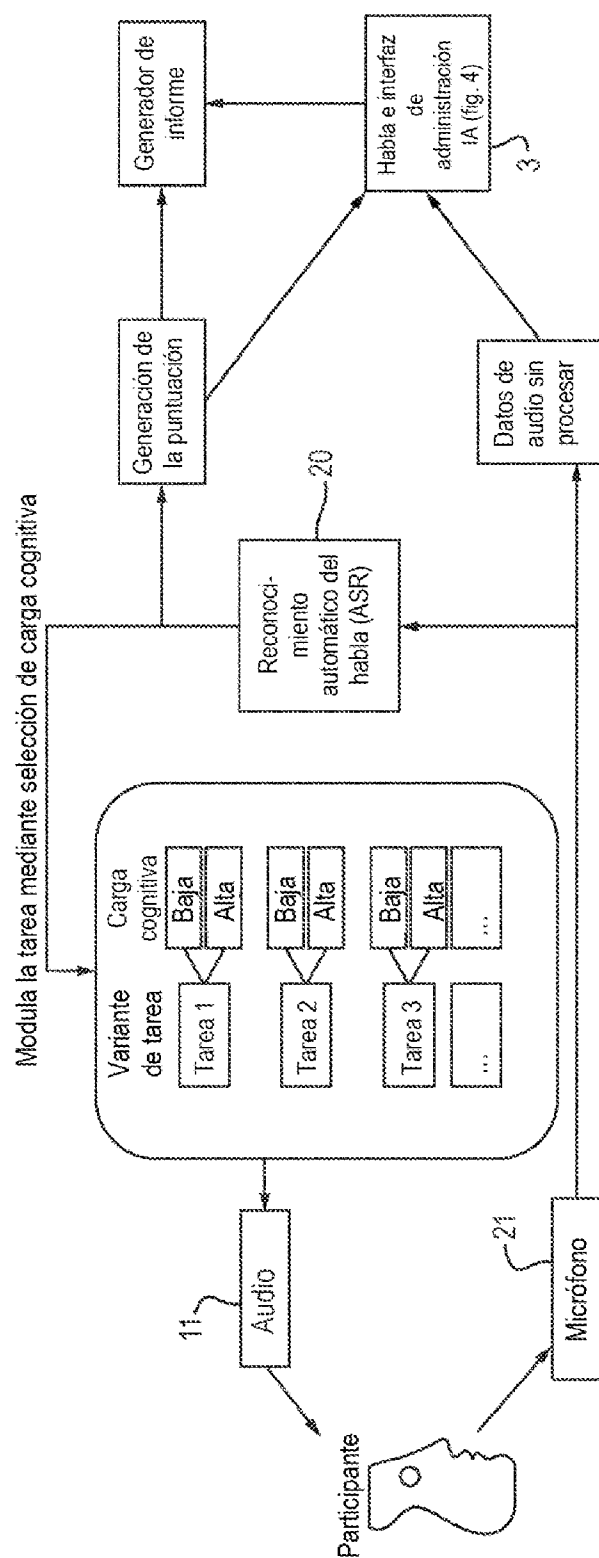


Fig. 4

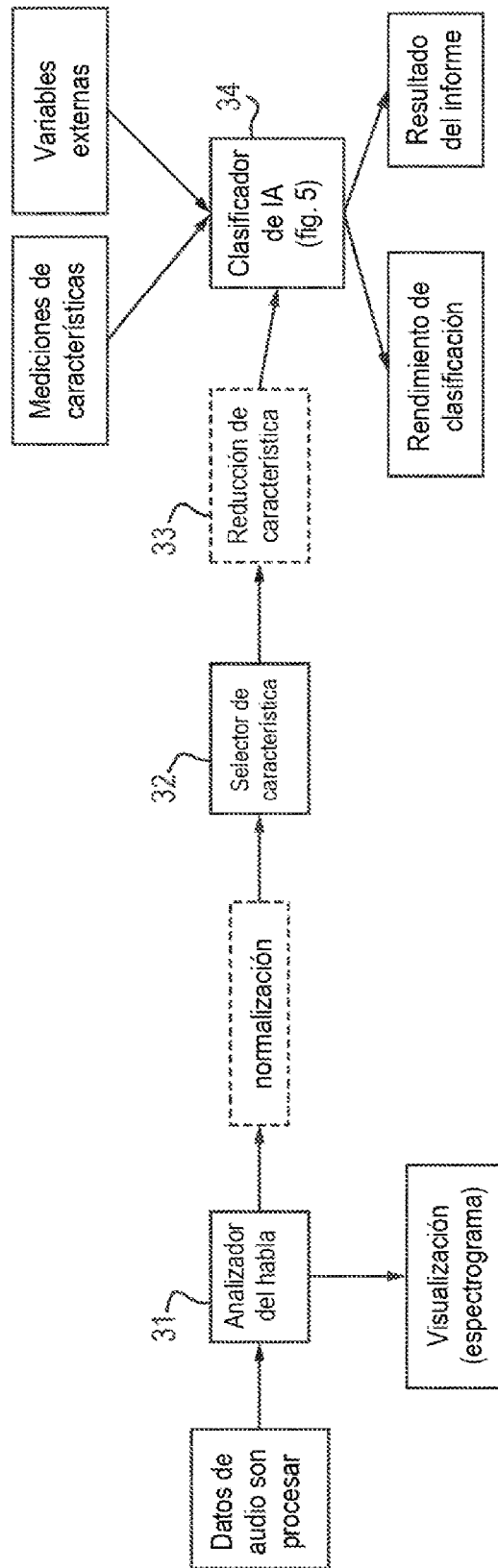


Fig. 5

